

Washington University in St. Louis  
**Washington University Open Scholarship**

---

All Theses and Dissertations (ETDs)

---

January 2011

# Integrity, Identity, and Why Moral Exemplars Do What Is Right

David Bauman

*Washington University in St. Louis*

Follow this and additional works at: <http://openscholarship.wustl.edu/etd>

---

## Recommended Citation

Bauman, David, "Integrity, Identity, and Why Moral Exemplars Do What Is Right" (2011). *All Theses and Dissertations (ETDs)*. 34.  
<http://openscholarship.wustl.edu/etd/34>

This Dissertation is brought to you for free and open access by Washington University Open Scholarship. It has been accepted for inclusion in All Theses and Dissertations (ETDs) by an authorized administrator of Washington University Open Scholarship. For more information, please contact [digital@wumail.wustl.edu](mailto:digital@wumail.wustl.edu).

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Philosophy

Dissertation Examination Committee

Christopher Heath Wellman, Chair

Eric Brown

Kurt Dirks

John Doris

Larry May

Simine Vazire

INTEGRITY, IDENTITY, AND WHY MORAL EXEMPLARS DO WHAT IS RIGHT

by

David Bauman

A dissertation presented to the  
Graduate School of Arts and Sciences of  
Washington University in  
partial fulfillment of the  
requirements for the degree  
of Doctor of Philosophy

May 2011

Saint Louis, Missouri

Copyright by  
David Curt Bauman  
2011

## Abstract

Many people think that integrity is a central concept for moral reasoning. Political, educational, and business leaders tout the importance of integrity for our society's moral health. But there are reasons to doubt that it is solely a moral concept. Our intuitions seem to confirm that a committed Mafia boss may have some form of integrity. Or one might say that integrity is the mere expression of other moral commitments or depict it as a formal virtue lacking any moral content. Others question whether or not it is even desirable or achievable. In this dissertation, I develop an account of integrity that defends integrity from these doubts and present it as a central moral concept.

In Chapter 1, I distill a basic notion of integrity from existing discussions to answer the question, "What is integrity?". My analysis provides a "bare bones" account that captures the central features of integrity without addressing its moral or immoral content. The identity account of integrity I present requires a person of integrity to have coherent commitments that constitute her identity as well as the determination to actualize them.

In Chapter 2 I answer the question, "Is integrity a moral concept?" I start by reviewing recent attempts to resolve a dilemma facing all integrity accounts. The first horn of the dilemma is the *intuition problem* which is that it seems that both tyrants and heroes could have some kind of integrity. The second horn is the *moral integrity problem* which is that integrity is commonly regarded as a moral concept that cannot apply to

tyrants. Any optimal account of integrity should explain our diverging intuitions about integrity while recognizing integrity as a moral concept. I resolve the dilemma by first making a distinction between substantive integrity and formal integrity. When I say, “Mary is a person of integrity,” most speakers presuppose that she is morally trustworthy. When I say that a tyrant has integrity, however, I must qualify my statement by pointing to some strongly held non-moral commitments that explain my attribution. I argue that substantive integrity refers to a person who has a whole and coherent identity that is morally uncorrupted, while formal integrity refers to a person who merely has a whole and coherent identity. In the remainder of Chapter 2, I develop a *moral identity account* of integrity that resolves the dilemma by accounting for the similarities and differences between tyrants and heroes as well as explaining the moral content of integrity.

In Chapter 3 I answer the question “Is integrity a virtue?” by addressing claims that integrity is not a virtue or that it is merely the expression of other virtues. Against these positions I argue that integrity is a virtue, namely the disposition to “be true” to oneself by maintaining a coherent self. I first build on the moral identity account to explain what constitutes a coherent “self.” I then explain how a person of integrity organizes her commitments according to her most important ends and how a desire for self-consistency motivates her to fulfill her commitments. This should suffice to show that integrity is a virtue. But my account faces questions about how the person of integrity maintains coherence, and I end Chapter 3 by responding to four such questions.

- 1) Can “honest thieves” have substantive integrity?
- 2) Can a person of substantive integrity ever lie in order to actualize another virtue?
- 3) Can a person of substantive

integrity alter or change her identity and still have integrity? 4) Can two people have opposed moral beliefs and both have substantive integrity?

In Chapter 4 I respond to an important empirical objection to the moral identity account of integrity. John Doris and Gilbert Harman argue that consistency of character is doubtful because situational factors often overwhelm a person's moral identity. The purpose of this chapter is not to contradict the findings of the social psychology experiments presented in defense of their position, but rather to show that the moral identity account of integrity can better explain their findings. I argue that integrity based on a person's moral identity is not as situationally flexible as Doris and Harman claim. Recent research on how a person's moral commitments can become more or less accessible to her working memory demonstrates that a person's moral identity can actually mediate behavior across situations.

In Chapter 5 I take up another empirical challenge to integrity as a moral virtue. David Luban claims that the quest for integrity is dangerous for two reasons. First, in our quest to maintain harmony between our beliefs and conduct, we tend to change our moral principles to justify our immoral behavior. Second, we also tend to rationalize our immoral behaviors because we want to appear upright to ourselves and others. Against Luban's position, I argue that what he labels a quest for integrity is neither necessary nor sufficient to produce the moral corruption and rationalization that he fears. A person may rationalize her conduct in a quest for achievement and wealth, but rarely would she do so in a quest for substantive integrity.

## Acknowledgements

Attaining a doctoral degree in philosophy has been a great challenge because of its rigorous requirements and long-term commitment. I benefitted greatly from the support and encouragement of many people who helped me achieve this milestone. Below I acknowledge those who made a significant contribution to my completion of this dissertation and degree. Anyone who I do not mentioned for lack of space should certainly know that I appreciate their contribution as well.

I want to thank the Graduate School of Washington University in St. Louis for its financial support during my years of study. I also want to thank the graduate students who entered the PhD program with me and who gave me the support to start and finish the program. Thanks to Jill Delston, Ben Floyd-Clapman, Amy Goodman-Wilson, Don Goodman-Wilson, Zach Hoskins, Matthew James, and Rawdon Waller for their humor and encouragement.

Thanks to the many professors who guided me through my classes, qualifying papers, and eventually my dissertation. In particular, Larry May patiently helped me narrow my dissertation topic into something that would actually fit into one book, Kit Wellman guided me through the long and arduous first draft stage, and Eric Brown gave me insightful and candid feedback on my outlines and concepts. Thanks also to John Doris whose work inspired the last two chapters. I also appreciate the insightful questions and comments of my other committee members, Simine Vazire and Kurt Dirks.

I am also indebted to my friends and family at New Creation in Christ who reminded me that God, music, prayer, and friendship make for a full life. Thanks also to Grace Pope who encouraged me to “keep writing” and made sure I stayed on track.

Finally I acknowledge my family. My parents, Harvey and Judy Bauman, who gave moral support as their son pursued his dream of studying and teaching ethics. Art and Marie Docken who provided a home-away-from-home to rest, think, and talk out my ideas. My wife and friend Lila Docken Bauman whose constant support and sacrifice carried me to this finish line. And last but not least, my son Scott Bauman who embraced my journey and joyfully skipped along with me to the top of “mount PhD.”

I dedicate my dissertation to all those who have courageously and humbly stood up for what is ethically right, even in the face of powerful opposition.

## Table of Contents

Abstract.....	ii
Acknowledgements.....	v
Introduction.....	2
Chapter 1 - Integrity: An Account for Heroes and Tyrants .....	4
Section 1: What Is Integrity? .....	4
Section 2: Integrity without Morality .....	11
2.1 - Condition 1 - Capacity to Form and Control Commitments .....	15
2.2 - Condition 2 - Reasons for Selecting Values.....	19
2.3 - Condition 3 - Identifying with Values.....	25
2.4 - Condition 4: Coherence.....	39
2.5 - Summary of Necessary Conditions .....	44
Section 3: Three Objections to the Identity Account.....	45
Chapter 2 - Integrity and Morality: Of Forms and Substance .....	48
Section 1: Reconciling the Identity Account with Morality .....	50
1.1 - Insulating Integrity from Immorality.....	50
1.2 - The Personal Integrity/Moral Integrity Distinction.....	55
Section 2: Non-Identity Accounts of Moral Integrity.....	61
2.1 - The Deliberation Account of Integrity .....	62
2.2 - Integrity as Moral Trustworthiness.....	66
Section 3: The Moral Identity Account .....	70
3.1 - Moral Trustworthiness and ID-Commitments.....	70
3.2 - Moral Values and Personal Values.....	73
Conclusion .....	86
Chapter 3 - Integrity and Virtue: A Case of Being True to Yourself .....	89
Section 1: Is Integrity a Virtue?.....	91
Section 2: Integrity and the Coherent Self.....	101
2.1 - What Is a Self?.....	101
2.2 - Beliefs as Guides .....	104
2.3 - Desires, Dispositions, and Higher-Order ID-Commitments .....	108
2.4 - Coherence and the Virtue of Integrity .....	114
2.5 - Moral Integrity.....	120
2.5 - Sub-Section Summary .....	128



Section 3: Concerns and Questions.....	129
3.1 - Specific Virtues and Integrity.....	130
3.2 - Can you lie and still have integrity?.....	131
3.3 - Can a person change the contents of her identity and still have integrity? ....	137
3.4 - Can two people have opposed moral beliefs and have substantive integrity? 142	
Conclusion .....	147
 Chapter 4 - Integrity, Moral Identity, and the Power of Situations .....	149
Section 1: Situationism and Mental Constructs .....	151
1.1 - The Situationist Challenge.....	151
1.2 - Princeton Seminarian Experiment.....	155
1.3 - Stanley Milgram’s Obedience Experiments .....	158
1.4 - The Problem of Consistent Behavior.....	163
Section 2: Moral Identity and Consistency.....	165
2.1 - Internalizing and Accessing Moral Principles.....	167
2.2 - Pro-Social Behavior and Social Responsibility.....	176
2.3 - Summary of Moral Identity Research .....	180
Section 3: Revisiting the Experiments.....	182
3.1 - Revisiting the Seminarians .....	183
3.2 - Revisiting the Milgram Experiments.....	191
3.3 - Situationists, Moral Identity, and Integrity.....	225
Conclusion .....	231
 Chapter 5 - The Quest for Integrity: Of Self-Deception and Self-Defense .....	233
Section 1: Integrity as a Dangerous Quest.....	235
1.1 - The Low Road of Integrity .....	238
1.2 - The High Road of Integrity .....	256
1.3 - Montaigne and Integrity .....	263
1.4 - Section Summary.....	272
Section 2: Integrity and Defending One’s Self.....	273
 Bibliography .....	283

“The integrity of the upright guides them, but the unfaithful are destroyed by their duplicity.” (*Proverbs 11:3*)

"The Governor of She told Confucius, 'Among my people, there is a man of unbending integrity: when his father stole a sheep, he denounced him.' To this Confucius replied, 'Among my people, men of integrity do things differently: a father covers up for his son, a son covers up for his father--and there is integrity in what they do.'" (*The Analects of Confucius*, 13.18, 63.)

## Introduction

Integrity is a concept with an identity crisis. In recent philosophical discussions, it is argued that integrity can describe both a Mafia wise guy who refuses to rat out his associates and a company whistleblower who risks her job to protect the public. It appears that integrity can refer to both a dedicated sinner as well as a dedicated saint. For the general public, this contradiction is puzzling but easy to rectify by limiting integrity to the realm of morality. For several philosophers and business ethicists, however, no contradiction exists because integrity is a non-moral concept that only becomes a moral notion when a person has commitments to moral principles. On this philosophical interpretation, integrity is a morally neutral concept that can describe both the Mafia wise guy and the company whistleblower.

In this dissertation I attempt to resolve this identity crisis by answering the questions, "What is integrity?" and "Is integrity worth pursuing?" In Chapters 1, 2 and 3, I answer the first question by discussing different integrity accounts while developing what I take is a more complete account. In Chapter 1, I review the philosophical literature concerning integrity and the general consensus that integrity is fundamentally a non-moral concept. I refine this consensus and offer some necessary conditions for having a non-moral form of integrity. A puzzle arises, however, because integrity commonly refers to people who are committed to moral principles. In Chapter 2 I argue that integrity is fundamentally a moral concept. I distinguish between *substantive integrity* which requires commitments to moral values and *formal integrity* which also requires commitments, but not necessarily to moral values. I then present my account of

substantive integrity. Chapter 3 addresses the question of whether or not integrity is a virtue. I argue that it is the virtue of *being true to oneself*, in particular a self with strong moral commitments. My answer to the first question is that integrity is a moral concept and a virtue that is anchored in one's commitments to moral values and that these commitments partly constitute our identity or self-conception.

In the last two chapters I answer the second question which addresses concerns that maintaining one's integrity may not be possible and that pursuing integrity could be dangerous. In Chapter 4 I address concerns of situationists who contend that objective situational factors, such as those found in Stanley Milgram's obedience experiments, seem to overwhelm stable mental structures such as character. They are skeptical that stable mental structures exist, which presents a direct challenge to integrity. I respond that recent research about how we access our moral commitments and constitute our self-conceptions reveals stable mental structures that make integrity possible. In Chapter 5 I address the concern that we may unconsciously change our moral principles to match our unethical conduct. David Luban argues that in our *quest* for integrity, we rationalize our immoral behavior and deceive ourselves. I respond that a quest for integrity is neither necessary nor sufficient for producing the rationalization and self-deception he fears. My answer to the second question is that integrity is possible and worth pursuing, but that situational factors and a desire to rationalize immoral behavior can threaten to undermine a person's integrity.

## Chapter 1 - Integrity: An Account for Heroes and Tyrants

In this chapter, I develop a basic notion of integrity from the literature to begin answering the question, “What is integrity?” My analysis provides a “bare bones” account of integrity that captures the central features of a person’s integrity *without* addressing the moral, non-moral, or immoral content of her commitments. In section 1, I take a broad look at the notion of integrity and the problem of defining it as a purely moral or non-moral concept. In section 2 I review the integrated-self account and the identity account which both represent integrity as a morally-neutral concept. I use the majority of section 2 and this chapter to present a generic version of the identity account and propose revisions to make the account more accurate given our usage of the word. I conclude section 2 with a summary of the necessary conditions for having integrity. In section 3 I present three general objections to identity accounts to introduce the remainder of the dissertation.

### Section 1: What Is Integrity?

The word integrity literally means the state of being untouched. John Beebe states, “*Tag*, its Sanskrit root, as the game we still call by this name implies, means to touch or handle. Out of this root come words like tact, taste, tax, and contaminate. *Integ* means *not* touched or handled.”<sup>1</sup> The earliest use of the Latin form of *integer* meant fresh, unimpaired, virgin, as well as whole and complete. Eventually the abstraction *integritas* as a moral term entered the Latin language.

---

<sup>1</sup> John Beebe, *Integrity in Depth*. (College Station, Texas: Texas A&M University Press, 1992), 6.

The Greeks do not list integrity as a virtue like courage, justice, temperance or prudence, but some Roman philosophers developed it as a moral notion. Around 70 BCE, Cicero was prosecuting the governor of Sicily for embezzling. In a speech calling for moral leadership, Cicero states, “nor can a greater disaster come upon us all than a conviction, on the part of the Roman people, that the Senatorial Order has cast aside all respect for truth and integrity, for honesty and duty [. . . rationem veritatis, integritatis, fidei, religionis ab hoc ordine abiudicar].”<sup>2</sup> Later, Seneca writing in the 60’s CE uses *integritum* as a moral quality in his book *De Beneficiis*. When considering likely candidates for his patronage, Seneca states, “I shall choose a man who is upright [*integritum*], sincere, mindful, grateful, who keeps his hands from another man’s property, who is not greedily attached to his own, who is kind to others.”<sup>3</sup>

Over the centuries, integrity was used in many contexts while still retaining its core meanings of purity of character and wholeness. In the 1913 edition of *Webster’s* dictionary, integrity is defined as:

1) the state or quality of being complete; wholeness; entireness; unbroken state; 2) Moral soundness; honesty; freedom from corrupting influence or motive; -- used especially with reference to the fulfillment of contracts, the discharge of agencies, trusts, and the like; uprightness; rectitude. 3) Unimpaired, unadulterated, or genuine state; entire correspondence with an original condition; purity.<sup>4</sup>

---

<sup>2</sup> Beebe, 1992, 7. Quoted from Cicero, “Against Verres: Part Two” in *The Verrine Orations*, Trans. L. H. G. Greenwood (Cambridge: Harvard University Press, Loeb Classical Library, 1928), Volume I, book One, II, para. 4-5.

<sup>3</sup> Quoted from “On Benefits,” in Lucius Anneaus Seneca, *Seneca Moral Essays*, Trans. John W. Basore. (Cambridge: Harvard University Press, Loeb Classical Library, 1989) Volume III, Book IV, xi.1.

<sup>4</sup> “Integrity.” Wordnik. Accessed July 25, 2010. <http://www.wordnik.com/words/integrity>.

The 1913 definitions follow almost verbatim from the 1828 edition of *Webster's* dictionary. In the 1828 edition, the example for integrity as “wholeness” refers to an individual state’s integrity as guaranteed by the US constitution. The 1828 description of “moral soundness” includes the following: “Integrity comprehends the whole moral character, but has a special reference to uprightness in mutual dealings, transfers of property, and agencies for others.”<sup>5</sup> In this sense, integrity is a term that specifically indicates a moral trustworthiness in human interactions more than a general evaluation of a person’s moral character.

The main adjustment in our current usage of integrity is that the moral concept has become more prominent. According to the 2005 *New Oxford American Dictionary*, in order of usage, integrity is “1) The quality of being honest and having strong moral principles; moral uprightness; 2) The state of being whole and undivided; the condition of being unified, unimpaired, or sound in construction.”<sup>6</sup> The American Heritage Dictionary defines integrity as “1) steadfast adherence to a strict moral or ethical code, 2) the state of being unimpaired; soundness, 3) the quality or condition of being whole or undivided; completeness.”<sup>7</sup>

The moral meaning of integrity, in particular “moral uprightness,” appears to have surpassed the non-moral notion of wholeness which is not surprising. Words with the same root as integrity are used in other languages to identify people who are morally

---

<sup>5</sup> Noah Webster, “Integrity.” *Webster's 1828 Dictionary*. Accessed October 30, 2010. <http://1828.mshaffer.com/d/search/word/integrity>.

<sup>6</sup> *New Oxford American Dictionary*, (New York: Oxford University Press, 2005).

<sup>7</sup> “Integrity.” Wordnik. Accessed July 25, 2010. <http://www.wordnik.com/words/integrity>.

uncorrupted or who resist compromising their moral commitments. The German word *integrität* means honesty and wholeness while the word *einheit* specifically means wholeness. In Immanuel Kant's *Groundwork of the Metaphysics of Morals* (Cambridge edition), the translator uses the English word integrity for the German word *Rechtshaffenheit* which can also be translated *righteousness*. The translator appears to have chosen the English word "integrity" because it best translates the moral steadfastness that Kant is describing. Kant says that an action of integrity is, "done with steadfast soul, apart from every view of advantage of any kind in this world or another and even under the greatest temptation of need or allurements, it leaves far behind and eclipses any similar act that was affected in the least by an extraneous incentive."<sup>8</sup> The idea conveyed by Kant is that an action of *Rechtshaffenheit* is done from a steadfast, fixed, and whole character. The English concept of integrity in common usage also captures this sense.<sup>9</sup>

When applied to people, integrity as "wholeness" points to a broader notion of maintaining a complete and coherent self. As mentioned above, integrity can be derived from the Latin root *integritas* which can mean unity, wholeness, and unbroken completeness, and also from *integer* meaning whole or intact. Both definitions of

---

<sup>8</sup> Immanuel Kant. *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy*. Trans. Mary J. Gregor. (New York: Cambridge University Press. 1999), 65.

<sup>9</sup> English and Latin are not the only languages that connect completeness with moral uprightness. Ancient Hebrew has a similar word derivation for integrity. One ancient Hebrew word for innocence and uprightness is *tôm* which comes from the word *táman* which means "to complete in a good or bad sense . . . come to an end." (Strong's Exhaustive Concordance). In this case, the word for uprightness comes from a morally neutral word for complete.



integrity rely on a notion of completeness, but the moral meaning specifically describes a person's complete and uncorrupted *moral* character.

We begin then with integrity meaning moral uprightness and/or wholeness. Dictionary definitions, however, only narrow the field of possible meanings and tell us how the word is used generally. They sometimes miss the nuances that have attached themselves to the concept in different contexts. Because the philosophical discussion of integrity often considers how the concept applies to people, I seek to explain what we mean when we say, "X is a person of integrity." In what follows I use two cases to draw sharper distinctions around the concept of integrity.

Case 1: Joey Scar is a member of the Mafia in the custody of the police. Even under harsh interrogation and a promise of immunity, he has refused to reveal Tony Soprano's involvement in five murders.

Case 2: During WWII, John Weidner led an operation that helped Jews flee Holland and Switzerland. At one time he was tortured by the Gestapo, but he did not give them the names of his fellow rescuers.<sup>10</sup>

Intuitively, it seems that both Scar and Weidner, two people of disparate moral character, are both candidates for the attribution of integrity. In Case 1, Scar is a person who is refusing to reveal information about five murders in the face of harsh questioning and the promise of immunity. In Case 2, Weidner is refusing to reveal information in the face of torture. Though these two men have quite different moral purposes, their resolve not to compromise in the face of adversity seem to reveal that they have some form of integrity.

---

<sup>10</sup> Kristen Renwick Monroe, *The Hand of Compassion: Portraits of Moral Choice During the Holocaust*. (New Jersey: Princeton University Press, 2004).

It is in the face of this seeming contradiction that some philosophers argue that integrity primarily picks out the “wholeness” or self-integration of a person rather than her moral uprightness.<sup>11</sup> For these philosophers, both Scar’s and Weidner’s unwillingness to compromise their “wholeness” displays *personal integrity* which is regarded as separate from *moral integrity*.<sup>12</sup> On this view, a person cannot have integrity, which I am defining initially as a deep commitment to some value or life project, unless she maintains her wholeness. In her analysis of integrity, Lynne McFall claims that, “There are conceivable cases in which we would want to grant that someone had personal integrity even if we were to find his ideal morally abhorrent; if moral justification is what we are after, moral integrity is the place to look.”<sup>13</sup> McFall gives the example of a person who is committed to deterring radicals who are burning books by burning some radicals. According to McFall, while we may attribute personal integrity to the killer, we would not say she is a person of moral integrity.

---

<sup>11</sup> Philosophers who argue that *personal integrity* is fundamentally a non-moral concept are Bernard Williams, Lynne McFall, Gabriel Taylor, Jeffrey Blustein and John Rawls. See Bernard Williams and J.J.C. Smart. *Utilitarianism: For and Against*. (New York: Cambridge University Press, 1973) and *Moral Luck*. (New York: Cambridge University Press, 1981); Lynne McFall. “Integrity.” *Ethics*, Vol. 98, No. 1 (1987); Gabriele Taylor, “Integrity.” *Proceedings of the Aristotelian Society*, Supplementary Vol. 55 (1981) and *Pride, Shame and Guilt: Emotions of Self-Assessment*. (Oxford: Oxford University Press, 1985); Jeffrey Blustein. *Care and Commitment: Taking the Personal Point of View*. (New York: Oxford University Press, 1991); John Rawls, *A Theory of Justice – Revised Edition* (Massachusetts: Harvard University Press, 1999), 455.

<sup>12</sup> A case for integrity as a moral virtue is made by Damian Cox, Marguerite La Caze, Michael P. Levine, *Integrity and the Fragile Self*. (Vermont: Ashgate, 2003). I take up the link between morality and integrity in Chapter 2.

<sup>13</sup> McFall, 1987, 14.

Jeffrey Blustein similarly concludes that personal integrity can be had by a sadist and a tyrant if they act according to their deeply held commitments.<sup>14</sup> Gabriel Taylor claims that a person of integrity must have deep commitments, but not necessarily to moral values. She states, “The person of integrity need not be a morally good person, she may not be much, or possibly not be at all, moved by other-regarding reasons.”<sup>15</sup> Ayn Rand based much of her account of morality on integrity as wholeness. She wrote in one of her journals, “Integrity – the first, greatest and noblest of all virtues—is a synonym for independence. Integrity is that quality in man which gives him the courage to hold his own convictions against all influences, against the opinions and desires of other men; the courage to remain whole, unbroken, untouched, to remain true to himself.”<sup>16</sup> Again, this definition can apply to both tyrants and heroes.

These philosophers take integrity as “wholeness” to the logical conclusion that even a tyrant who maintains her commitments can be a person of integrity. In section 2, I evaluate two integrity accounts that seek to justify the claim that the concept of integrity should be understood as primarily personal integrity (i.e. integrity as wholeness) of which moral integrity (i.e. integrity as moral uprightness) is a subset or addition.

---

<sup>14</sup> McFall, 1987, 16.; Blustein, 1991, 123.

<sup>15</sup> Ibid., 1985, 128.

<sup>16</sup> Ayn Rand, *The Journals of Ayn Rand*, Ed. David Harriman. (New York: Penguin Putnam Inc., 1997), 260. I present her definition to demonstrate the total emphasis on integrity as wholeness. Rand argues that a truly independent man would not be a tyrant because that would display a weakness – the need for other people.

## Section 2: Integrity without Morality

Two of the most widely discussed contemporary accounts of integrity are the integrated-self account and the identity account. I briefly review the integrated-self account and some objections to it as an introduction to the identity account. The identity account is the most well developed account of integrity that addresses the personal integrity/moral integrity distinction.

Cheshire Calhoun labels and describes the integrated-self account of integrity in her article, "Standing for Something."<sup>17</sup> Relying on the notion of integrity as wholeness, the integrated-self account requires a person of integrity to endorse particular desires that constitute a self. She is not easily moved by random desires nor does she endorse desires because of peer pressure. Following Harry Frankfurt's work on self-integration, Calhoun explains that the integrated-self is created by deciding which desires we should make part of ourselves and which ones we should reject. Frankfurt states, "It is these acts of ordering and of rejection – integration and separation – that create a self out of the raw materials of inner life."<sup>18</sup>

Calhoun argues that the integrated-self account of integrity has intuitive appeal.

She states,

It captures our sense that people with integrity decide what they stand for and have their own settled reasons for taking the stands they do. They are not wantons or crowd followers or shallowly sincere. Nor are they so weak willed or

---

<sup>17</sup> Cheshire Calhoun, "Standing for Something." *The Journal of Philosophy*. Vol. 92, no. 5 (May 1995).

<sup>18</sup> Frankfurt did not present his work as an account of integrity, but Calhoun has used it to create an account. Harry Frankfurt, "Identification and Wholeheartedness." *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. Ed. Ferdinand Schoeman. (New York: Cambridge University Press, 1987), 39.

self-deceived that they cannot act on what they stand for. The actions of persons of integrity express a clearly defined identity as an evaluating agent.<sup>19</sup>

Robert Audi and Patrick E. Murphy present a similar version of the integrated-self account of integrity. Confronted with a plethora of integrity definitions and usages, they argue that integrity should be taken in an *integrational sense*, or as an “integration among elements of character.”<sup>20</sup> The account stresses the coherence and unity among elements of a person’s character as well as between character and conduct. Integrity is presented as a complement to both moral and non-moral virtues in this sense of integration. Integrity is not a self-sufficient ethical standard, but integrates ethical and unethical standards into a coherent whole. Audi and Murphy paint integrity as a significant but value-neutral concept. They state, “To say that integrity, in the distinctive and wide integrational sense, is not a moral virtue implies neither that it is not good in itself, nor even that it is not essential for strong moral character.”<sup>21</sup> They do argue that integrity can and should be used as a “blunt instrument” in a moral sense to motivate moral conduct, but this is not its core meaning.

Calhoun along with Damian Cox, Marguerite La Caze, and Michael P. Levine raise two objections to the integrated-self account of integrity. The first objection is that it appears that having a fully integrated self is not sufficient to what most would identify as a person of integrity. Realistically, everyone has conflicts among their desires, intentions, and actions, and a fully integrated person would stand out as an automaton

---

<sup>19</sup> Calhoun, 1995, 237.

<sup>20</sup> Robert Audi and Patrick E. Murphy, “The Many Faces of Integrity.” *Business Ethics Quarterly*, 16, No. 1. 2006: 9.

<sup>21</sup> Audi and Murphy, 2006, 13.

more than a person of integrity who would manage these conflicts.<sup>22</sup> It seems that a person of integrity would manage these conflicts rather than automatically follow her integrated self.

The second objection is that a person could have integrity and lack self-integration. As described above, Scar or Weidner could act on the values that they most want to actualize and therefore be people of integrity in a particular domain, but at the same time could lack a high degree of overall self-integration.<sup>23</sup> To be a person of integrity typically does not require near-perfect integration of character and behavior, but possibly only in socially important domains (e.g. honesty, keeping promises). Calhoun and Cox, et. al. conclude then that even though a person of integrity must bring together various aspects of herself through “integrating and rejecting” particular desires, having an integrated self is not sufficient or the same as having integrity.<sup>24</sup> I leave the integrated-self account at this point not because it cannot be adjusted to address these objections, but because its core components of self-construction and identity are also found in the identity account. I do not answer these objections until I present the identity account to see if it can avoid them all together.

A second and more developed account of integrity is the *identity account*.

According to Calhoun, a central feature of identity accounts is that a person of integrity has a deep commitment to “those projects and principles that are constitutive of one’s

---

<sup>22</sup> Cox et al., 2003, 19-20.

<sup>23</sup> Calhoun makes the point that people can have reasons to resist resolving conflicting commitments and so self-integration may show a lack of integrity. Calhoun, 1995, 241.

<sup>24</sup> Cox et al., 2003, 26.

core identity.”<sup>25</sup> Bernard Williams, Jeffrey Blustein, Gabriel Taylor, and Lynne McFall all claim that a person’s *identity-conferring commitments* are the core from which integrity arises.<sup>26</sup> The identity account of integrity has its origin in Williams’ claim that a person’s integrity is constituted by her identity-conferring commitments to values, principles, and life projects. These commitments are ones with which a person is “deeply and extensively involved and identified.”<sup>27</sup> Williams states, “There is no contradiction in the idea of a man’s dying for a ground project – quite the reverse, since if death really is necessary for the project, then to live would be to live with it unsatisfied, something which, if it really is his ground project, he has no reason to do.”<sup>28</sup> For Williams, a person’s integrity is not based on just any commitments, but commitments to values or projects that constitute her identity.

After carefully reviewing the identity accounts offered by these philosophers, a generic identity account of integrity emerges that captures the four necessary conditions a person must meet to have integrity. The conditions are: 1) capacity to form and control one’s commitments; 2) specific reasons for selecting commitments to values; 3) identification with values; and 4) coherence among commitments as well as among commitments, motivations, and actions.

Because integrity is a fairly familiar concept, one way to identify the characteristics of a person who has integrity is to consider examples of people who would

---

<sup>25</sup> Calhoun, 1995, 235.

<sup>26</sup> Williams, 1973 and 1981; McFall, 1987; Taylor, 1981 and 1985; Blustein, 1991.

<sup>27</sup> Williams, 1973, 116.

<sup>28</sup> Williams, 1981, 13.

commonly be judged as either having or lacking integrity. I use examples not solely to pump our intuitions, but rather to mine our *understanding* of what it means to have or lack integrity. Integrity is a concept that is easier to identify than to explain, so I go from identification to explanation as I break down the necessary conditions for having integrity. I then describe the four necessary conditions in detail and conclude that a person must display them to some extent in order to have integrity. Consider Case 3 as we begin analyzing the necessary conditions for having integrity.

Case 3: Lefty and Righty are longtime members of the Mafia in the custody of the police. They have never informed to the police in the past. Before being brought in for questioning, they both tell Tony Soprano that they will not reveal his involvement in five murders. The police separate Lefty and Righty and begin questioning them. After four hours of harsh interrogation and a promise of witness protection, Lefty refuses to reveal Soprano's involvement. After four hours of harsh interrogation and a promise of witness protection, Righty reveals Soprano's involvement.

A common response to this case is that Lefty appears to be a person of integrity and Righty appears to lack integrity. But what do these different actions tell us about the necessary conditions for having or lacking integrity? We must first look at the assumptions we make about Lefty and Righty and their capacity to form and control their commitments.

### **2.1 - Condition 1 - Capacity to Form and Control Commitments**

The integrated-self account of integrity relies heavily on Frankfurt's description of how people form their identities, and the identity account similarly builds on his insights. Philosophers use Frankfurt's work on self-concept creation because it explains a plausible process of how one creates a stable identity, which is fundamental to any



account of integrity as wholeness.<sup>29</sup> According to identity theorists Taylor and Blustein, a person cannot have integrity unless she has the capacity to form stable commitments from her desires and can then control these commitments.<sup>30</sup> But to evaluate this claim, we must first understand Frankfurt's description of how people form their identities by selecting and rejecting desires.

In the article "Identification and Wholeheartedness," Frankfurt explains how a person commits to particular desires and how this commitment can constitute her identity. He explains the process in terms of first-order desires, second-order desires, and second-order volitions. The object of a first-order desire is to perform or not perform some action, like smoking or sleeping in late. The desire that ultimately results in action is thought to be what the person wanted to do, such as smoking a cigarette or hitting the snooze button. A person also has second-order desires that have as their objects the "first-order desires they want; and . . . second-order volitions concerning which first-order desire they want to be their will"<sup>31</sup>. To continue the smoking example, a person may have a second-order desire *not* to have the desire to smoke which conflicts with her first-order desire *to* smoke. She can also have a second-order volition to want not-smoking to be her will, and this conflicts with her first-order desire to smoke.

According to Frankfurt, the process of forming a second-order volition, or what he also considers a *commitment*, starts when a person "cuts off" a certain sequence of

---

<sup>29</sup> I use the term *identity* throughout the chapter to describe the core self-concept by which a person defines who she is, who she is not, and who she aspires to be. I am not making any claims about a person's identity through time or issues regarding the identity/body connection.

<sup>30</sup> Taylor, 1985; Blustein, 1991.

<sup>31</sup> Frankfurt, 1987, 32.

desires and makes a decision that a particular desire (e.g. to tell the truth) now partly constitutes her identity (e.g. “I am an honest person”). Whether through deliberation or a non-conscious process (e.g. a parent’s example or a developed disposition), the decision to commit to a desire forms an *intention* to actualize the desire and that intention is the second-order volition or commitment.<sup>32</sup> Of course forming a commitment does not mean that other desires and intentions will not interfere with her actualizing this particular desire, but deciding to have the commitment means that she is not equivocating about which desire she wants to make her will.

I agree with Blustein and Taylor that to have integrity a person must first have the *capacity* to form commitments concerning first-order desires.<sup>33</sup> If integrity requires a whole self, then a person must be able to select those desires that will constitute her identity while rejecting others. For example, a person who acts only on whatever desire happens to be the strongest at the time is not a candidate for integrity because this is contrary to having anything that could be considered a whole self. On Blustein and Taylor’s view, only a person who has the capacity to commit to one desire rather than another (i.e. form a second-order volition) can select which desires to actualize and which

---

<sup>32</sup> Frankfurt does not imply that these decisions must be highly articulate and reflective mental acts, but rather he accepts that the decision can occur when a person non-consciously “makes up her mind” that a particular desire is her own. Frankfurt, 1987, 38, 40-42.

<sup>33</sup> Blustein observes that to possess integrity, a person “must be capable of wanting certain of their first-order desires to be effective and of rejecting others.” Blustein, 1991, 96.

ones to reject, and only a person with the capacity to make these decisions can have integrity.<sup>34</sup>

In addition to having the capacity to make commitments, a person of integrity must also be able to control her commitments and their resulting actions. According to Taylor, “An account of integrity as being primarily a matter of the agent’s control over what she is doing explains the type of behavior we expect of a person we regard as having integrity.” She goes on to say that the person’s choices are her own and her identifications with her desires (i.e. commitments) are not subject to “unconsidered change.”<sup>35</sup> For example, Lefty may have a commitment to actualize his fidelity to Soprano, but he must be in control of that commitment to have integrity. It appears that he lacks integrity if the strongest desire in his mind at the time consistently trumps the commitments that make up his identity. On Taylor’s account, a person of integrity controls her commitments and her first-order desires, and to lack this control is a sign that the person lacks integrity.

In a less-rigid version of Taylor’s view, Blustein does think control is necessary to have integrity; however he allows a person of integrity to act on what comes “naturally.” He states, “In these cases, the agent’s knowledge of what their life is about, their commitments and the actions that flow from them, are second nature.” It seems that having a tight grip on every action is not necessary to be a person of integrity if acting on

---

<sup>34</sup> Both Taylor and Blustein agree that the capacity to form second-order volitions is not sufficient for having integrity because a person could form second-order volitions without a sufficient commitment to them. Taylor calls these people “shallowly sincere” because they lack the consistent commitment necessary to be people of integrity. Blustein, 1991, 98; Taylor, 1985, 113-115.

<sup>35</sup> Taylor, 1985, 126.

one's commitments is effortless. While I agree with Taylor and Blustein that to have integrity a person must control what she is doing, I suggest that Blustein's description more accurately reflects how we identify a person of integrity. On this account, a person who actualizes her commitments without exceptional effort is also in control.

My revised first necessary condition is that for a person to have integrity as wholeness, she *must have the capacity to form commitments (i.e. second-order volitions) and the capacity to control her commitments and actions*. The control she has can be expressed in exceptional effort or effortless assent to her most strongly held commitments.

Before analyzing the second necessary condition, I need to broaden the object of commitments from desires to *values*. This change is useful because integrity as wholeness involves much more than forming commitments toward desires. People make commitments to desires as well as projects, virtues, values, purposes, and dispositions. From this point forward I use the word *value* in place of Frankfurt's word *desire*, and by value I mean any desire, value, virtue, and/or project that is an object of a person's commitments.

## **2.2 - Condition 2 - Reasons for Selecting Values**

The identity account of integrity also constrains how a person with integrity should select particular values to become commitments. The second necessary condition requires the person to have *reasons* for selecting particular values. Identity theorists accept that most values are not reflectively and consciously selected as commitments, but Taylor in particular argues that a person of integrity must have reasons for her

commitments. A related concern of Taylor's is that one may have reasons for selecting a particular value, but that other people may have significantly influenced the selection. I consider these two requirements below.

According to Taylor a person of integrity will, "engage in some form of reasoning: if he does not just act on whatever inclination happens to be the strongest he must have some reason for wanting one value rather than another to be effective, though he need not necessarily be able to articulate that reason."<sup>36</sup> While she allows that some reasons for having commitments may not be articulated, these reasons must still exist for the agent. She states, "if an agent is to value something and to have control over her values then (at least) her wanting some desire [value] to be effective must be based on some reason such that any reason she accepts in favor of a change in identification must be thought to override it. For this to be possible it, the earlier reason, must have a role to play in that person's practical reasoning."<sup>37</sup> Taylor's standard for reasons is that the person must have some reason, even if it is not articulated, for her commitments and the resulting actions. To lack reasons is to lack integrity.

I disagree with Taylor's standard because it may disqualify people who act on commitments that arose from non-rational sources such as a personal disposition or a pattern of behavior inspired by a respected person. For example, Lawrence Blum describes Magda Trocmé, a courageous French rescuer of Jews in WWII, as a person who had what appear to be non-rational dispositions for her actions. She "did not come

---

<sup>36</sup> Taylor, 1985, 113.

<sup>37</sup> Taylor, 1985, 116.

morally armed with a worked out set of general ideals which she searched for ways to implement . . . : `I do not hunt around to find people to help. But I never close my door, never refuse to help somebody who comes to me and asks for something.’”<sup>38</sup> Trocmé appears to have integrity because of her commitment to the value of helping people who ask, even when her reasons are not articulated. Trocmé’s example implies that a person can have integrity as evidenced by her commitments, even when she cannot articulate reasons that justify her commitments.

Taylor also holds that a person of integrity “must get her practical reasoning right and act on that reason which, *all things considered*, she thinks best.”<sup>39</sup> In other words, the person of integrity must first be clear that she has chosen the best reasons before acting and second, not be self-deceived about her reasons for acting. But while these conditions may be intuitively correct if we assume an extremely high standard of reflection for agents, they are unnecessary because it is well within the notion of integrity that a person of integrity could not know the reasons that justify her actions or is self-deceived about her reasons for acting. Consider the Weidner case with a fictional person named Beidner. What if after refusing to talk during a brutal interrogation session with the Gestapo, Beidner realizes that she does not have a second-order volition to fidelity. Instead she has a second-order volition to thwart any project of the Nazi’s.<sup>40</sup>

---

<sup>38</sup> Lawrence Blum, *Moral Perception and Particularity*. (New York: Cambridge University Press, 1994), 86.

<sup>39</sup> *Ibid.*, 1985, 128. Emphasis mine.

<sup>40</sup> Self-deception reveals a lack of integrity if it is the kind in which a person believes she has a commitment to a particular value but rarely actualizes the value when given the opportunity. Unlike Beidner above, she believes she is committed to a value, but actually has none. Beidner, however, may act

Two factors point to Beidner having integrity even if she does not act for all-things-considered reasons and is self-deceived. First, her actions show that she has the first necessary condition of having a commitment that she can control. By meeting this first necessary condition without knowing which commitment she holds, Beidner appears to be a person of integrity. The second factor is that reasoning correctly or incorrectly would not change an attribution of integrity. If Beidner does not inform, it makes little difference if she has all-things-considered reasons for not speaking or if she is mistaken about her reasons but resists because she has a strong disposition to thwart evil people. I suggest that her reasoning adds little or no weight in attributing integrity to her because *acting* in line with her commitment supports the claim that she has integrity. Similarly, we can imagine a whistleblower at a company saying that she sacrificed her job because it was, “The right thing to do.” The attribution of integrity may assume some correct *reasoning* but primarily rests on her acting on a commitment to “do what is right” rather than the quality of her reasons or reasoning.

Another constraint on reasons in the integrity literature is *how* a person develops a commitment to actualize a value. Typically a person develops a commitment in a variety of ways such as through interactions with society or a group, interactions with individuals she respects, and her own dispositions and personal choices. Taylor claims that a person may not have integrity if she has a commitment merely because a group or another person holds that commitment. While most identity accounts agree that we rarely decide

---

consistently on her commitment to thwart any actions taken by the Nazis while believing she is committed to fidelity. I discuss self-deception and cognitive dissonance in detail in Chapter 5.

to commit to a value and then instantly begin to act on it, these accounts do expect a person to determine *for herself* which values are important to her.

Taylor argues that a person who looks to a group to determine which values to identify with has abdicated her ability to decide which values are most important for her. Her assessment appears to be correct given what integrity seems to require, but I would add that a person who has a commitment to the value of *following a group and its code* has integrity regarding her commitment to the group. Taylor claims that a person who has, “a *general* desire always to be guided by the group” would not have integrity, and this also seems correct given the personal nature of integrity. But if the person has a personal commitment to be guided by the group, then I claim that her actions justify an attribution of integrity.

Consider the case of Lefty in this context. What if he has a commitment to fidelity because he accepts the Mafia code that, “You never rat out the family.” Lefty still appears to be a person of integrity whether or not his refusal to inform comes from a commitment to fidelity or a commitment to follow the Mafia code. My point is that adopting the values of a group instead of deciding which individual values one wants to actualize does not disqualify a person from being a person of integrity *as long as* the person has a commitment to follow the group and its code. Lacking a commitment, Taylor’s concerns are well founded. Of course, any cases of brain washing, coercion, or any coercive group mechanism that removes a person’s capacity to control her actions would nullify an attribution of integrity.



Where does this leave reasons and reasoning in regard to integrity? On my revised account, all that is necessary to have integrity is that a person have the capacity to consider reasons for or against her volitions and actions. A person who has integrity, as opposed to someone who is blindly stubborn or brainwashed, is generally thought to be open to reasons even when these reasons may not cause any change in her commitments or actions. I can imagine the Gestapo explaining why Weidner should give up the names of his comrades. They may say, "If you talk you can avoid being beaten, save your life, and maybe receive a shorter prison sentence." For Weidner to have integrity, it seems like he must at least have the capacity to consider reasons and evaluate them. A person who is mentally unstable or only hears gibberish when reasons are given is not a candidate for integrity. This necessary condition distinguishes people with integrity from people who are merely stubborn or obsessed.<sup>41</sup>

To summarize the second necessary condition, contrary to Taylor I have argued that a person may have integrity if her commitments arise from dispositions or other non-conscious causes. Also, on my account a person such as Beidner has integrity even if she is mistaken about her reasons or self-deceived about her exact commitments. On my account, Beidner has integrity as long as she acts from some commitment that she can control. When it comes to acquiring commitments, a person can follow a group's code and still have integrity as long as she has a commitment to follow that group's code. Finally, for a person to have integrity she must have the capacity to reason and consider reasons, even if they do not change her commitments or actions.

---

<sup>41</sup> I address the differences between dogmatism and integrity more fully in Chapter 3.

My revised second necessary condition is that for a person to have integrity, *she must have the capacity to reason and consider reasons.*

### **2.3 - Condition 3 - Identifying with Values**

Consider the case of Lefty and Righty above and ask yourself why it appears that Lefty has integrity and Righty does not. The most obvious characteristic that distinguishes Lefty and Righty is that Lefty appears to act on a commitment because his steadfast resistance demonstrates that he has more than just a mere desire or preference to maintain his fidelity. To elaborate, by commitment I mean a stable, second-order volition to actualize some value even in the face of external pressure to compromise. On the other hand, Righty does not appear to have a commitment because he appears to bow to external pressure. Righty's action reveals at best a wavering desire to keep quiet in contrast to Lefty's commitment to maintain his fidelity.

Identity theorists make the third necessary condition, *identifying with values*, the corner-stone of their account because it captures the "wholeness" required to have integrity. As mentioned above, Williams claims that a person's integrity is constituted by those commitments to values, principles, and life projects that confer an identity on her. These specific identity-conferring commitments are ones with which a person is "deeply and extensively involved and identified."<sup>42</sup> Following Williams, McFall argues that it is a conceptual truth that, "personal integrity requires unconditional commitments" that confer an identity to the person. Blustein also notes that, "Personal integrity, in one of its

---

<sup>42</sup> Williams, 1973, 116.

senses, is the consistency that obtains when people act according to their commitments.”<sup>43</sup>

Taylor also recognizes that a person who has integrity must identify with and commit to particular values. The identification cannot be shallow, for example changing one’s “commitments” from one day to the next, but should be “reasonably consistent” such that she cannot identify with some value while ignoring its implications on other occasions.<sup>44</sup> Integrity requires consistency of action and this type of consistency comes from identifying with particular values. The central thesis that these philosophers agree on is that in order to have integrity, a person must identify with some values to such an extent that it makes losing her integrity possible, or to quote McFall, “In order to sell one’s soul, one must have something to sell.”<sup>45</sup>

I analyze this third necessary condition by reviewing Frankfurt’s discussion of how a person “identifies with” a particular desire and how this decision can constitute a person’s identity. I then present an account of how identifying with a value can be the same as committing to actualize that value. I explain how a person’s identity is constituted and revised to the extent that the person commits to actualizing a value.

First, a note on terminology. In what follows, I use the word *constitute* in regards to a person’s identity *not* to imply that the person’s entire identity is made up of her commitments, nor to imply that a person’s identity cannot be changed. Instead, I use the word to indicate that when a person identifies with a particular value, the commitment to

---

<sup>43</sup> McFall, 1987, 16.; Blustein, 1991, 105.

<sup>44</sup> Taylor, 1985, 115.

<sup>45</sup> *Ibid.*, 1987, 10.

actualize the value becomes part of the person's identity. I broadly define *identity* as who the person believes she is as well as who she wants to be. To have an identity is to have a self-concept of who one is and is not.

Frankfurt states that a person forms a second-order volition when she "cuts off" a certain sequence of desires and makes a decision that a particular desire (e.g. to tell the truth) now partly constitutes her identity (e.g. "I am an honest person"). According to Frankfurt, when a person terminates the unstable give-and-take among various first-order desires and decides, consciously or non-consciously, that a particular first-order desire will be her will, the act of deciding is described as *identifying with* a specific desire. The question then is *to what extent* will the decision to commit to a desire constitute a person's identity? On Frankfurt's account, "The decision determines what the person really wants by making the desire upon which he decides fully his own. To this extent the person, in making a decision by which he identifies with a desire, *constitutes himself*."<sup>46</sup>

I interpret Frankfurt as saying the following: To the extent that the person *really wants* the value (i.e. desire) to be fully her own is the extent to which she *identifies with* the value, and this is the extent to which the value constitutes her identity. In other words, deciding to commit to a value does not automatically result in a revised identity. On this interpretation, the factor that determines how much one's identity is constituted is how much one "really wants" the value to become one's will. A person who "really wants" honesty to be her will makes honesty a more fundamental part of her identity than a person who minimally wants being honest to be her will.

---

<sup>46</sup> Frankfurt, 1987, 38.

An example should make the process clearer. Assume that as a teenager, Lefty desires that his actions correspond with the value of keeping his word to others, but he has other values that conflict with this particular desire. At some point in his life, he decides that the value of fidelity is an important value and he forms a commitment to make actualizing fidelity his will. Note, however, that creating a commitment may not substantially constitute his identity. Now imagine that a person Lefty respects displays great fidelity and her example so impresses him that he “really wants” fidelity to be a central part of who he is. Lefty then deepens his identification with fidelity which further constitutes his identity and increases his level of commitment to fidelity. He also reinforces his commitment by regularly keeping his word. It seems that when a person “really wants” a value to become her will, identifying with the value and committing to the value are the same thing.

I propose that when considering commitments in the realm of integrity, a person who has integrity must not only identify with a value (i.e. make a commitment), but must also be determined to actualize that value. As mentioned earlier, the act of identifying with a value does not necessarily make the person more “whole” because a person could identify with one value in the morning and another in the evening. Integrity requires identifying with a value to such an extent that one makes a commitment that constitutes one’s identity. To account for the strength of a person’s commitment to a value which Frankfurt identifies as “really wanting” a desire, I bring in the notion of determination. *Determination* is an important factor to consider when talking about integrity as can be seen in Case 3 above. It appears that Righty initially identifies with the value of fidelity

when he speaks to Soprano, but he lacks the determination to actualize that value under pressure. In Frankfurt's words, Righty does not "really want" to make fidelity his will or his commitment is at best weak.

Frankfurt does not discuss different levels of commitment and determination. Instead, he focuses on the strongest possible case of identifying with a desire and its effect on constituting one's identity. The type of identity-constituting decision he discusses is one that is taken "without reservation" and "is a decisive one."<sup>47</sup> He calls the strongest commitment to actualize a desire a *wholehearted* commitment and this type of commitment "resonates" throughout other desires and volitions.<sup>48</sup> For Frankfurt, a commitment is either wholehearted or not, and I return to his description later. To analyze integrity, however, a finer-grained account of commitment which accounts for a person's determination would be helpful because just as commitments can be held with varying strengths so can the stability of one's identity.

To better distinguish among commitments, I describe three distinct commitment *levels* which differ based on the extent to which the person identifies with a particular value and her degree of determination to actualize the value. As I sketch a rough description of possible commitment levels, I explain how changes in identification and determination vary with each level of commitment.

### *2.3.1: Level 1 Commitments*

---

<sup>47</sup> Frankfurt, 1987, 38.

<sup>48</sup> Frankfurt, 1987, 42 and 44.

The lowest level of commitment is a *Level 1 commitment*. In Frankfurt's terms, a Level 1 commitment is the mere decision to form a second-order volition to make a particular value one's will. For example, I may decide that good health is important to me and I form a second-order volition to actualize good health. For me to have a Level 1 commitment to health, I must *decide* that health is more important to me than other values and form a second-order volition to make health maintenance my will. I minimally *identify with* the value of maintaining my health (e.g. "I want to be a person who values my health when it is convenient.") which means I am minimally *determined* to actualize that value.

A Level 1 commitment may not have a significant influence on how a person lives her life. If I have a Level 1 commitment to health, I consider the value as minimally important to my identity and I am determined to exert minimal effort to actualize it. For example, if I go to a party and can choose between a bacon cheeseburger and a salad, my commitment to the value of health may be overridden by a conflicting desire to eat a trans-fat-filled bacon-cheeseburger. While I have decided to make health an important value, I have not determined "in my heart" to actualize this value. A Level 1 commitment lacks the strong identification and determination required to actualize the value in most circumstances.

If my commitment to a value is easily overthrown, it may appear that a Level 1 commitment is only a *preference* for health or simply *caring* about my health, assuming that I am not deceiving myself about actually having health as a value. Nancy Schauer argues that a Level 1 commitment is the same as caring. She states, "We find ourselves

able to care about some things more than others, but what we are able to care about and be committed to may change from time to time, and cannot be preserved merely by our trying to prolong it.”<sup>49</sup> I disagree because a Level 1 commitment appears to be different from a preference or a care in two important ways. First, a person who decides to commit to a value is no longer vacillating about which value she wants to actualize in her life as opposed to other values. I may *prefer* a bacon-cheeseburger over a salad because it tastes better, but the preference could change because it does not relate to my identity or how I want to live my life. Likewise, I may *care* for my health but still not identify with being a person who wants to maintain his health. However, if I care for my health and eat the salad instead of the cheeseburger, a plausible explanation for my action is that I have committed to the value of health. If I have a Level 1 commitment and fail to actualize my value of health, one possible explanation is that I do not strongly identify with the value and am only weakly determined to actualize the value in the face of competing preferences, inclinations, or other reasons.

The second difference between having a commitment and having a preference or care is the feelings of regret and guilt that typically accompany a failure to fulfill a commitment. These feelings do not typically follow when a person does not actualize a value that she *prefers* or *cares* about. I base my claim on the fact that regret naturally accompanies actions that are contrary to a person’s commitments and identity. Most people can relate to the regret and guilt felt when they are minimally committed to a value and yet fail to fulfill that commitment. For example, if I eat the bacon-cheeseburger

---

<sup>49</sup> Nancy Schaubert, “Integrity, Commitment, and the Concept of a Person.” *American Philosophical Quarterly*, Vol. 33 No. 1. (January 1996): 123.



given my commitment to health, I am justified in feeling regret driving home because I acted contrary to my commitment. If I merely had a preference for the value of health, I may be justified in feeling disappointed but not feeling regret and guilt. In some cases I may care about my health to a high degree and may feel regret when I act contrary to what I care about. But then I should recognize that I care about my health to such a high degree that I probably have a Level 1 commitment to health. I do not consider the line between a Level 1 commitment and preferring or caring to be bright, but these concepts are unique and Schaubert's analysis misses these distinctions.

To summarize, a person with a Level 1 commitment makes a decision that a particular value is more important to her than other values, she minimally identifies with the value, and she is minimally determined to actualize that value in her daily life. Most people would rarely attribute integrity to a person with a Level 1 commitment because her commitment is minimal. In Case 3 above, it appears that Righty fits the description of a person with a Level 1 commitment because he has a commitment to fidelity when he is safe, but lacks the identification and determination to act on it when pressed.

Higher level commitments, on the other hand, are less likely to be compromised. The main differences between a Level 1 commitment and higher level (i.e. Level 2 and Level 3) commitments are the extent to which the person identifies with the value and the determination she has to actualize the value. I present a clearer picture of these two conditions by first explaining how identifying with a value strengthens a commitment and then describing the increasing degrees of determination found in higher level commitments.

### 2.3.2: Levels 2 and 3 as Identity-Conferring Commitments

As mentioned above, some philosophers classify commitments with which a person strongly identifies as *identity-conferring commitments* or what I call *ID-commitments*. Any commitment above a Level 1 commitment is an ID-commitment because rather than the person minimally identifying with the value, her identification with the value is strong enough to constitute her identity and provide the determination to actualize the value. A person who forms an ID-commitment has to some extent altered her identity.

Philosophers have argued that having integrity requires having some form of an ID-commitment. Taylor argues that an agent values (i.e. commits to) something, “only if his relevant identifications are reasonably consistent.”<sup>50</sup> On Taylor’s view, the agent with ID-commitments has stable identifications that consistently result in second-order volitions (i.e. commitments) being actualized across situations and times. McFall also notes that ID-commitments “reflect what we take to be most important and so determine, to a large extent our (moral) identities.”<sup>51</sup> She states that these ID-commitments are, “what it means to have a ‘core’: a set of principles or commitments that make us who we are.”<sup>52</sup> On this account, to have integrity requires having a high level commitment to a particular value, which means that a person identifies with the value to such an extent that it constitutes her identity. The extent to which the value constitutes the person’s identity will vary with the level of commitment and the measure of determination.

---

<sup>50</sup> Ibid., 1985, 117.

<sup>51</sup> Ibid., 1987, 13.

<sup>52</sup> Ibid., 1987, 13.

It is important to note that a person who *identifies with* a value can do so in a positive sense or a negative sense. In the positive sense that I am using, the person wants to be, or sees herself as, the kind of person who actualizes the value. In the negative sense, the person may see that the value is part of her identity and she does not want to be the kind of person who actualizes the value. For example, “I identify with the desire to drop all my responsibilities and quit, but I won’t do it.” In this chapter, I use *identify with* in the positive sense. I discuss identity construction in greater detail in Chapter 3.

If we look at the cases of Scar, Weidner, and Lefty, it appears that they have at a minimum what I define as *Level 2 commitments* to fidelity. The evidence for the Level 2 commitment is their unwillingness to compromise the value even under great pressure. They appear to have an ID-commitment to the value of fidelity which means that they strongly identify with the value and are determined not to compromise it even in the face of adversity. Level 2 commitments, however, do not represent the strongest identification with a value.

A person with a *Level 3 commitment* has an ID-commitment to a value *and* she has integrated this particular value into her identity to such an extent that she cannot compromise this value without substantially compromising her self-concept. Imagine Scar finally breaking under extreme pressure and reluctantly informing on Soprano. His compromise may indicate that his ID-commitment to fidelity constituted his identity, but was not *integrated* into his identity.<sup>53</sup> In contrast, if we look at the real life case of

---

<sup>53</sup> I am not arguing here for a combined identity/self-integration account of integrity. I am, however, arguing that there are deeper levels of identifying with a value. At Level 2 the value constitutes a person’s identity, but at Level 3 the value is indistinguishable from the person’s identity.

Weidner it appears that his commitment to fidelity is integrated into his identity. He was imprisoned, beaten, and his life was threatened, but he did not inform. Weidner provides an example of a Level 3 commitment because it appears that his identity and his commitment to fidelity are the same thing. He would rather lose his life than compromise the commitments that constitute his identity.

Williams calls the phenomenon of not being able to act or even think of acting against a certain value “practical necessity.” He describes practical necessity as the “conclusion not merely that one should do a certain thing, but that one *must*, and that one cannot do anything else.”<sup>54</sup> The “must” used by the person with a Level 3 commitment is one that is unconditional and “goes all the way down” to a person’s very being. It is the commitment that Frankfurt calls wholehearted. The values the person holds are so integrated into her identity that to compromise them would be to fundamentally compromise who she is. What is unique about Level 3 commitments, therefore, is that the values are held in a necessary way so that they are indistinguishable from the person’s identity.

### 2.3.3: *Determination at Levels 2 and 3*

A person may identify with a value, but not be willing to exert the required effort to actualize it. We can imagine Righty identifying with the value of fidelity but not acting on it because it would require too much effort. In the other cases above, it appears that Scar, Weidner, and Lefty identify with particular values and endure significant hardship in order to actualize the values. Following on these examples, it appears that it

---

<sup>54</sup> Ibid., 1985, 188.

is not enough for a person to have only an ID-commitment (i.e. Level 2 or 3), but she must also have the *determination* to actualize the value with which she identifies.

Consider the following case in this regard.

Case 4: Weidner describes a woman in his rescue organization named Suzy who was captured by the Gestapo. Suzy broke the first rule of the resistance organization: never write down the addresses of other members. Suzy refused to reveal the names of the rescuers, but when the Gestapo threatened to kill her mother and father she gave the Gestapo the names and addresses. Within two days almost 150 members had been arrested except for Weidner and four other officers.<sup>55</sup>

In Case 4, Suzy reluctantly informs on Weidner because the Gestapo made her an offer that trumped her ID-commitment to protect the rescue organization. If we assume that Suzy had a Level 2 commitment walking into the interrogation room, then we know that she identifies with keeping her promises and that she is determined to keep them. In fact, she remained silent until the cost of not informing became much too high. In the end, Suzy strongly identifies with promise keeping but she lacks the determination to actualize it in all circumstances.

Determination is the degree of resolve to actualize a value given its cost on one's time, wealth, effort, pleasure, physical comfort, family or other considerations. Degrees of determination could be placed on a continuum. At the low end is "Minimum Determination/Minimum Cost" and at the high end is "Maximum Determination/Maximum Cost." A person with a Level 1 commitment has determination at the "Minimum Determination/Minimum Cost" end of the continuum and would lack integrity in most circumstances. With a Level 2 commitment, Suzy is somewhere in between these two poles because her determination is strong but ultimately defeasible

---

<sup>55</sup> Monroe, 2004, 126.

because she does not have the necessary determination to have a Level 3 ID-commitment.<sup>56</sup> She may keep her promises in almost every circumstance, but she has a limit on the cost she is willing to incur.

Frankfurt recognizes that circumstances can reveal the difference between a Level 2 and Level 3 ID-commitment, or what he calls a lack of wholeheartedness. The evidence is the person's determination. Frankfurt states: "We do not know our hearts well enough to be confident whether our intention that nothing should interfere with a decision we make is one we ourselves will want carried out when – perhaps recognizing that the point of no return has been reached – we come to understand more completely what carrying it out would require us to do or to sacrifice doing."<sup>57</sup> This is the point where a person's wholeheartedness (i.e. strength of one's identification and one's determination) is tested.

If we again consider Case 2, one explanation of Weidner's actions is his strong identification with the value of fidelity and his resolute determination to actualize the value. His determination appears to be at the "Maximum Determination/Maximum Cost" end of the continuum as demonstrated by his uncompromising resistance even when tortured. Based on his strong identification and unconditional determination, Weidner

---

<sup>56</sup> I do not take integrity as wholeness to mean that a person of integrity must be unconditionally committed to a value, but rather that her identity is constituted by a value to a significant degree. A person like Suzy who has a Level 2 ID-commitment rarely if ever breaks her promises and if it were not for this extreme circumstance she would still have kept her word. She is a "person of integrity" in most contexts. However, the circumstance reveals that her commitment lacks determination. If Suzy were a person with a Level 3 ID-commitment to promise keeping, she would have kept her promise regardless of the circumstances because she is not a person who breaks promises.

<sup>57</sup> *Ibid.*, 1987, 44.

has a Level 3 ID-commitment. Using Frankfurt's term, we would say that Weidner has a wholehearted commitment to fidelity.

#### 2.3.4: *Commitment Summary*

While I have used a broad brush to paint the details of the different commitment levels, the central features and differences among the levels are clear. According to my account of commitments, anyone who has decided that a particular value is more important than other values has some level of a commitment. Like Righty's commitment to fidelity, a person with a Level 1 commitment has decided that she wants it to be her will, but she minimally identifies with the value and is minimally determined to actualize it. Beyond Level 1 commitments are ID-commitments which are those commitments to values that more fully constitute a person's identity. Suzy, a person with a Level 2 commitment, strongly identifies with the value of fidelity and she is determined (i.e. willing to pay the cost) to actualize this value. A person at Level 2, however, may discover that the cost of maintaining a value is too high and may reluctantly compromise her identity. If it were not for extreme circumstances that demand the person pay a high cost, she would have actualized the particular value. Finally Weidner, who I claim has a Level 3 commitment based on his uncompromising stand in the face of death, is a person whose ID-commitments are synonymous with his identity. Because of the seamless integration of the value with his identity and his steadfast determination, actualizing the value of fidelity is a practical necessity.

Base on my analysis and arguments, I revise the third necessary condition as follows. For a person to have integrity she must *have an ID-commitment (i.e. Level 2 or*

3 commitment) to a value and the determination to actualize it. The most obvious cases of integrity are those people with Level 3 commitments.

#### 2.4 - Condition 4: Coherence

Identity accounts of integrity claim that a person who has integrity has internal coherence which can be defined as coherence among different commitments as well as among commitments, motivations, and actions. In this sub-section I briefly explain this condition and make one revision. In Chapter 3 I develop my account of the structure of a person's set of ID-commitments and explain how this structure prioritizes the actions the person takes.

In line with the identity account of integrity presented in this chapter, McFall defines integrity as the state of being “undivided; an integral whole” and interprets this as a necessary but insufficient condition of coherence.<sup>58</sup> She explains that for a person to have integrity, she must have *internal coherence* which is constituted by three types of coherence: 1) consistency within one's set of commitments; 2) consistency between commitments and actions, and 3) consistency between commitments and motivations.

The first coherence is simple consistency within one's set of commitments, which means that the person orders and prioritizes commitments (e.g. justice and happiness) so that they do not conflict or so that conflicts are resolved. Constant conflict between one's commitments betrays a disorganized self that does not have the wholeness necessary for integrity. Righty may be a person with conflicting commitments because when he is with Soprano he wants to be a person of fidelity, but when he is with the police he is not sure

---

<sup>58</sup> McFall, 1987, 7. She takes the definition of integrity from the *Oxford English Dictionary*. She does not specify the edition.



which commitment to follow. His actions may betray the fact that he has two commitments, such as to fidelity and avoiding pain, that are vying for dominance. This type of conflict indicates that Righty does not have integrity. It seems that to have integrity, a person must organize her commitments so that they are coherent and consistent rather than unpredictable or in constant conflict.

This is not to say that commitments will never conflict or that a person must have a perfectly consistent set of commitments to have integrity. Realistically, our families, jobs, and society make many conflicting demands on us and these can bring some commitments into tension. We can imagine Suzy facing the conflict between her commitment to the rescue organization and her commitment to her parents. A person with integrity, however, should manage her commitments based on the situation without fundamentally changing her identity. Suzy, in one sense of the word, shows fidelity toward her family at the expense of her fidelity to Weidner. While I discuss the structure of a person's commitments in greater detail in Chapter 3, it is enough at this point to claim that a person who has integrity as wholeness must have coherence among her commitments to the extent that they constitute a whole identity and not an identity that is unmanageable and conflicted.

The second coherence McFall mentions is that between commitments and actions, especially in the face of temptation.<sup>59</sup> McFall argues, again from the notion of integrity as wholeness, that it is commonly held that incoherence between commitments and actions indicates a lack of integrity. These types of incoherence are weakness of will (i.e.

---

<sup>59</sup> McFall, 1987, 7.

giving in to temptation) and self-deception (i.e. giving in to temptation but rationalizing that one did not act against one's commitments). Weakness of will and self-deception display a lack of integrity because integrity as wholeness requires that a person's actions are in accord with her identity, as constituted by commitments. To have a commitment to honesty while deceiving others reveals either that one's will is too weak to follow through on the commitment which means that one lacks determination and/or one may be self-deceived about having an ID-commitment to honesty in the first place. Either way, a lack of coherence between commitments and actions reveals a lack of integrity.

The last form of coherence is that between commitment and motivation. McFall correctly observes that one can do the right action from an incompatible motivation. In these cases the action *corresponds* with a commitment but the motivation is not appropriate for attribution of integrity.<sup>60</sup> Her example is someone who tells the truth only when it serves his immediate selfish interests. While the person may tell the truth even in the face of adversity, his motive for telling the truth is not grounded in a commitment to be honest; hence a lack of coherence exists.

*Affective reactions* are one form of motivations that lack coherence because they are not linked to commitments. Affective reactions are typically selfish (i.e. self-regarding) or altruistic (i.e. other regarding). Selfish reactions to a situation may be fear, greed, lust, and envy while altruistic reactions may be compassion and generosity. To understand this form of incoherence, consider three motives for Lefty not informing on Soprano.

---

<sup>60</sup> McFall, 1987, 8.

- 1) Lefty's motive for not talking is greed. He hopes that Soprano will give him \$10,000 for not talking.
- 2) Lefty's motive for not talking is fear. He is afraid Soprano will have him killed if he talks.
- 3) Lefty's motive for not talking is compassion. He feels sorry for Soprano because he had a difficult childhood.

In statement 1) Lefty lacks integrity because greed is a selfish reaction that can change a person's actions depending on the situation and therefore his action does not reveal the ID-commitment required for integrity. It is difficult to attribute integrity to Lefty if his motive is greed because if the police offer him more money he may inform on Soprano. Similarly, the motive of fear in statement 2) reveals a lack of integrity because whoever can create the strongest fear in Lefty can change his actions. If the police threaten him with torture and death, then he may inform. The reason why these selfish reactions cause us to withdraw attributions of integrity is that they are not linked to any ID-commitment.

Statement 3) presents us with an altruistic reaction that intuitively makes us reconsider an attribution of integrity for Lefty.<sup>61</sup> Compassion, like greed and fear, is inherently unstable because if the situation changes in a way that reduces the compassionate reaction, then Lefty may be motivated to inform. For example, if the

---

<sup>61</sup> Bill Puka makes a similar contrast between solid character and altruism. He states, "Character brings organization and good order where morality had been out of sorts . . . Altruism neither gets us organized nor keeps us on the straight and narrow. Moreover, it does not seem dependent on any structure of propriety whatsoever, strict or casual." A person of integrity characteristically governs her actions while a person who shows compassion may act from affect and not ID-commitments. Bill Puka, "Altruism and Character." *Moral Development, Self, and Identity*. Ed. Daniel K. Lapsley and Darcia Narvaez (New Jersey: Lawrence Erlbaum Associates, 2004).

police show Lefty a video of Soprano living a pampered life as a child and being spoiled by wealthy parents, then Lefty's motive of compassion may disappear and he may inform. Or the police could use Lefty's compassion to their advantage by introducing him to the families of those who were killed by Soprano. In any case, if Lefty's affective reactions are not linked to any ID-commitments, then they are incoherent or at least contrary to the "wholeness" required for integrity.

In this sub-section I have presented the necessary condition of coherence and argued along with McFall that at least three types of coherence are necessary for a person to have integrity. The first type of coherence is the consistency among a person's commitments and I explained that this consistency does not have to be set in stone, but it must be not be in conflict and unmanageable. The second type of coherence is between commitments and actions. Two signs of incoherence are weakness of will and self-deception defined as believing one has a commitment but acting *contrary* to that commitment. The third type of coherence is between commitments and motivations, and I explained how actions driven by affective reactions with no connection to ID-commitments may correspond to actions that look like integrity but lack the coherence between commitments and motivations to be actions of integrity.

Given the above discussion, I revise the fourth necessary condition to the following: *To have integrity, a person must have coherence among her ID-commitments as well as coherence among her ID-commitments, her actions, and her motivations.*

## 2.5 - Summary of Necessary Conditions

The identity account of integrity posits that a person's identity, which is partly constituted by her ID-commitments, is the core from which a person who has integrity acts. A person with no ID-commitments does not have integrity. In this section I have revised a generic identity account that posits at least four necessary but not sufficient conditions for having integrity defined as wholeness.

The first necessary condition is that the person must have the capacity to form second-order volitions or commitments and the capacity to control her commitments and actions. The second necessary condition is that she must also have the capacity to reason and consider reasons. As I argued earlier, it is not necessary that she actually *have* reasons for her commitments and actions because they may arise from dispositions or other sources. If a person meets the first two conditions, then she is a *candidate* for integrity attributions.

The third necessary condition is that a person must have an ID-commitment at Level 2 or 3 which means that she strongly identifies with a value and has the determination to actualize it. A person who has no ID-commitments has no substantial identity to which an attribution of integrity can refer. Related to these ID-commitments is the fourth necessary condition that a person must have coherence among her ID-commitments as well as coherence among her ID-commitments, her actions, and her motivations. The fourth condition expands on the third condition and requires a person's actions and motivations to be in accord with her ID-commitments.

For the remainder of the dissertation, I take the first two conditions as a given for a person to be a candidate for an integrity attribution. In the upcoming chapters, I analyze the last two conditions in more depth because they involve forming and maintaining a self-concept and acting or refusing to act in accord with one's ID-commitments. For now, I summarize the latter two necessary conditions into one necessary condition with which I begin the next chapter. A person of integrity must have *coherent ID-commitments and the determination to act in accordance with them*.

### **Section 3: Three Objections to the Identity Account**

Most identity theorists conclude that a person who meets the necessary conditions above has what they call *personal integrity*. But it is at this point that the identity account must explain why integrity is most often defined as moral uprightness. Identity theorists explain that to have *moral integrity* a person must meet the four necessary conditions and also have ID-commitments to moral values. Moral integrity requires personal integrity.

Following the identity account to its logical conclusion, McFall, Blustein, Williams, Audi and Murphy, and Taylor all allow a tyrant as well as a hero to have integrity. Some find this conclusion unsatisfactory because as noted above, integrity commonly refers to morally upright people and not tyrants who have coherent ID-commitments to unjust values and the determination to actualize them.

Now that a revised identity account is on the table, I consider three objections to identity accounts in general.<sup>62</sup> The first objection is that the identity account allows a person of integrity to do great good or great evil. The identity account does not say much

---

<sup>62</sup> Calhoun, 1995, 242-246; Cox et al., 2003, 29-36.

about the content of one's commitments, so this is a problem if one thinks integrity fundamentally requires moral content. I address the moral content objection in Chapter 2 by further developing the identity account to include moral content.

The second objection is that many consider integrity to be a moral virtue, but according to the identity account it is merely a formal virtue (i.e. lacking any substance) or not a virtue at all. Williams argues that integrity is not a virtue because it is not a disposition and it is not a formal virtue like courage that would enable the practice of other virtues. In Chapter 3 I respond to these objections and argue that integrity is the disposition to be true to oneself. I also address other concerns regarding the identity account. Three prominent concerns are that the identity account 1) does not seem to allow a person of integrity to change her commitments; 2) does not allow an honest person to lie in order to save someone's life; and 3) does not account for how two people can disagree on a moral issue and both have integrity.

The third objection concerns whether or not we can or should pursue integrity. The first form of the objection is raised by situationists such as Gilbert Harman and John Doris. They present empirical evidence that questions whether or not stable mental constructs, a necessary basis for integrity, actually exist. They strongly suggest that situational factors often overwhelm a person's character and therefore they are skeptical about the existence of stable mental constructs that can guarantee consistent behavior across varied situations.<sup>63</sup> In other words, if stable mental constructs such as character

---

<sup>63</sup> John Doris, *Lack of Character*. (New York: Cambridge University Press, 2002); Gilbert Harman, "Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error," *Proceedings of the Aristotelian Society*, 1999.

are not possible, then integrity which relies on the same constructs may not be possible when confronted with situational factors. The second form of the objection is David Luban's claim that the *quest* for integrity is extremely dangerous because our desire to see ourselves as "morally good" causes us to unconsciously weaken our moral principles to match our unethical conduct. He also argues that our quest for integrity causes us to rationalize our behaviors and deceive ourselves.<sup>64</sup> I respond to the situationist challenge to integrity in Chapter 4 and Luban's concerns in Chapter 5.

Because the identity account lacks rigorous development in certain areas, these objections have led some to reject the account all together. In the remainder of this dissertation I develop and defend an identity account that responds to these objections.

---

<sup>64</sup> David Luban, *Legal Ethics and Human Dignity* (New York: Cambridge University Press, 2007).



## Chapter 2 - Integrity and Morality: Of Forms and Substance

If an alien visitor were to learn about integrity from the media, it would most likely conclude that it is one of the most valued attributes of leaders in the public and private spheres. Political candidates advertise themselves as having integrity while questioning the integrity of their opponents. Universities, corporations, and governments trumpet the importance of integrity as a moral concept for their organizations. Billionaire Warren Buffet is quoted as saying, "In looking for people to hire, you look for three qualities: integrity, intelligence, and energy. And if they don't have the first, the other two will kill you."<sup>1</sup> When leaders and organizations talk of integrity they are using it to identify the quality of being morally upright. In my review of current definitions which I elaborate on below, integrity's most common usage is, "The quality of being honest and having strong moral principles; moral uprightness."<sup>2</sup>

Contrary to this view, the identity account that I presented in Chapter 1 claims that integrity is fundamentally about wholeness and coherence. A person of integrity has coherent identity-conferring commitments (i.e. ID-commitments) and the determination to act in accordance with them. These commitments can be moral, non-moral, or immoral. Because the identity account as presented is silent on the *content* of one's commitments, the main objection it must answer is that identity accounts are incomplete — they do not explain how integrity is commonly used to indicate moral uprightness.

---

<sup>1</sup> "Integrity." Leadershipnow.com. Accessed September 9, 2010.  
<http://www.leadershipnow.com/integrityquotes.html>.

<sup>2</sup> *New Oxford American Dictionary*, (New York: Oxford University Press, 2005).

The purpose of this chapter is to answer the question, “Is integrity a moral concept?” To answer this question I must solve both the intuition problem and the moral integrity problem. The *intuition problem* arises because many have the intuition that both tyrants and heroes can have integrity as long as they have ID-commitments to something. The second problem is the *moral integrity problem* which arises because if integrity is a moral concept, then tyrants cannot have integrity. The challenge facing any account of integrity is to solve both of these problems in a way that explains our diverging intuitions about integrity as well as its common usage as a moral concept.

In the literature, integrity accounts follow one of two strategies to address the two problems. Either they account for the intuition problem and then make adjustments to solve the moral integrity problem, or they account for the moral integrity problem and then make adjustments to solve the intuition problem. In section 1, I explain how identity account defenders solve the intuition problem and then evaluate two attempts to solve the moral integrity problem. In section 2, I review and evaluate two non-identity accounts of integrity that attempt to solve the moral integrity problem first before making adjustments to solve the intuition problem. I argue that while both of these accounts offer important insights about integrity, they still fall short of solving the intuition and moral integrity problems. In section 3 I revise the identity account from Chapter 1 to explain the moral content of integrity and the similarities and differences between tyrants and heroes. I conclude that my revised account solves both problems adequately.

## **Section 1: Reconciling the Identity Account with Morality**

Any integrity account must address why some people have the intuition that a person with an uncompromising ID-commitment to an immoral value seems to have integrity. Identity theorists find this the most interesting problem to solve because they often begin their analysis of integrity by defining it as “wholeness,” from one Latin meaning of the word *integer*. Using wholeness as an anchoring concept, the identity account explains how both tyrants and heroes can have integrity if they have coherent ID-commitments that constitute a whole identity. With the intuition problem solved, however, the identity account must avoid the moral integrity problem which states that integrity, as it is commonly used and defined, cannot describe a tyrant.

Identity theorists use two strategies to solve the moral integrity problem while preserving the gains made by solving the intuition problem. First they insulate integrity from immoral ID-commitments by placing limits on the values and reasoning capabilities a person of integrity can have. Second, they distinguish between personal integrity and moral integrity to solve the intuition problem without totally abandoning integrity as a moral concept. I first consider the strategies to insulate integrity from immoral ID-commitments and then the personal/moral integrity solution.

### **1.1 - Insulating Integrity from Immorality**

Most identity theorists conclude that immoral people can meet the necessary conditions for having integrity as described in Chapter 1.<sup>3</sup> Some theorists, however, want

---

<sup>3</sup> Philosophers who argue that *personal integrity* is fundamentally a non-moral concept are Bernard Williams, Lynne McFall, Gabriel Taylor, and Jeffrey Blustein. See Bernard Williams and J.J.C. Smart. *Utilitarianism: For and Against*. (New York: Cambridge University Press, 1973) and *Moral Luck*. (New York: Cambridge University Press, 1981); Lynne McFall. “Integrity.” *Ethics*, Vol. 98, No. 1 (1987);

to maintain some moral meaning for the notion of integrity without restricting integrity to morally upright people. Identity theorists have sought to insulate integrity as wholeness from immorality by restricting the values to which one may commit and by requiring a high standard of rational thinking. I consider these attempts in order.

When it comes ID-commitments to values, Lynne McFall argues that a person cannot have integrity if she is committed to the values deemed most inappropriate for attributions of integrity: pleasure, acquiring wealth, and seeking approval.<sup>4</sup> She does not argue against these values on moral grounds but rather on the grounds that integrity by definition precludes committing to these values.<sup>5</sup> For example, what if a Mafia wise-guy refuses to inform on his boss during a harsh interrogation, not because he is committed to fidelity, but because he values his boss' approval above all else? Or what if a rescuer of Jews during World War II will not turn in his comrades to the Gestapo, not because he is committed to the value of protecting the innocent, but because he values the huge rescue fees he is receiving? McFall claims that most people would say that these people lack integrity because the values they hold are inappropriate for attributions of integrity by definition.<sup>6</sup>

In addition to her argument from definition, McFall uses the *reasonable person standard* to determine which values are integrity-appropriate. She states that a person of

---

Gabriele Taylor, "Integrity." *Proceedings of the Aristotelian Society*, Supplementary Vol. 55 (1981) and *Pride, Shame and Guilt: Emotions of Self-Assessment*. (Oxford: Oxford University Press, 1985); Jeffrey Blustein. *Care and Commitment: Taking the Personal Point of View*. (New York: Oxford University Press, 1991).

<sup>4</sup> McFall, 1987, 11.

<sup>5</sup> *Ibid.*, 1987, 11.

<sup>6</sup> She takes her definition from *Webster's Third New International Dictionary*. McFall, 1987, 11.

integrity has an ID-commitment to values that we must at least recognize, “as ones a reasonable person might take to be of great importance and ones that a reasonable person might be tempted to sacrifice to some lesser yet still recognizable goods.”<sup>7</sup> If a reasonable person takes a value to be of great importance and/or one that she may be tempted to sacrifice, then it is an integrity-appropriate value.

McFall’s value standard, however, seems too high if the central condition for having integrity is having ID-commitments. It seems that if the wise-guy and the rescuer have ID-commitments to *any* values, then they have integrity regardless of what a reasonable person thinks about the values. The seducer Don Giovanni as portrayed in Mozart’s opera of the same name is a clear counter-example to McFall’s claim. In the last scene he is offered the opportunity to repent of his ID-commitment to pleasure or face the fires of Hell. He boldly declares that he will not repent even as dark images from Hell surround him. Following McFall’s standard, Don Giovanni’s ID-commitment to sensual pleasure would most likely *not* be one that a reasonable person would take to be of such great importance that he would rather be tortured for eternity than give it up. Yet, on the identity account it is not a stretch to attribute integrity to Don Giovanni because he has a coherent ID-commitment to a value and is determined not to compromise it.

Jeffrey Blustein similarly critiques McFall’s position when he states, “Sometimes we grant or deny integrity based on our own conceptions of importance, and sometimes we see integrity involved in an act on the basis of the role some commitment plays in a

---

<sup>7</sup> Ibid., 1987, 11.

person's life.”<sup>8</sup> If a person such as Don Giovanni has an ID-commitment to the value of pleasure and seems to have integrity, then having an ID-commitment to values like pleasure, artificiality, and self-seeking may not automatically disqualify a person from having integrity. On the identity account of integrity, the *type* of values do not carry as much weight in integrity attributions as the ID-commitment to them. Even though McFall attempts to move integrity away from immoral values by definition and a reasonable person standard, she must still explain how Don Giovanni can seem to have integrity given his supposed ID-commitments.

Note that I am not arguing that the “wealth-seeking” rescuer mentioned above has integrity if he decides on a whim that acquiring wealth is the best reason for enduring torture. He only has integrity on the identity account if wealth is the object of an ID-commitment that he is determined to actualize. As mentioned in Chapter 1, if he acts from a transient motive to acquire wealth (e.g. he wants the money for a one-time purchase) or from an affective motive such as greed (e.g. he has a desire to acquire wealth) that is not from an ID-commitment, then he does not have integrity. In both cases, he does not meet the necessary condition of having a coherent ID-commitment and the determination to actualize it.

A similar attempt to insulate integrity from immorality is Gabriel Taylor's claim that a person cannot have integrity unless she takes the interests of others into account. According to Taylor, a person can have integrity and be immoral, but she must at the same time consider her impact on others. Taylor states, “If she [a person of integrity] is

---

<sup>8</sup> Blustein, 1991, 123.

to have a sane view of herself and the life she wants to lead she cannot ignore the evidence of her impact on others and their reactions to her.”<sup>9</sup> She also states that, “At least the person of integrity cannot be a moral solipsist, for she will recognize that others . . . act on reasons which they regard as justificatory.”<sup>10</sup> She “must therefore give some recognition to others as persons who have views and interests and intentions of their own.”<sup>11</sup>

I interpret Taylor as presenting a minimum and a maximum condition for recognizing the interests of others. I agree with a minimum condition that a paradigmatic person of integrity must not be disconnected from the reality of who she is and the lives of other people. I noted a similar point in Chapter 1 when I argued that a candidate for integrity must have the *capacity* to understand reasons. I consider this a given for any candidate of integrity. But I disagree with a maximum condition which would require a person of integrity to recognize other people as having needs and intentions to such an extent that she would be inconsistent if she were to act against their interests for her own satisfaction.

If we go back to the Don Giovanni example, one that Taylor also cites as an example of integrity, it seems that he meets the minimum condition because he is not disconnected from the reality of his seductions and the people who fall for them. This makes him even more effective in his quest. But Don Giovanni does not meet the maximum condition. He does not consider the intentions and interests of other people as

---

<sup>9</sup> Taylor, 1985, 128.

<sup>10</sup> Ibid., 1985, 128.

<sup>11</sup> Ibid., 1985, 128.

having the same weight as his own, nor would the identity account require him to hold their interests in such a way. Other people are simply objects in his game and he is unwilling to sympathize with them.

Taylor's position on integrity seems equivocal.<sup>12</sup> She allows Don Giovanni to have integrity, but she later establishes a requirement of considering the rights and interests of others before acting. In her attempts to insulate integrity from immorality, she re-exposes herself to the intuition problem. Taylor's attempt to use rationality as a moral firewall is not unique to discussions of integrity. But being rational does not provide moral content to the identity account. A tyrant can have consistent and logical reasons for acting unjustly just as a hero can have consistent and logical reasons for acting justly.

To summarize, McFall's restriction on integrity-appropriate values does not apply because a person with an ID-commitment to a value such as pleasure seems to have integrity on a strict identity account. Also the "reasonable person" and sympathetic person standards for having integrity are not necessary to have integrity as wholeness. Whether insulating integrity from immorality is effective or not, identity theorists offer another solution to the moral integrity problem. They argue that two kinds of integrity exist: personal integrity and moral integrity.

## **1.2 - The Personal Integrity/Moral Integrity Distinction**

A person who has coherent ID-commitments to any values at all has *personal integrity* and she may have *moral integrity* if some of her ID-commitments are to moral

---

<sup>12</sup> Damian Cox, Marguerite La Caze, and Michael P. Levine make a similar point regarding Taylor's view. *Integrity and the Fragile Self*. (Aldershot: Ashgate, 2003), 67.



values. On McFall's account, moral integrity requires personal integrity. She argues that, "If one held personal but no moral principles, then one could have personal integrity without moral integrity . . . Thus moral integrity presupposes personal integrity."<sup>13</sup>

Blustein makes a similar distinction when he notes that a sadist and a tyrant can have personal integrity if they act according to their commitments, but it is obvious that they do not have moral integrity.<sup>14</sup> If the sadist and tyrant repent of their ways and take on an ID-commitment to respect others, then they would have personal integrity as well as moral integrity.

The personal/moral distinction attempts to solve the moral integrity problem while explaining why we have the intuition problem. My objection to this solution is that by positing personal integrity as a necessary condition for having moral integrity, it elevates personal integrity to the same level as what I would characterize as genuine integrity (i.e. moral integrity). The identity account leads theorists to this distinction because it uses ID-commitments to solve the intuition problem and then must determine how to add moral integrity into the account. The personal/moral distinction, however, ignores or at least subordinates common usage for conceptual simplicity.

As demonstrated by the multiple dictionary definitions presented in Chapter 1, common usage since the early 1800's emphasizes integrity as a moral term while integrity as "wholeness" primarily refers to non-human entities such as states and agreements. Forwarding to today, the preeminence of integrity as a moral term is also

---

<sup>13</sup> McFall, 1987, 16.

<sup>14</sup> Blustein, 1991, 123.

supported by dictionary usage measures as well as other sources. In July 2010 I searched Amazon.com for books with “integrity” in the title and found that 60 of the first 101 books listed focused specifically on moral behavior while a substantial number of the other 41 books deal with being true to one’s ID-commitments to specific moral values. I also considered the field of business scholarship where scholars use the term often. In an extensive literature review of business articles and books that use the word integrity, Michael Palanski and Francis Yamarrino found that twenty separate sources use integrity to indicate a particular moral virtue such as honesty or morality in general.<sup>15</sup>

A recent psychology study bears out my contention that most people associate integrity with moral commitments. Psychologists Lawrence Walker and Karl Hennig undertook a study examining people's conceptions of moral exemplarity, in particular their conceptions of the prototypical just person. They asked 131 participants to rank the attributes that most accurately describe a just exemplar on a scale from (1) (*extremely inaccurate*) to (8) (*extremely accurate*). Out of 113 attributes that were ranked, the attribute “has integrity” scored a 7.15 on the 8-point scale and was the 7<sup>th</sup> highest descriptor of a just person.<sup>16</sup> While not conclusive, this data shows a strong tendency for people to associate integrity with just people and not only “whole” people.

Taking a step further into common usage, consider how a non-philosopher would interpret the following sentence:

A: Mary is a person of integrity.

---

<sup>15</sup> Michael E. Palanski and Francis J. Yammarino. “Integrity and Leadership: Clearing the Conceptual Confusion.” *European Management Journal* (June 2007).

<sup>16</sup> Lawrence J. Walker and Karl H. Hennig, “Different Conceptions of Moral Exemplarity: Just, Brave, and Caring,” *Journal of Personality and Social Psychology* Vol. 86 No. 4. (2004): 646.

When most non-philosophers read sentence A, they interpret it as stating that Mary is morally upright because when the term “integrity” stands alone it simply describes her as a morally trustworthy person. To make my point even clearer note how the following sentence sounds redundant: “Mary is a person of integrity, and she is morally upright.” It is redundant because the second part adds nothing to the first part. Also consider the sentence, “Mary lacks integrity.” Again, most people would take the sentence as saying that Mary is *not* morally upright and this common interpretation is based on an understanding that integrity is a moral concept.

If the personal/moral distinction is accurate, sentence A means that Mary has some coherent ID-commitments, the determination to actualize them, and some of these ID-commitments may or may not be moral. But most non-philosophers would rarely if ever interpret the sentence as leaving open the possibility of moral uprightness. From the perspective of common usage, the personal/moral distinction does not exist and possibly cannot exist. The gap between the distinction and common usage reveals a problem with proposing a personal integrity that is completely unrecognizable from moral integrity.<sup>17</sup>

Now consider Sentence B1 in reference to common usage.

B1: Jerry is a person of integrity. He is consistently dishonest, regardless of the consequences.

---

<sup>17</sup> Some philosophers disagree with the distinction from the point of view that integrity does not require any moral content and requiring moral content only complicates the concept. Palanski and Yammarino take this route and define integrity with the non-moral definition “consistency of an acting entity’s words and actions” (Palanski and Yammarino, 2007, 178). They argue that more specific terms such as honesty, authenticity, and courage should be used instead of integrity depending on the speaker’s meaning. John Bigelow and Robert Pargetter argue that integrity is the capacity to exercise strength of will. While they claim that integrity in most contexts is inherently good, they avoid giving it any substantive moral content. Both of these accounts, while appealing to the benefits of simplicity, leave unanswered the question of what integrity means in common usage and the moral integrity problem. John Bigelow and Robert Pargetter, “Integrity and Autonomy,” *American Philosophical Quarterly*, Vol. 44 No. 1. (2007).

Sentence B1 is different from A because the second part of the sentence seems to contradict the first part. Most speakers to whom I have presented similar sentences deny Jerry is a person of integrity, but with some explanation some can see a vague connection between integrity and consistency. To understand B1, speakers must adjust the meaning of the word “integrity” to match the context provided in the second sentence. The adjustment requires they must abandon integrity as a moral concept and instead take integrity to mean having a steadfast commitment to something else. To demonstrate this adjustment, note how sentence B2 adds nothing new to B1.

B2: Jerry is a person of integrity. He is consistently dishonest, regardless of the consequences. He won't compromise his commitment to dishonesty.

Sentences A, B1, and B2 reveal that integrity has at least two meanings that depend on the context in which it is used. Integrity fundamentally means moral uprightness, but contextual features can change it to mean merely having steadfast commitments.

We can use these insights to solve the intuition problem while coming closer to solving the moral integrity problem. If the identity account as presented in Chapter 1 has identified necessary conditions for having integrity, then a person has integrity only if she has coherent ID-commitments and the determination to actualize them. This means that tyrants and heroes can both have some form of integrity. However, given the meaning of integrity with and without contextual cues, I recommend that the term *substantive integrity* refer to people with ID-commitments to moral values (e.g. Weidner) and the

term *formal integrity* refer to people with ID-commitments to non-moral or immoral values (e.g. Don Giovanni).

By distinguishing between substantive and formal integrity, the identity account can explain how genuine integrity matches our common usage and also how some people may have a form of integrity. For example, in line with common usage, a moral exemplar can have substantive integrity which means he is a person whose identity is constituted by coherent ID-commitments to moral values. Similarly, Don Giovanni has formal integrity which means he is a person whose identity is constituted by coherent ID-commitments to non-moral or immoral values. I elaborate on what it means to commit to moral, non-moral or immoral values in section 3.

One may ask what we gain by moving from the personal/moral distinction to the substantive/formal distinction. The new distinction provides two benefits over the old distinction. First, it recognizes that integrity attributions are primarily moral attributions and this recognition correlates with common usage. Instead of so-called moral integrity being an add-on to personal integrity, integrity with moral content becomes central and acknowledges that anything called personal integrity is merely a form or shadow of substantive integrity. A second benefit of the distinction is that it maintains gains made in solving the intuition problem while moving us closer to solving the moral integrity problem. The personal/moral distinction implies that to have integrity in general one only needs coherent ID-commitments. While this strategy solves the intuition problem by accounting for moral and immoral agents, it makes the moral integrity problem more difficult to solve because now both tyrants and heroes can have integrity.

The substantive/formal distinction, on the other hand, acknowledges that integrity is primarily a moral concept and that a person of integrity is a person with ID-commitments to moral values. At the same time the distinction recognizes that integrity can sometimes refer to tyrants because they share something in common with morally upright people. I am not saying that the personal/moral distinction is useless, but that the substantive/formal distinction better classifies the two uses of integrity without subordinating common usage for conceptual simplicity. I do not jettison the personal/moral distinction, but instead use it in section 2 to distinguish between commitments and values.

In this section I have presented identity theorist attempts to insulate the identity account from claims that it lacks moral content. I argued that these replies leave a gap between the identity account and common usage. I then argued that by distinguishing between substantive integrity and formal integrity we can explain our diverging intuitions regarding tyrants and heroes while anchoring the account to common usage.

Before revising the identity account presented in Chapter 1 to include the notions of substantive and formal integrity, I review two prominent non-identity accounts of integrity that seek to solve the moral integrity problem in the shadow of the intuition problem.

## **Section 2: Non-Identity Accounts of Moral Integrity**

In this section I continue answering the question, “Is integrity a moral concept?” by reviewing and critiquing two non-identity integrity accounts (i.e. accounts in which having coherent ID-commitments is not sufficient and may not be necessary to have

integrity). The first account argues that integrity is a social virtue that enables co-deliberators to work together. The second account argues that integrity is primarily a moral concept even though our intuitions are often split about tyrants having integrity. I label the two accounts the deliberation account and the moral trustworthiness account.

## 2.1 - The Deliberation Account of Integrity

Cheshire Calhoun's influential account of integrity appeals to a notion of integrity that requires a person, as a community member, to stand up for what she judges is best.<sup>18</sup> For Calhoun, having and protecting one's identity may play a part in a person having integrity, but it is not necessary. Instead, integrity is primarily a social virtue that requires a person to see herself as part of a community rather than as a lone individual trying to maintain her identity. Calhoun argues that integrity is a social virtue because when we "stand for" what is our best judgment we do so in a context of other people. Broadly stated, Calhoun argues that integrity is a social virtue that fits us for community membership as co-deliberators who are trying to answer some form of the question, "What is worth doing?"<sup>19</sup> Because each deliberator comes to the question with only her own view points, all she can offer to the other deliberators is her best judgment. Calhoun states,

As one among many deliberators who may themselves go astray, the individual's judgment acquires gravity. It is, after all, not *just* her judgment about what would be wrong or not worthwhile to do. It is also her *best* judgment. Something now hangs for all of us, as co-deliberators trying to answer correctly the 'What is worth doing?' question, on her sticking by her best judgment. Her standing for something is not just

---

<sup>18</sup> Cheshire Calhoun, "Standing for Something." *The Journal of Philosophy*. Vol. 92, no. 5 (May 1995).

<sup>19</sup> Calhoun, 1991, 257. Calhoun does not elaborate on why this question in particular is relevant to a community.

something she does for herself. She takes a stand for, and before all deliberators who share the goal of determining what is worth doing.<sup>20</sup>

Requiring a person to stand up for her best judgment is what Jody Graham calls *taking epistemic responsibility*.<sup>21</sup> Standing up for her best judgment matters because it serves the common goal of all co-deliberators. Calhoun states, “Persons of integrity treat their own endorsements as one’s that matter, or ought to matter to fellow deliberators.”<sup>22</sup> Calhoun supports her claim with the common intuition that a person lacks integrity if she gives up her best judgment too easily in the face of criticism or because of external pressures. To have integrity, a person must demonstrate her epistemic responsibility to stand up for her best judgment as a fellow co-deliberator. If she stands up for her best judgment, she fulfills her epistemic responsibility and has integrity.

One must be careful, however, when standing up for one’s best judgment because a person who digs in her heels too much may lack integrity. Calhoun argues that arrogance, close-mindedness, and other traits of fanaticism and dogmatism typically demonstrate a lack of integrity. “All [these traits] reflect a basic unwillingness or inability to acknowledge the singularity of one’s own best judgment and to accept the burden of standing for it in the face of conflict.”<sup>23</sup> For a person to have integrity, she must also have *epistemic humility*. When we stand for our best judgment about what is

---

<sup>20</sup> Calhoun, 1991, 257.

<sup>21</sup> Jody L. Graham. “Does Integrity Require Moral Goodness?” *Ratio*. XIV, No. 3. September 2001. 243.

<sup>22</sup> Calhoun, 1991, 258.

<sup>23</sup> Calhoun, 1991, 260.



worth doing, integrity requires us “simultaneously to stand behind our convictions and to take seriously others’ doubts about them.”<sup>24</sup>

Calhoun’s account of integrity offers two helpful insights about integrity. First, it branches integrity out from a mere defense of one’s identity and into the realm of the community. While identity accounts in general have not developed a community component, Calhoun’s account recognizes that integrity is an important community-relevant virtue.<sup>25</sup> The second insight is that dogmatism and closed-minded stubbornness can demonstrate a lack of integrity. Her requirement goes beyond the necessary condition that a person be capable of understanding the reasons presented by others. It seems that a person of integrity must be *open* to consider the reasons others give.

While Calhoun’s account is helpful in understanding these aspects of integrity, it does not resolve the moral integrity problem because it does not specify any moral content that a person of integrity must have. Instead her account offers appropriate epistemic attitudes for discussing, “What is worth doing?” which could be a moral or non-moral undertaking. I agree that her account correctly identifies the intuition that a person lacks integrity if she surrenders her epistemic responsibility (e.g. gives up her best

---

<sup>24</sup> Calhoun, 1991, 260. An account similar in content is presented by Cox et. al, 2003. They argue that, “Integrity is a virtue located at the mean of various excesses. On the one side we find conditions of capriciousness, wantonness, weakness of will, disintegration, hypocrisy, dishonesty, and an incapacity for reflection or self-understanding. On the other side we find conditions of fanaticism, dogmatism, monomania, sanctimoniousness, hyper-reflexivity and the narrowness and hollowed out character of a life closed off from the multiplicity of human experience” (15). In other words, integrity is the mean between being too open minded and uncommitted on one side and being dogmatically closed minded on the other side. I address their account in Chapter 3.

<sup>25</sup> Larry May also emphasizes the importance of community when describing integrity. He presents a communitarian account of integrity that accounts for social influences on developing an identity and how a person can change her identity. I consider some of his insights in Chapter 3. Larry May, “Integrity, Self, and Value Plurality.” *Journal of Social Philosophy*, Vol. 27 No. 1. (Spring 1996).

judgment too easily) too easily or does not have enough epistemic humility (e.g. dogmatically holds to her ideas too tightly). However, having these two characteristics does not necessarily result in an attribution of substantive or moral integrity.

The missing link between Calhoun's account and substantive integrity is the link between one's best judgments and the standards for determining which of these are the best *moral* judgments. Graham vividly exposes this problem using the example of the 1930's Tuskegee Study in which 399 African American men from Alabama were diagnosed with syphilis and purposely not treated. The group of doctors and scientists responsible for this study observed the disease as it progressed and the study was reviewed up into the late 1960's with the results published in *The Journal of the American Medical Association*. Graham observes that a group of highly educated co-deliberators presented their best judgment about, "What is worth doing?" for 30 years, and yet they continued down an immoral path. She states,

While it is reasonable to assume that some of the members were close-minded, arrogant, and deaf to criticism, many of them were not. Unfortunately, the collective decision-making failed morally, and my intuition is that none of the participants were persons of integrity; they were not persons of integrity for the very reason that they were unable to see the disrespect paid to the Tuskegee men participating in the study.<sup>26</sup>

The reason Calhoun's integrity account does not directly address morality is that demonstrating epistemic responsibility and humility are consistent with immoral actions. Standing up for one's best judgment to the right degree while being open-minded does not necessarily result in moral values or moral action. Graham observes that, "open-mindedness, conceptual clarity, and logical consistency, or any other internal individual-

---

<sup>26</sup> Graham, 2001, 245.

based constraints, while often indicative of good character are not guarantors of it.”<sup>27</sup> In all fairness, Calhoun is only indirectly describing the content of moral integrity. She is answering the question, “What is a person who tries to have integrity trying to do?” Her answer is, “trying to stand for what, in her best judgment, is worth persons doing.”<sup>28</sup> While her account has some moral implications for members of the community, it does not clearly explain the moral content that common usage expects.

## 2.2 - Integrity as Moral Trustworthiness

After Graham critique’s Calhoun’s account using the Tuskegee Study, she develops her own account of moral integrity based on the common intuition that a person must respect others to be a person of integrity. She begins her account by explaining why integrity as epistemic responsibility and epistemic humility is incomplete. Her concern is that a person can demonstrate both of these characteristics and have no experience or track record in successfully navigating quandaries. It seems that a typical person of integrity must be *good* at knowing which best judgment is *the* best judgment. In other words, she must be *epistemically trustworthy*. In general a person who is trustworthy takes the interests of others into consideration, and a person who is epistemically trustworthy has the trait of respecting the views and opinions of others. But epistemic trustworthiness is still not sufficient to have moral integrity. Referring to those Tuskegee physicians who were epistemically trustworthy, Graham states, “All, individually and most certainly collectively, failed to take the lives of the subjects seriously, and without

---

<sup>27</sup> Graham, 2001, 246. Mark Halfon presents an account of integrity that also relies on epistemic trustworthiness to insinuate some moral content. Mark S. Halfon, *Integrity: A Philosophical Inquiry* (Philadelphia: Temple University Press, 1989).

<sup>28</sup> Calhoun, 1995, 260.

this moral compass, no one can be said to be deserving of the admiration that is signaled by justified attributions of integrity.”<sup>29</sup>

Graham’s intuition is that we are not justified in attributing integrity to someone as a form of admiration unless she has some moral virtues. For example, the Tuskegee physicians appear to have some *form* of integrity because they have epistemic trustworthiness, and being trustworthy implies respect for *all* others and not just those within a particular community. But the physicians were lacking *moral trustworthiness*. In other words, they could not be trusted to respect the humanity within each individual. According to Graham and following on my discussion in section 1, our intuitions confirm that most attributions of integrity presuppose moral trustworthiness and the Tuskegee physicians lack substantive integrity for this reason. The implication of this conclusion is that because tyrants are not morally trustworthy, they cannot have genuine integrity and the moral integrity problem is solved.

Having solved the moral integrity problem, Graham must solve the intuition problem. How does she explain attributions of integrity to tyrants or the Tuskegee Study physicians who were committed and determined to learn more about a debilitating disease regardless of the human cost? Graham acknowledges that language allows us to attribute integrity to tyrants, but questions what this tells us about the nature of integrity. Graham explains that we can condemn an act or kind of act and not the person performing it. For example, we may condemn the *act* of bombing a warehouse in South Africa to protest apartheid while not disapproving of the person who did it. In this way

---

<sup>29</sup> Graham, 2001, 245.

the activist has integrity as moral trustworthiness but we still disapprove of the kind of act.<sup>30</sup>

The identity theorist may respond that we can attribute personal integrity to a person like Don Giovanni even though we disapprove of his acts as well as his moral judgment in the past. Graham rejects this response. For the sake of clarity, I extensively quote her suggested alternative:

If we are tempted to ascribe integrity to Don Giovanni we are so tempted because we fail to mind the distinction between actions done in accordance with virtue and actions that result from the possession of it. . . . Once we mind this distinction, then we can make sense of our ascription of integrity based solely on a particular performance or one aspect of character rather than on whole character assessments. . . Such [morally questionable] characters manifest behavior of the virtuous person without being virtuous themselves.

Graham argues that acting with resolve like a virtuous person or even having the trait of being committed to one's principles is not enough to be a person of integrity as we commonly use the term. To say that Don Giovanni has integrity is to say that he manifests a trait that is an aspect of integrity. But if integrity requires a person to be morally trustworthy, then it is incompatible with people like Don Giovanni who are morally untrustworthy. Therefore, tyrants cannot have integrity and the only reason the intuition problem exists is because we do not distinguish between a person's full character and her acts or aspects of character. If we made these distinctions, then it would be false to say Don Giovanni is a person of integrity.

I find Graham's arguments convincing because they build on the same intuitions that led to the introduction of the formal/substantive distinction in section 1. Her overall

---

<sup>30</sup> Graham, 2001, 249.

account, however, leaves two important questions unanswered about the character or identity of the person who has or lacks integrity. In regards to integrity as moral trustworthiness, she correctly notes that attributing integrity to a person is to recognize that she is morally trustworthy. However, Graham does not describe what constitutes a morally trustworthy person's character. Without this explanation we are left with an intuitively correct answer to the moral integrity problem but little understanding of what makes a person morally trustworthy. Given this lack of explanation the first question that needs to be answered is, "What is the content of moral trustworthiness?"

Regarding Graham's response to the intuition problem, I agree that if attributions of integrity are attributions of moral trustworthiness, then attributing integrity to tyrants must be mistaken. I am also convinced that she is correct in distinguishing between the acts and traits that are in *accordance* with a virtue and the person who actually has the virtue. But Graham does not explain how a person acts in accordance with the virtue while not having it. When Don Giovanni demonstrates his commitment to a life of seduction, Graham says he is steadfast like a person who has integrity. But where does his steadfastness come from and does it differ from the steadfastness of the person of integrity? Also, what aspects of character do tyrants have that produces the intuition problem? The second question that needs to be answered is, "What features do the tyrant and the hero share and what features make them fundamentally different?" To understand our intuitions about integrity, this question needs to be answered.

In the next section I revise the identity account to answer these two questions and integrate insights from sections 1 and 2.

### **Section 3: The Moral Identity Account**

The two questions left by the moral trustworthiness account are, “What is the content of moral trustworthiness?” and “What features do the tyrant and the hero share and what features make them fundamentally different?” These questions are not necessarily new because the first one goes to the heart of the moral integrity problem and the second aims to explain the moral intuition problem. In this section, I revise the identity account presented in Chapter 1 and argue that it best solves the two problems while answering the two open questions. I first show how the moral trustworthiness account is effective in so far as it relies on the identity account’s necessary conditions. Second, I describe the features that substantive integrity and formal integrity share and the main differences between them. Finally, I argue that the revised identity account best solves the intuition problem and moral integrity problem.

#### **3.1 - Moral Trustworthiness and ID-Commitments**

In section 1 I presented attempts to adjust the identity account to solve the moral integrity problem and argued that all fell short. I then suggested that a distinction should be made between substantive integrity which requires ID-commitments to moral values and formal integrity which requires ID-commitments to non-moral or immoral values. I now consider the content of these moral values to which a person can commit and use Graham’s concept of moral trustworthiness as a springboard for the discussion.

To understand Graham’s concept of moral trustworthiness, we must first understand what she means by “moral” and “trustworthy.” According to Graham, to be moral, “one respects, first and foremost, the humanity in each individual regardless of the

individual's relation to oneself. One must see individuals as having worth because of an equal standing, not to be lessened by sacrificing one individual for the good of others."<sup>31</sup> Graham explains that a person of integrity judges and acts from the broad notion of equal respect. A person of integrity is also *trustworthy*. In contrast to a person who is just reliable, "The trustworthy person has a regard for equal treatment, not simply loyal treatment. There is a flexibility and moral know-how – a depth of character – that need not be present in the reliable person."<sup>32</sup> To rephrase these definitions slightly, a morally trustworthy person has a *character trait of equally respecting other people*.

Socrates is an example of a person who fits this description of moral trustworthiness. In a story he relates to his jury, Socrates explains that he was told by a rogue government to bring Leon of Salamis to be executed. Because Socrates believed the charges against Leon were unjust, he disobeyed the government command to deliver Leon and went home knowing that his action could lead to his own death. Socrates states, "Then I showed again, not in words but in action, that, if it were not rather vulgar to say so, death is something I couldn't care less about, but that my whole concern is not to do anything unjust or impious."<sup>33</sup> The decision to go home was easy for Socrates because he had a character trait of not doing anything unjust. It appears that Socrates was morally trustworthy on this account and was a person of integrity.

---

<sup>31</sup> fn. 23. Graham, 2001, 247.

<sup>32</sup> fn. 22. Graham, 2001, 247.

<sup>33</sup> Plato, *The Apology* in *The Trial and Death of Socrates*. Trans. G. M. A. Grube. (Indianapolis: Hackett Publishing Company, 2000): 32d.



If moral trustworthiness is sufficient for moral integrity, we may be tempted to jettison the identity account. However, the question “What is the content of moral trustworthiness?” is still partially unanswered. Graham’s account suggests that it is a character trait of equally respecting other people. But what is a character trait and how does it relate to attributions of integrity? The identity account answers this question. According to the identity account, a person’s character or identity is constituted by her ID-commitments. Both Graham’s account and the identity account require a person of integrity to have a stable character, and the identity account explains that a stable character can be constituted by coherent ID-commitments to values and the determination to actualize them. When describing a person of integrity’s stable character, the identity account offers a more detailed explanation than the moral trustworthiness account.

The identity account, however, still lacks an account of morality. My revisions from section 1 only state the intuitive proposition that a person of substantive integrity has ID-commitments to moral values, but I did not describe the content of these moral values. The moral trustworthiness account defines morality as “equal respect for others.” While this definition seems intuitively correct, it does not explain why these moral values are important or why we are justified in admiring people who are committed to them and condemning those who are not. Another question arises about which “others” one must equally respect. Graham means all other humans, but can a Mafia wise-guy in a closed community have integrity if he shows equal respect to only other Mafia members? To understand moral values and integrity, we need to account for how community memberships and values within those communities contribute to the intuition problem.

In what follows I describe the values to which people have ID-commitments and how they determine integrity attributions.

### **3.2 - Moral Values and Personal Values**

A plausible account of integrity needs to solve the intuition problem and the moral integrity problem, and I have argued that the accounts above either miss the target or come close but lack important details. To explain how an identity account can solve these problems, I use three putative cases of integrity to draw out our understanding of integrity and then identify the common and distinct characteristics of each agent.

Case 1: Don Giovanni is committed to seducing women. He refuses to compromise his commitment even when threatened with severe punishment.

Case 2: Joey Scar is a Mafia wise-guy who always keeps his word within the crime organization, even if it upsets his fellow criminals. He does lie to those outside of the crime syndicate when it will help the Mafia.

Case 3: Socrates refuses to obey a rogue government and bring Leon of Salamis to Athens for execution because to do so would be unjust. Because Socrates would not allow himself to do anything unjust, even to a stranger, he went home knowing his action could result in his own death.

These three agents have two characteristics in common which justifies us in attributing some form of integrity to them. Following the identity account as presented in Chapter 1, the first common characteristic is that all three agents meet the necessary condition of having a coherent ID-commitment to some value and the determination to actualize that value. As a reminder, to have an ID-commitment is to wholeheartedly identify with a value and reject competing values. The second common characteristic is that all three agents refuse to compromise their identities. It appears that they desire to be

true to themselves even when this may cause personal pain or loss.<sup>34</sup> By “themselves,” I am primarily referring to their ID-commitments that constitute their identities. Because discussing this disposition in detail would take me away from the current task, I save my analysis of “being true to oneself” for Chapter 3. In Chapter 3 I argue that this disposition is a necessary condition for being a person of integrity.

Given these two common characteristics, all three agents appear to have integrity because they have coherent ID-commitments to some values and are motivated to be true to themselves. Regardless of the moral content of a person’s values, we tend to admire these two characteristics because they are associated with trust and reliability, two important characteristics for communities. Having ID-commitments and being true to oneself opens the door for an integrity attribution, but these two characteristics solve the intuition problem at the expense of the moral integrity problem. I next consider the substantive differences among the three agents to investigate the moral integrity problem.

It is at this point where the personal/moral distinction can help distinguish between kinds of *values* rather than kinds of integrity. One clear distinction between Case 1 and Cases 2 and 3 is that Don Giovanni’s actions are in accordance with his *personal values* while Scar’s and Socrates’ actions are in accordance with some type of *moral values*.<sup>35</sup> Broadly speaking, personal values are those that a person commits to for

---

<sup>34</sup> Larry May notes that “being true to oneself” is a common way of characterizing integrity (May, 1996, 124) and John Kekes defines integrity as “being true to oneself.” John Kekes, *Enjoyment*, (New York: Oxford Press, 2008).

<sup>35</sup> I am using the term *values* broadly to mean a value, desire, belief, principle, virtue, or a project that is important to the person. The personal/moral values distinction is made by Damian Cox, Marguerite La Caze, and Michael P. Levine, “Integrity”, *Stanford Encyclopedia of Philosophy*. Updated August 10, 2008. (<http://plato.stanford.edu/entries/integrity/>).

her own reasons and that she has no community obligation to actualize. Personal values are justified as important by personal reasons, needs, and expectations. Examples of personal values include becoming a chess Grandmaster, belief in the superiority or inferiority of racial groups, capitalist or socialist principles, the importance of stamp collecting, and the principle that one must “follow the evidence wherever it leads.”

I define moral values as those values that the moral community expects its members to practice and that each member has an obligation to actualize.<sup>36</sup> The idea of being a member of a moral community has been developed in the writings of Peter Strawson and most recently Stephen Darwall.<sup>37</sup> Darwall argues that our moral obligations and responsibilities to others are in force from the perspective of the moral community because other members of the moral community expect us to act according to moral values. Henry Sidgwick similarly observed that, “the most important part of the function of morality consists in maintaining habits and sentiments which seem necessary to the continued existence, in full numbers, of a society of human beings under actual

---

<sup>36</sup> Even though I separate these values for explanatory purposes, I acknowledge that moral values like honesty can become personal values and personal values such as faith may help the community function and survive.

<sup>37</sup> P.F. Strawson, “Freedom and Resentment”, in *Studies in the Philosophy of Thought and Action*. (New York: Oxford University Press, 1968). Stephen Darwall, *The Second Person Standpoint: Morality, Respect, and Accountability*. (Massachusetts: Harvard University Press, 2006).

circumstances.”<sup>38</sup> Examples of these moral values include honesty, fidelity, keeping one’s word, justice, and respecting each other.<sup>39</sup>

The main difference between moral values and personal values is the justification for why these values are important for the moral community and not easily dismissed. Moral values are important because they are justified by both instrumental and moral reasons that directly relate to the functioning of the moral community.<sup>40</sup> For example, moral values are important because the moral community has an instrumental need for reliable negative action (e.g. we won’t lie to each other) and positive action (e.g. we will help each other in emergencies) for it to function and survive.<sup>41</sup> James Wallace has also argued that community-oriented character traits or *conscientiousness*, which I loosely

---

<sup>38</sup> Sidgwick also argued that this regulatory role of morality in no way exhausted “our ideal of good or desirable human life.” Henry Sidgwick. “Essays on Ethics and Method.” Ed. Marcus G. Singer. (New York: Oxford University Press, 2000), 94.

<sup>39</sup> At this time, I do not consider *values of conscience* which could fit within moral values. In Chapter 3 I discuss values of conscience that include commitments to standards of religious purity and to projects that further the well being of all humans. Like moral values, values of conscience are often justified as important by the moral community which may include one’s deity. In this Chapter I limit my focus to the moral and personal values distinction in order to solve the intuition and moral integrity problems.

<sup>40</sup> The notion of *community* that I am using throughout this paper is one proposed by Saul Kripke in his analysis of Ludwig Wittgenstein’s discussion of rule following in his *Philosophical Investigations*. For the purposes of this paper I am operating under the notion that a community is a group of individuals with shared concepts and rules. These concepts and rules, though always evolving, provide a standard for community members to assert that an individual is or is not following a rule, such as meeting the characteristics required for the attribution of integrity. To follow a rule or fall under a concept makes reference to the individual as a member of the community, and the “community must be able to judge whether an individual is indeed following a given rule in a particular application, i.e. whether his responses agree with their own” (635). These judgments take into account the context of the action. See Saul Kripke, “On Rules and Private Language.” *The Philosophy of Language – Fifth Edition*. Ed. A.P. Martinich. (New York: Oxford University Press. 2008), 626-638. Ludwig Wittgenstein, *Philosophical Investigations*. Trans. G.E.M. Anscombe. (Massachusetts: Blackwell. 2001).

<sup>41</sup> I loosely follow Bernard Williams’ description of things that are *simply important* and those that are *important relative to a person*. Not killing other community members is something that is simply important while finding a stamp to finish a stamp collection is important relative to a person. Bernard Williams, *Ethics and the Limits of Philosophy*. (Massachusetts: Harvard University Press, 1985), 182-187.

equate to an ID-commitment to moral values, are essential for communities to function.<sup>42</sup> The character traits that fit within conscientiousness are “honesty, fairness, truthfulness, and being a person of one’s word.”<sup>43</sup> Wallace correctly states that these traits and values “make possible activities and institutions that are necessary for communities and [are] highly beneficial for the members of communities.”<sup>44</sup> The benefits of trust, efficiency, and survival are instrumental reasons that justify the importance of moral values and also justify community members expecting each other to actualize these values.

Moral values are also justified by moral reasons which are those demands for certain behaviors that community members can make upon each other. Consider two neighbors Jane and Sam who talk to each other while doing yard work. Most people would agree that they are justified in demanding honesty from each other because they are members of the same moral community. If Sam is justified in demanding honesty from Jane, then most people would say that he is also justified in feeling upset if she lies to him.<sup>45</sup> The demand for honesty as well as the resentment from being lied to are both justified by the fact that moral community members are expected to actualize positive and negative moral values in their relations with other members. It follows that moral values are important and not easily dismissed by members. Members of the moral

---

<sup>42</sup> I take character traits to be similar to ID-commitments to values because both character traits and ID-commitments constitute a stable identity.

<sup>43</sup> James Wallace, *Virtues and Vices*. (Ithaca: Cornell University Press, 1978), 90.

<sup>44</sup> *Ibid.*, 1978, 15.

<sup>45</sup> Strawson describes specific attitudes that we are justified in having toward other members of the moral community who violate community values. These “reactive attitudes” include indignation, resentment, gratitude, guilt, and blame. He also notes that children and mentally challenged adults are not appropriate subjects of the reactive attitudes. Strawson, “Freedom and Resentment.”

community are justified in demanding that community members fulfill their obligation to actualize these values for instrumental and moral reasons and are justified in holding them accountable if they fail to do so.

Because these instrumental and moral reasons justify moral values as being important for all community members, members cannot easily dismiss moral values as not applying to them.<sup>46</sup> The justification for moral values creates a *reasonable obligation* for all moral community members to actualize moral values for the sake of its functioning, survival, and the mutual expectations of its members.<sup>47</sup> For example, if I am walking down the street, the person coming toward me has a reasonable obligation not to hit me in the face or to lie to me about my house being on fire. *Reasonable* describes those actions that require minimal costs to me and my interests, which is often how it is used in philosophical discussions regarding our obligations to help others. In other words, the reasonable obligation is defeasible because it would not require the person coming toward me to give his life rather than hit me in the face or to sell everything he owns in order to keep a promise. Moral values only exert a reasonable and defeasible obligation that one can ignore in some circumstances.<sup>48</sup>

---

<sup>46</sup> Wallace argues that a community can reasonably “require of one another adherence to certain generally beneficial forms of behavior. . . . General conformity in our community to certain modes of behavior is enormously beneficial, and these benefits are very widely distributed.” (Wallace, 1978, 115). Both Wallace (1978, 144-158) and Williams (1985, 174-196) provide detailed discussions about the origin of our obligations, in particular the origin of our obligation to provide mutual aid.

<sup>47</sup> I follow Wallace on this point (Wallace, 1978, 115). By *obligation* I mean that agent-neutral reasons exist and they justify community members expecting each other to actualize these values (e.g. tell the truth). Williams also explains that, “Obligation works to secure reliability, a state of affairs in which people can reasonably expect others to behave in some ways and not others” (Ibid., 1985, 187).

<sup>48</sup> It is worth noting that it is often those who consistently push past the “reasonable” limit regarding moral values who are shining examples of integrity. A person who acts in accord with the moral value of honesty

How do these rough observations about personal and moral values relate to integrity and the three cases above? In section 1 I explained that the sentence “Mary is a person of integrity,” is interpreted by most people to mean that she is morally upright. According to the identity account of integrity, this means that at a minimum Mary has a coherent ID-commitment to *moral values*. I also suggested that she has substantive integrity which is in line with common usage for the term integrity. If this is correct, then Scar and Socrates both appear to have some type of substantive integrity. Scar has an ID-commitment to the moral value of fidelity and Socrates has an ID-commitment to the moral value of justice.

Don Giovanni, however, has an ID-commitment to act in accordance with his personal values. Unlike moral values, personal values are justified as important by the agent’s personal reasons. For example, the importance of Don Giovanni’s personal value of seducing women is not justified by reasons that refer to the function and expectations of the moral community nor to his reasonable moral obligations to community members. Instead, he justifies the importance of his value by his own reasons, or those reasons that are relative to him and have no power to obligate anyone else. Don Giovanni can change his personal reasons and thus remove his obligation to his personal values. He cannot, however, remove himself from the moral community’s reasonable obligation that he not deceive and harm others. Moral values apply to Don Giovanni whether he chooses to recognize them or not. Because he has ID-commitments to personal values, and immoral

---

even when it is reasonable for her to lie reveals an ID-commitment and determination which is a sign of integrity.



ones at that, it appears that he has formal integrity, which is a shadow of substantive integrity.

To summarize the first distinction among the three cases above, moral values are justified as important for instrumental and moral reasons that support the function and expectations of the moral community. Scar and Socrates seem to meet this condition and therefore appear to have substantive integrity defined as having ID-commitments to moral values. Don Giovanni, however, has formal integrity because he lacks ID-commitments to moral values and actually violates common moral values in his pursuit of personal values.

This first distinction moves us closer to solving the intuition problem and the moral integrity problem, but both problems loom over the discussion because intuitively Scar may have substantive integrity even though he rarely keeps his word to those who are not in the Mafia. The next task is to determine what distinguishes Scar from Socrates. A significant difference between Cases 2 and 3 is not found in the fact that they hold different moral values, but the *scope* the agents apply when actualizing their moral values. A person can have an ID-commitment to moral values like honesty and the result is her community members trust her, but she may at the same time restrict the salient community for these values.

Let's return to Scar who always keeps his word to those in the Mafia. The community that is salient to ascribing integrity to Scar is the Mafia which requires a certain amount of trust and moral behavior for the community to function and survive. Scar's ID-commitment to keeping his word reveals that his community can trust him and

can function without worrying about its survival. Not only can Mafia members trust that Scar will do what he says, the police may eventually “trust” that Scar will do what he says. As pointed out by Graham, trust in this case does not mean that Scar is *trustworthy* in the full sense of that word because it does not extend to the larger human community. Given Scar’s ID-commitment to keeping his word, it appears that he has integrity in the Mafia community but not in the larger moral community.

Socrates demonstrates his integrity through his ID-commitment to the moral value of justice. Unlike Scar, Socrates’ ID-commitment to justice connects him to humankind in general, or the *human community*. Socrates’ actions demonstrate his ID-commitment to not do an injustice to anyone. The concept of belonging to and acknowledging the human community is not a new notion.<sup>49</sup> The Stoics specifically distinguished the roles one plays in particular groups and the role one plays as a citizen of the world.<sup>50</sup> These distinctions are quite natural because they capture the insight that we may be expected to avoid violating moral values that increase the trust and effective functioning of a particular community (e.g. university, nation) as well as the human community.<sup>51</sup>

The central difference between Scar and Socrates is that Scar is only partly committed to a moral value while Socrates is wholeheartedly committed to a moral value.

---

<sup>49</sup> T. H. Irwin has argued that, according to Aristotle, virtuous acts reference the overall community’s good. T. H. Irwin, “Aristotle’s Conception of Morality.” *Proceedings of the Boston Area Colloquium in Ancient Philosophy*, Vol. 1. 1985.

<sup>50</sup> Julia Annas, “My Station and Its Duties: Ideals and the Social Embeddedness of Virtue”. *Proceedings of the Aristotelian Society* (January 2002).

<sup>51</sup> The distinction between communities adds an additional layer to Graham’s definition of morality as equal respect for others. On my account of moral values, one can have equal respect for others *in your community* (e.g. Tuskegee Physicians) and also equal respect for others in the *human community*. The limited scope of the former explains why a person may genuinely believe that she has integrity while not realizing she has limited the scope of her moral concern.

To clarify this point, consider integrity attributions and their relationship to moral values. When a person of substantive integrity has an ID-commitment to moral values, it means that she will not *violate* these values because her identity is constituted by this ID-commitment. As a reminder, the original meaning of integrity combines the Sanskrit roots *tag* meaning “to touch or handle” and *in* meaning “not.” The result is the word *integ* which can be interpreted as not touched, pure, healthy, and uncorrupt. The definition does not imply that integrity requires a person to be altruistic and to help people. It *does* mean that she will not violate moral values lest she be unfaithful to her deepest moral commitments and thereby corrupt herself. In the cases above, Scar’s values are corrupted because he limits their scope and Socrates’ are uncorrupted.

Another difference between Scar and Socrates is highlighted by what it means to have an *ID-commitment* to a moral value. Having an ID-commitment to any value strongly implies a level of consistency in one’s behavior. I argued in Chapter 1 that integrity requires coherence between values and actions, and this also implies consistency. In the context of ID-commitments to moral values, consistency means that the person’s moral behavior does not unreasonably vary across situations and people. Scar’s supposed ID-commitment to keep his word lacks consistency as demonstrated by his behavior outside of the Mafia. Even if he is consistent within the Mafia, he is living two lives from two ID-commitments. One commitment is to keep his word to his fellow criminals, but the highest ID-commitment in his life is to ensure the success and survival of a crime syndicate — hardly the ID-commitment of a person of substantive integrity. Seneca describes a similarly duplicitous person who dutifully cares for a sick friend, but

who is also thinking about receiving part of the inheritance. He states: “As birds of prey that feed upon carcasses keep watch near by the flocks that are spent with disease and are ready to drop, so such a man gloats over a death-bed and hovers about the corpse.”<sup>52</sup> I suggest that Scar, like this duplicitous person, keeps his word for the sake of the benefits he receives from the Mafia and its activities and not because he has a coherent ID-commitment to a moral value.

Unlike Scar, Socrates demonstrates consistency by refusing to bring Leon of Salamis to trial. His action, if we assume that it is typical of him, reveals a consistent application of his ID-commitment to justice, regardless of the situational factors or the persons involved. I again emphasize that it seems that his integrity is not revealed by any act of altruism, but by being true to his ID-commitment to justice.

In summary, Scar is not a person of substantive integrity because his ID-commitment to a moral value is corrupt and inconsistent. His ID-commitment to keeping his word only applies to a small group of people and he reserves the right to violate this moral value when it benefits him and his organization. Socrates, on the other hand, will not do an injustice to anyone. His action reveals that he is a trustworthy person who will not stray from his commitment to justice. I conclude that given common usage, Socrates has substantive integrity because the human community can trust that he will not violate his moral values which could result in harming the human community.<sup>53</sup> Scar has formal

---

<sup>52</sup> “On Benefits,” in Lucius Anneaus Seneca, *Seneca Moral Essays*, Trans. John W. Basore. (Cambridge: Harvard University Press, Loeb Classical Library, 1989) Volume III, book IV, xx.3.

<sup>53</sup> Eugene V. Torisky notes the increase of trust in a community when supererogatory acts combine with integrity. He states, “My contention is that an ethics of supererogation, played out over an entire society by agents acting with integrity, functions as a sort of social ‘forgiveness rule.’ Supererogatory actions, like

integrity because he cannot be trusted to be true to his moral values in the human community. In fact, we know that he will keep his word when doing so advances an organization that thrives on intimidating members of the human community.

At this point I can revise the identity account and include the insights discussed above to solve the intuition problem and the moral integrity problem. I call this account the *moral identity account of integrity* to indicate that moral values are necessary to have substantive integrity. First, a person of substantive integrity has coherent ID-commitments to moral values. Second, she has the determination to act on them consistently. This means that she is true to her moral values across situations and people. The person of integrity is not necessarily a moral altruist, but is necessarily a person who can be counted on to maintain her moral values even under great pressure to compromise.

By positing these two necessary conditions, I am *not* saying the following. I am not saying that a person cannot have ID-commitments to personal values as well as moral values. However, a lack of substantive integrity is evident when one acts on personal values that violate moral values. The Tuskegee physicians fit this description because they acted on their personal and professional values to the exclusion of their moral values. I am also not claiming that a person of substantive integrity will not violate a moral value to further a qualitatively better moral value, such as lying to save someone's life. I address this concern in Chapter 3. Finally, I am not saying that these two

---

actions in accordance with duty, help to build trust, the ability to sustain the social good without continual or face-to-face enforcement." "Integrity and Supererogation in Ethical Communities," Paper given at the Twentieth World Congress of Philosophy, Boston, Massachusetts. August 10-15, 1998. Accessed September 9, 2010. <http://www.bu.edu/wcp/Papers/Soci/SociTori.htm>.

conditions are sufficient to be a person of substantive integrity. I still need to address several concerns and questions about the identity account.

I propose that the moral identity account as presented solves the intuition problem, the moral integrity problem, and answers the two questions not covered by the moral trustworthiness account. The intuition problem as stated above is that our intuitions are pulled in different directions when it comes to attributing integrity to tyrants (e.g. Don Giovanni) and heroes (e.g. Socrates). The moral identity account explains that tyrants only appear to meet the two necessary conditions for substantive integrity. Their lack of ID-commitments to moral values and/or their inconsistent behavior regarding moral values does not match the common notion of a person of integrity. Meanwhile, heroes such as Socrates meet the necessary conditions for having substantive integrity. The intuition problem is solved when we recognize that tyrants have formal integrity because they have coherent ID-commitments to some values and the determination to act on them. The difference is that heroes have coherent ID-commitments to *moral values* and the determination to act on them consistently.

The identity account also solves the moral integrity problem. To say that “Socrates is a person of integrity” is just to say he has an ID-commitment to moral values and the determination not to violate them. This does not imply that people like Don Giovanni and Scar cannot display characteristics that resemble the substantive integrity held by Socrates. I solve the moral integrity problem by highlighting the differences between Socrates’ consistent maintenance of his moral values and Don Giovanni’s and Scar’s inconsistent maintenance of moral values. If this distinction holds, substantive

integrity cannot refer to tyrants. Instead, tyrants have a formal integrity that is similar to substantive integrity in form but not in substance.

The moral identity account also answers the questions, “What is the content of moral trustworthiness?” and “What features do the tyrant and the hero share and what features make them fundamentally different?” The content of moral trustworthiness is coherent ID-commitments to act in accordance with moral values and the determination to act on them consistently. The moral identity account includes moral values while introducing the coherent ID-commitments and determination that provide the stable character needed to be trustworthy. The second question is answered when one considers the necessary conditions presented above. Tyrants and heroes both have coherent ID-commitments. Tyrants, however, either lack the ID-commitment to moral values and/or are inconsistent in their application of these moral values across the moral community.

### **Conclusion**

The purpose of this chapter is to develop an integrity account that answers the intuition problem and the moral integrity problem. In section 1 I presented identity theorist replies to the moral integrity problem and argued that these leave a gap between the identity account and our common usage of integrity as a moral concept. I also argued that the personal/moral integrity distinction is contrary to common usage. I then suggested that distinguishing between substantive integrity and formal integrity can explain our diverging intuitions regarding tyrants and heroes while also moving the identity account closer to common usage.

In section 2 I reviewed two non-identity accounts of integrity and showed how they did not adequately solve the moral integrity problem. Calhoun's account allows rational tyrants to have integrity and Graham's account left open two major questions underlying the intuition problem and the moral integrity problem. In section 3 I presented the moral identity account of integrity. I distinguished between moral values that are justified by instrumental and moral reasons and personal values that are justified by personal reasons. I concluded that to have substantive integrity one must have coherent ID-commitments to moral values and the determination to act on them consistently.

Several concerns and questions have been raised against identity accounts in general and these objections apply to the moral identity account as well. One concern is that integrity may not be a moral virtue because it does not appear to be a dispositional virtue like compassion nor does it enable a person to be virtuous in the way that self-discipline or courage may enable generosity. A second concern is that we attribute integrity to some people who will not violate their consciences and yet their consciences do not fall into the realm of moral values. Another concern is that the identity account would not allow a person to change her commitments when presented with new evidence. Related to this concern is the objection that a person of integrity would not be able to violate the moral value of honesty to save someone's life. A final concern is that the identity account does not explain how two people can have different moral values and still be people of integrity. I address these questions and concerns in Chapter 3 and



further construct an account of how a person of integrity coherently structures her ID-commitments, desires, and ends.

### Chapter 3 - Integrity and Virtue: A Case of Being True to Yourself

In the preceding two chapters I outlined the moral identity account of integrity by revising a generic identity account of integrity. I argued that in common language to say that a person has integrity is to indicate that she is morally trustworthy. I also explained the differences between substantive integrity and formal integrity. To have substantive integrity a person must have coherent identity-conferring commitments (i.e. ID-commitments) to *moral values* and the determination to consistently act on them. Formal integrity is a *form* or shadow of substantive integrity in that a person has ID-commitments to some values but does not consistently adhere to moral values across situations and groups. Unlike the person of substantive integrity, the person of formal integrity could not be trusted to live by basic moral values such as honesty, keeping one's word, or acting justly. A person of substantive integrity is not necessarily an altruistic person, but she does not violate moral values, particularly those that increase the trust of other community members.

As discussed in Chapter 1, most identity accounts of integrity are based on Bernard Williams' claim that a person's integrity is constituted by her ID-commitments to values, principles, and life projects. These commitments are ones with which a person is "deeply and extensively involved and identified."<sup>1</sup> Williams states, "There is no contradiction in the idea of a man's dying for a ground project – quite the reverse, since if death really is necessary for the project, then to live would be to live with it unsatisfied,

---

<sup>1</sup> Bernard Williams and J.J.C. Smart. *Utilitarianism: For and Against*. (New York: Cambridge University Press, 1973), 116.

something which, if it really is his ground project, he has no reason to do.”<sup>2</sup> For Williams, a person’s integrity is not based on just any commitments, but commitments to values or projects that constitute her identity.

I have also relied on Williams’ claims in developing the moral identity account. Williams, however, takes his account of integrity in a completely different direction. He argues that because integrity is being true to one’s deepest life projects and commitments, it is not a moral concept. His famous example of integrity is an artist who leaves his family to fulfill his ID-commitment to painting masterpieces. The painter disregards his responsibilities to his wife and children to pursue his ground project and therefore he has integrity. By defining integrity as being true to one’s ground projects, Williams excludes integrity from the realm of morality and he also argues that integrity is not a virtue.

In this chapter, I addressed the question of whether or not integrity is a virtue by presenting different perspectives in this debate as well as my arguments for why integrity is a virtue. I also present an in-depth description of how a person develops a self-concept in order to respond to several questions about what it means to have substantive integrity. As in earlier chapters, I rely on common intuitions to draw out the boundaries of integrity as it is a concept that we understand but often struggle to articulate.

In section 1 I present Williams’ and John Rawls’ claims that integrity is not a virtue or at least only a secondary one. I argue that both of them offer important but incomplete accounts of integrity and present an account that best matches common usage of the concept and our intuitions. In section 2 I present an extended account of what

---

<sup>2</sup> *Moral Luck*. (New York: Cambridge University Press, 1981),13.

constitutes a self-concept to combine my virtue argument with the moral identity account from Chapter 2. In section 3 I answer four questions that my revised account of integrity can now answer: 1) Do “honest thieves” have substantive integrity? 2) Can a person with substantive integrity ever lie in order to actualize a “higher” moral value? 3) Can a person with substantive integrity alter or change her identity and still have integrity? 4) Can two people have opposed moral beliefs and both have substantive integrity?

### **Section 1: Is Integrity a Virtue?**

In this section I answer the question: “Is integrity a virtue?” For the sake of clarity, I am going to use the word “integrity” in a broad sense that does not necessarily require moral content. My purpose in re-expanding the usage of integrity is to consider accounts that appear to track a non-moral kind of integrity. I first consider two virtue accounts of integrity. Bernard Williams denies that integrity is a virtue and John Rawls argues that it is only a “virtue of form” that can have a variety of content. I then argue that integrity is a virtue directly related to being true to one’s identity.

Williams is well known for his arguments against the requirements of impartiality built into utilitarianism and Kantianism. His main objection to these theories is that they seem to require that we alienate ourselves from the projects, people, and ID-commitments that provide meaning to our lives. According to Williams, a person’s integrity is constituted by the ID-commitments and ground projects that give her the motivation and reason to live.<sup>3</sup> This definition fits with common intuitions that integrity refers to a non-

---

<sup>3</sup> Bernard Williams, *Ethics and the Limits of Philosophy*, (Massachusetts: Harvard University Press, 1985), 13.

moral “whole self,” but it conflicts with the common notion that integrity is a moral concept.<sup>4</sup>

To defend his definition of integrity, Williams provides two arguments for why integrity is not a virtue. First, integrity is not a virtue because it is not a *disposition* that yields motivations like the motivational virtues of generosity and benevolence. Williams says that dispositional virtues are excellences of character, “which are internalized dispositions of action, desire, and feeling.”<sup>5</sup> His second argument is that integrity is not like the virtues of courage and self-discipline which “enable one to act from desirable motives in desirable ways.”<sup>6</sup> Integrity is merely the *expression* of the dispositions that form a person’s identity as well as any virtues that enable one to act on these dispositions. In summary, “Integrity does not enable him to do it [virtuous action], nor is it what he acts from when he does.”<sup>7</sup>

Williams grounds his first argument on the common definition that virtues are internalized dispositions of action. Robert Merrihew Adams also categorizes some virtues as motives and dispositions that often aim at actualizing some good for others, such as benevolence and compassion. He calls these *motivational virtues* because they engage the will of a person to act for some human good.<sup>8</sup> For example, a person with the virtue of compassion is motivated to help someone in need and will act on that

---

<sup>4</sup> See an excellent critique of Williams’ use of the word *integrity* in Damian Cox, Marguerite La Caze, and Michael P. Levine. *Integrity and the Fragile Self*. (Vermont: Ashgate, 2003).

<sup>5</sup> Williams, 1985, 35.

<sup>6</sup> Bernard Williams, *Moral Luck*, (New York: Cambridge University Press, 1981), 49.

<sup>7</sup> *Ibid.*, 49.

<sup>8</sup> Robert Merrihew Adams, *A Theory of Virtue*. (New York: Oxford University Press, 2006), 33.

disposition. Another characteristic of motivational virtues is that they are stable and consistent across situations. For example, we would not attribute the virtue of compassion to someone who only acts compassionately twice a year or restricts her compassion to her own children. Williams argues that integrity is not a virtue because it is not a disposition to do anything and even if integrity implies consistency, the motivational virtues provide stability and consistency on their own.

Williams' second argument is that integrity is not a virtue that engages the will to *enable* one to act on dispositional virtues. Integrity is not like the enabling virtues of courage and self-discipline. Consider Jeff who feels compassion towards a beggar but knows that he will be made fun of by his co-workers if he stops the car and gives him money. On Williams' account, the virtue of courage provides the additional will-power that Jeff lacks to act on his virtue of compassion. Similarly, if Sarah is tempted to give less generously to a charity, the virtue of self-discipline can enable her to act on her disposition and sacrifice for the sake of a good cause.

Adams also notes the unique qualities of these enabling virtues and categorizes courage, patience, and self-discipline as *structural virtues* because they are not defined by motives nor by one's aims. Instead, they are "structural features of the way one organizes and manages whatever motives one has."<sup>9</sup> A person can call on structural virtues such as courage to make up for a lack of motivation.

To successfully argue that integrity is a virtue using Williams' description and Adams' classification, I must demonstrate that integrity is either a motivational virtue

---

<sup>9</sup> Adams, 2006, 33.

and/or a structural virtue. Also, I must demonstrate that integrity is not merely the expression of another virtue in a person's life. To meet this challenge, I consider two cases of people who most would agree have some type of integrity, whether it is a virtue or not. I then determine what characteristic reveals their integrity.

Case 1: Socrates is told by a rogue government in Athens to bring Leon of Salamis to be executed. Because Socrates believes the charges against Leon are unjust, he disobeys the command to deliver Leon and goes home knowing his action could lead to his own death. Socrates later told his own jury that, "Then I showed again, not in words but in action, that, if it were not rather vulgar to say so, death is something I couldn't care less about, but that my whole concern is not to do anything unjust or impious."<sup>10</sup>

Case 2: Thomas More became Lord Chancellor to English king Henry VIII in 1529. A strong defender of the Pope's authority above that of the king's in certain matters, More refused to sign a letter asking the Pope to annul Henry's marriage to Catherine of Aragon. He eventually resigned his commission, but was still pressured to compromise his religious beliefs. In 1534, More refused to swear his allegiance to the *Parliamentary Act of Succession* because of an anti-papal preface in the oath and his refusal to uphold Henry's divorce. He was imprisoned in the Tower of London and eventually tried as a traitor. When offered the king's mercy if he would change his mind about the marriage, More stated, "I beseech Almighty God that I may continue in the mind I am in, through His grace, unto death." He was convicted of treason and sentenced to death. Before the executioners ax fell he told the crowd, "I die the king's good servant, and God's first."<sup>11</sup>

At first glance, both Socrates and More appear to be people who have integrity in the sense that they are determined to actualize ID-commitments to particular values. But do they have a specific virtue of integrity in addition to the virtues of justice and piety? What is distinctive about these cases is that both individuals confront a direct threat not

---

<sup>10</sup> Plato, *The Apology in The Trial and Death of Socrates*. Trans. G. M. A. Grube. (Indianapolis: Hackett Publishing Company, 2000) 32a-e.

<sup>11</sup> Douglas Linder. 2006. "Trial of Thomas More" from <http://www.law.umkc.edu/faculty/projects/ftrials/more/moreaccount.html>.

only to their virtues and life projects, but to their identities.<sup>12</sup> Socrates will not compromise his identity as a just and pious person and More will not compromise his identity as God's servant. It appears that these men share a disposition to maintain their identities, and I claim that this disposition is the virtue of integrity. I define the virtue of integrity as the disposition to live out one's identity or as John Kekes calls it, "*being true to oneself*."<sup>13</sup>

Contrary to Williams' argument that integrity isn't about anything, integrity appears to be a disposition to act in accordance with one's identity. Socrates and More both display a disposition to maintain and protect their identities even in the face of death. Why? To most people, being true to oneself is extremely important because one's identity is arguably a good in itself in which we all invest heavily. Most people want to be a certain kind of person and cultivate the beliefs, emotions, desires, and ID-commitments that constitute that person. Kekes recognizes this motivation when he describes integrity as a form of constancy over time: "To have constancy is to be steadfast in adhering to one's deliberate pattern as it is transformed from a distant ideal to one's second nature and true self."<sup>14</sup> Even if someone adopts an identity for instrumental reasons, the act of adopting an identity is itself an important expression of who one is and wants to be.

---

<sup>12</sup> I define a person's identity as who she is as well as her clear vision of who she wants to be. Another word for identity could be self-conception, but I use *identity* to maintain consistency with Williams and the identity account. The content of a person's identity includes her beliefs, emotions, desires and ID-commitments. I discuss the content of a person's identity in more detail in section 2 below.

<sup>13</sup> John Kekes, *Enjoyment*, (New York: Oxford Press, 2008), 145.

<sup>14</sup> John Kekes, "Constancy and Purity" *Mind Vol. 92, No. 368*. (Oct. 1983), 499.



John Finnis, reflecting on the good of being oneself, writes that, “all one’s free choices go to constitute oneself, so that one’s own character or identity is the most fundamental of one’s ‘accomplishments’; the accomplishment most unequivocally one’s own . . .”<sup>15</sup> Finnis bases his observations on Aristotle’s claim that an excellent person desires to be himself and no one else. “For being is a good for the good person, and each person wishes for goods for himself. And no one chooses to become another person even if that other will have every good when he has come into being . . .”<sup>16</sup> Just as compassion is a disposition to help others in need, integrity is a disposition to be true to oneself. I discuss the different contents of a “self” in section 2.

Williams could object at this point and restate his argument that both cases show actual virtues in action and bringing in integrity unnecessarily multiplies virtues. Socrates is merely acting on the virtue of justice and More is merely acting on the virtue of piety. To attribute integrity to them only recognizes an expression of these virtues.

At first glance it appears that Williams is correct about multiplying virtues because it seems that ID-commitments to justice and piety may motivate action by themselves. But upon reflection, we can see that integrity can be a distinct disposition from other virtues. Consider the case of a fictional person Locrates. His country is invaded by a foreign army and he is told by an invading General to eat swine’s flesh or face execution. Over the years Locrates has developed his identity to be the kind of person who finds eating meat detestable and immoral. He boldly tells the rulers, “I will

---

<sup>15</sup> John Finnis, *Fundamentals of Ethics*, (Washington D.C.: Georgetown University Press, 1983), 40.

<sup>16</sup> Aristotle, *Nicomachean Ethics*. Trans. Terence Irwin. (Indianapolis: Hackett Publishing Company, 1999). (1166a 20).

not eat meat.” In this case where no other disposition motivates Locrates to avoid eating meat, he still has the disposition to be true to himself, i.e. the virtue of integrity.

I acknowledge that distinguishing traditional virtues from integrity can be difficult because each person has only one identity constituted by multiple ID-commitments and dispositions. Most discussions of virtue define each virtue as a distinct disposition and analyze it apart from a person’s identity and her other dispositions. Psychologist Augusto Blasi laments the “atomization and depersonalization” that happens in studies regarding moral understanding and virtues. He states, “It is as if the various virtues—whatever they are—are self-subsistent entities, and not qualities of a person; as if their being virtues did not depend on the whole personality in which they are imbedded.”<sup>17</sup> A virtue-centric approach as opposed to an identity-centric approach would naturally focus on a person’s just or pious dispositions without considering her disposition to be true to herself. An identity-centric approach, on the other hand, describes Socrates’ virtue as a disposition to justice that is one part of his identity. His disposition to never do anything unjust and his disposition to be true to himself cannot be easily separated. I consider the interaction of dispositions in more detail in section 3.

The Locrates case is intended to show that a person can have the virtue of integrity without attributing her actions to another disposition. On the other hand, can a person have a virtue *without* having the virtue of integrity? I think this is possible.

Consider another fictional person Thomas Poor. He has the virtue of piety as

---

<sup>17</sup> Augusto Blasi, “Moral Character: A Psychological Approach”. In *Character Psychology and Character Education*. Ed. Daniel K. Lapsley and F. Clark Power. (Indiana: University of Notre Dame Press, 2005): 96.

demonstrated by attending church, following each tenet of his faith, saying his daily prayers, and giving alms. However, when the authorities tell Poor that he must reject a core tenant of his faith or be an outcast from his city, he reluctantly agrees. He feels great remorse for compromising his identity, but his disposition to piety is not lessened. Poor is a person who has the virtue of piety but lacks the disposition to be true to himself under all circumstances.

In response to Williams' objection above, the Locrates and Poor cases imply that the virtue of integrity can be distinct from other virtues and that attributing it to a person does not unnecessarily multiply virtues. We can attribute the virtue of integrity to Socrates and More because they seem to have the virtue of being true to themselves in addition to their other dispositions. If my account is correct, then integrity is a motivational virtue and Williams has incorrectly rejected it as a virtue.<sup>18</sup>

It may appear from the Poor example that integrity is also a structural virtue. If Poor had piety *and* integrity, maybe he would not have compromised his faith. If integrity is like courage, it would have enabled him to fully act on his piety. As noted above, Williams' second reason for rejecting integrity as a virtue is that it does not enable a person to carry out her dispositions.

I agree with Williams on this point. Integrity appears to be a disposition to be oneself and the strength of that disposition may wane during trying circumstances. In line with the structural virtue category, Philippa Foot observes that virtues like courage

---

<sup>18</sup> In a related discussion, Williams has argued that to say "not through me" as a reason for not doing evil is not a motivating thought. It represents another motivation, possibly the fear of pollution or pride (e.g. "I won't do what you say!"). Neither of these examples, however, respond to the strong disposition to be true to oneself found in the examples above (1985, 50).

are “correctives” in that they make up for a lack of motivation in human nature which typically seeks to avoid pain.<sup>19</sup> In the cases of Socrates and More, they may have enlisted the virtue of courage to enable them to be true to themselves. I am *not* claiming that the disposition to be true to oneself cannot empower a weaker disposition into action as it appears to have done with Socrates and More. I am claiming that when fear or laziness challenge one’s ID-commitments, even the disposition of integrity may require courage and self-discipline to reach its end. Again, pulling these dispositions apart can be an inexact process.

Another objection to integrity as a motivational virtue could come from John Rawls. Rawls, in *A Theory of Justice*, states that integrity and associated virtues of “truthfulness and sincerity, lucidity, and commitment, or as some say authenticity,” are “virtues of form.”<sup>20</sup> According to Rawls, in times of social doubt when no one is certain what is actually moral, people run to these virtues of form to defend their own moral views. As such, integrity is a virtue that could justify a tyrant’s commitment to injustice and a hero’s commitment to justice. Rawls states, “It is impossible to construct a moral view from these virtues [e.g. integrity] alone; being virtues of form they are in a sense secondary.”<sup>21</sup> However, when the virtues integrity and authenticity are joined with an appropriate concept of justice, then “they come into their own.”

---

<sup>19</sup> Philippa Foot, “Virtues and Vices”. *Virtue Ethics*, Ed. Stephen Darwall (Massachusetts: Blackwell Publishing, 2007).

<sup>20</sup> John Rawls, *A Theory of Justice – Revised Edition* (Massachusetts: Harvard University Press, 1999), 455.

<sup>21</sup> *Ibid.*, 1999, 456.

Rawls offers a perspective on different kinds of virtues that is worth considering. He distinguishes between virtues that directly concern others (i.e. primary virtues) and those that can support any content (i.e. secondary virtues). Rawls does not make the motivational/structural distinction but instead focuses on the content and object of a virtue in determining the role it plays. In viewing integrity as “truthfulness and sincerity, lucidity, and commitment, or as some say authenticity,” Rawls is correctly observing that each of these secondary virtues can be held by a tyrant which means they are not “real” virtues that are necessarily concerned with others.<sup>22</sup> But lacking a concern for others does not exclude these secondary virtues from being dispositional. It is not controversial to say that people can have a disposition to be truthful, sincere, or authentic for good or evil ends. On this interpretation of Rawls, he would not deny that integrity is a motivational virtue, only that it can have a variety of content. Or, put another way, one can be true to oneself and this self may have ID-commitments to a variety of values.

To summarize my account of integrity as a virtue, I have argued that integrity is a motivational virtue because it is a disposition to be true to oneself. I have also argued that while it is not strictly a structural virtue like courage or self-discipline, it may call on these virtues to enable its fulfillment. I also agreed with Rawls that integrity may be a secondary virtue because its content may not require a direct concern for others. I have not addressed every objection to integrity being a virtue, but I do consider my explanation a strong support for the common belief that integrity is a virtue in some sense of the word.

---

<sup>22</sup> Rawls, 1999, 455.

In the next section I address two questions regarding integrity as “being true to oneself.” The first question is, “What constitutes this *self* to which one is true or committed?” and the second is, “Is integrity primarily a *moral* virtue?”

## **Section 2: Integrity and the Coherent Self**

In order to proceed with defining the virtue of integrity as “being true to oneself,” I must sketch out a general account of what constitutes a self. I have two purposes for sketching a general description of the self. First, to understand integrity we must understand what constitutes the self and how and why a coherent self is maintained. Because the self is the center around which integrity attributions refer, an understanding of integrity requires an account of the self. My second purpose is to prepare the conceptual framework needed to answer four questions about the moral identity account of integrity.

In this section I describe three accounts of the self and select one that is most widely accepted as a model for further discussion. I then describe three items that constitute the self: beliefs, desires, and dispositions. I offer an account of how the self manages these three items to develop a coherent identity to which one can be true. I complete this section by explaining how substantive integrity relates to the virtue of integrity. I explain that sometimes a person with ID-commitments to their conscience can also have integrity.

### **2.1 - What Is a Self?**

It is beyond the scope of this chapter to outline, much less settle, the debate about what a self is and its ontological status. Having some concept of the self and its inner

workings, however, is necessary to understand the foundations of integrity and its development. Below I present three views of the self and select one that matches common intuitions as well as empirical research.

The Cartesian view claims that the self is a real entity. This view is often connected with Descartes *Meditations* in which he states, “I think therefore I am.”<sup>23</sup> The self could be a soul, a spirit, a disembodied mind, or some other independent entity. A difficulty facing this view is that it requires a set of metaphysical and theological arguments that are beyond the scope of my dissertation and the most recent literature on identity. While I do not reject this view out of hand, I bracket the view for the sake of clarifying the self as it is presented in recent philosophy and moral psychology literature.

A second view of the self is that it is a mere fictional character that is identical with brain states. According to Leslie A. Howe, “Some, like Daniel Dennett, would argue that selves are merely fictional characters produced by the brain but with no further reality than that possessed by such fictional characters.”<sup>24</sup> No “me” considers questions or answers. It is only the brain processing information and “I” am a fictional third party narrating the process. David Velleman questions Dennett’s account. He asks why the self should be a (third person) fictional character rather than “the author of a veridical autobiography, who really is identical with the protagonist of his story?”<sup>25</sup> Another concern about Dennett’s theory is that in his attempt to take a scientifically objective

---

<sup>23</sup> Rene Descartes, *Meditations on First Philosophy*. Trans. Ronald Rubin. (California: Areté Press, 2001).

<sup>24</sup> Leslie A. Howe, “Self and Pretence: Playing with Identity”, *Journal of Social Philosophy* Vol 39, No.4. (2008): 566.

<sup>25</sup> David Velleman, “The Self as Narrator,” in *Self to Self: Selected Essays* (New York: Cambridge University Press, 2006). 207.

view of the self, he presents a “fictional and impoverished self that fails all the intuitive criteria of selfhood.”<sup>26</sup> Velleman and others appreciate Dennett’s view but find another view more rich and accurate to our experience of the self.

The third self view is the self as narrator. This account argues that the self is a narrator of a life and some would say also the narrative itself.<sup>27</sup> Because it seems that the self is both narrator and the narrative, Velleman describes the self as the *process* or activity of narrating one’s life. This does not mean that I can narrate just any life I want. My narrative includes my past experiences and reactions to the world that have created a self-conception. The narrative concept of self suggests that the self is the process of narrating a life in accordance with a coherent self-conception. A self-conception is hard to define but may include habits and dispositions by which a person interprets who she is and her place in the world given her narrative.

Based on an extensive review of empirical literature on self conceptions and identity, psychologists Jeremy A. Frimer and Lawrence J. Walker also describe the self, and the moral self in particular, as a unifying process. They state, “The moral self can have the quality of accountability with imperfect unification so long as there exists an agentic, unifying process whose job it is to cross-reference the contents of the self (beliefs, values, behavior), *attempt* to mend inconsistency and have some (incomplete)

---

<sup>26</sup> Joan McCarthy, *Dennett and Ricoeur on the Narrative Self*. (New York: Humanity Books, 2007), 98.

<sup>27</sup> Narrative views of the self have been presented by Alasdair MacIntyre, Paul Ricoeur, and David Velleman. Alasdair MacIntyre, *After Virtue*. (Indiana: University of Notre Dame, 1984) especially chapter 15. Paul Ricoeur, *One Self as Another*. Trans. Kathleen Blamey. (Chicago: University of Chicago Press, 1992). Velleman, 2006.



success.”<sup>28</sup> On their account, the self is an agent-like unifying process which fits with Velleman’s process account. Frimer, Walker, and Velleman also claim that the self attempts to mend inconsistencies within the self and among its actions. In other words, the self seeks coherence among the contents of the self-conception and its actions. In what follows, I adopt the self-as-process model because it is a plausible account supported by some empirical evidence and common experience while avoiding difficulties faced by the other two views.

Returning to the topic of integrity as being true to oneself, it is necessary to discuss the contents of a self-conception or identity. I have narrowed the contents of an identity to beliefs and desires.<sup>29</sup> On this account of the self, a person with the virtue of integrity is true to an identity constituted by coherent beliefs and desires, with desires potentially becoming ID-commitments.

## **2.2 - Beliefs as Guides**

Two specific beliefs are worth discussing in the context of developing and managing a self-conception. They are beliefs about one’s identity and beliefs about which principles and ends should guide one’s decisions.

The first beliefs about identity take the form of “who I was,” “who I am,” and “who I want to be.” These are beliefs that make up the narrative of a life and they play a significant role when choosing among desires to actualize. When a person considers who

---

<sup>28</sup> Jeremy A. Frimer and Lawrence J. Walker, “Towards a New Paradigm of Moral Personhood.” *Journal of Moral Education Vol. 37. No. 3.* (2008): 346.

<sup>29</sup> John Kekes identifies three items that are significant for a person of integrity: beliefs, desires, and emotions. I do not discuss emotions in detail as they are often items that a person of integrity manages and controls more than the primary motivations for action, unless they are central to one’s ID-commitments (Kekes, 2008, 149).

she was, she has certain beliefs about who she was as a child, as a student, as an adult, and any other periods in her life. Statements that reveal these beliefs may include, “I used to be afraid of the dark. I trained to be a nurse. I have always loved philosophy. I have never told a lie.” The importance of these beliefs is how they form a historical narrative upon which a person bases her current identity. Beliefs about “who I am” can also refer to the “who I was” beliefs. These beliefs often center on our roles and obligations: “I am a graduate student. I am a father. I am a Christian.” Within each of these roles are additional beliefs about what it means to fulfill these roles.

Based on my discussion of identity creation in Chapter 1, I suggest that one’s beliefs about “who I am” often match one’s ID-commitments and together constitute a person’s self-conception. Beliefs and ID-commitments play this role by setting boundaries for one’s thoughts and behaviors. Recall the statement by Socrates that, “death is something I couldn’t care less about, but that my whole concern is not to do anything unjust or impious.” He lets his jury know that he *is not* a person who does unjust or impious actions. Also recall More’s words, “I die the king’s good servant, and God’s first.” He tells the assembled crowd and executioner that his role as God’s servant takes priority over his role as a subject of Henry VIII. Both of these examples demonstrate the power of “who I am” beliefs and I later I discuss how they integrate with a person’s dispositions.

The “who I will be” beliefs refer to an emerging or ideal identity to which a person aspires. These may include beliefs such as, “I believe I can be a lawyer. I believe I can be a cancer survivor. I believe a new job will make me happy. I will never be a

good cook.” These beliefs can limit or broaden one’s identity because they are often about possible selves. If Stan believes that it is impossible to become a doctor, he will not take the MCAT test. If he believes he can become a doctor, he will take the MCAT and apply to medical school. “Who I will be” beliefs open or close possibilities and by doing so they also influence the commitments, principles, and relationships a person selects and rejects in order to actualize these beliefs. Anyone who has made a decision to do something in line with her identity understands the role these identity-specific beliefs play in guiding our decisions and commitments.

In addition to having concrete beliefs about his identity, Socrates had beliefs about the principles and ends for which he would be willing to die. The self seems to use beliefs about principles and ends to guide decisions about which desires to actualize and which actions to take. Having a belief in a principle means believing that the principle represents a truth about how the world works or how we want it to work. The self also uses beliefs about what ends or goods are worth pursuing. Three ultimate ends a person may desire are advantage, pleasure, or some objective good. By objective good I mean something that is good in itself and usually good for others. These may include moral values, truth, freedom, and musical harmony. In this rough account of beliefs *about* ends, it is sufficient to recognize that different people value different ends. One person may believe that gaining power and wealth (i.e. advantage and pleasure) are the most worthy ends while another may believe that relieving suffering (i.e. objective good) is the most worthy end. A belief about what ends are worth pursuing may spark a desire for that end or lead us to reject an end.

Philosopher Anthony Flew exemplifies how beliefs about principles and ends can result in an ID-commitment and a disposition to be true to oneself. In the late 1940's, Flew graduated from Oxford and stayed on to do graduate work under the philosopher Gilbert Ryle. In observing how Ryle responded to other philosophers, Flew observed that he obeyed Socrates' command in the *Republic*: "We must follow the argument wherever it leads." He states, "It is a principle I myself have tried to follow throughout a long and very widely controversial life."<sup>30</sup>

Flew, an atheist before entering the academy, refined his arguments supporting his lack of belief and argued against theist positions for decades as a professor and writer. All the while he maintained his commitment to "follow the argument wherever it leads" while defending atheism in debates from the 1960s to the 2000's.<sup>31</sup> At a symposium in 2004, however, Flew announced that he now accepted the existence of God. Flew based his change on the evidence that the complexity of DNA arrangements required to produce life could not have happened by chance. Reflecting on his statement to the symposium audience, Flew notes that it "represented a major change of course for me, but it was nevertheless consistent with the principle I have embraced since the beginning of my philosophical life—of following the argument no matter where it leads."<sup>32</sup>

In addition to Flew, Socrates and More appear to have held beliefs about which principles and ends should guide their lives. The end that Socrates aimed at was not

---

<sup>30</sup> Anthony Flew, *There Is a God: How the world's most notorious atheist changed his mind*. (New York: Harper Collins, 2007). 22.

<sup>31</sup> Flew, 2007, 68.

<sup>32</sup> Flew, 2007, 75.

pleasure or advantage but rather to act in accordance with something that is objectively good, i.e. to live a just life. More's end was not to avoid pain and imprisonment, but to obey the tenets of his faith as he understood them and to give up his present life for one that he believed could never be taken away. I believe Flew's end was to find the truth in whatever matter he investigated and change his life accordingly. I consider the seeking of truth for its own sake an objective good because it is done for the sake of itself, though it may bring some pleasure as well. While all of these ends are the *objects* of desires and dispositions, the beliefs *about* these ends guide the selection of the ends and how they are actualized. Beliefs about principles and ends help construct a more durable identity and narrative by clarifying what is and is not consistent with that identity.

While important, beliefs do not fully constitute an identity. The process of developing and maintaining an identity involves a dynamic interaction of beliefs and desires. While beliefs provide a structure and a coherent narrative about what is possible and impossible, acceptable and unacceptable, and important and irrelevant to a self, desires and dispositions identify specific objects to be actualized or avoided and the motivation to actualize or avoid those objects.

### **2.3 - Desires, Dispositions, and Higher-Order ID-Commitments**

A person may have beliefs *about* the possibility and worthiness of ends, but she probably has desires and dispositions for specific ends as well. In this sub-section I review how some desires become ID-commitments which constitute a person's identity. I then discuss how ultimate goods and higher-order ID-commitments prioritize and resolve conflicts between desires.

In Chapter 1 I presented Harry Frankfurt's description of how a person commits to particular desires (and their ends) and how the resulting ID-commitments constitute her identity. I take Frankfurt's account as a plausible explanation of how we construct a self because his observations mesh well with a common understanding of how we make commitments as well as the findings of recent research on moral motivation and identity.<sup>33</sup>

As a reminder, Frankfurt explains the process of self-formation as selecting desires and deciding to make them one's will or volition. The process involves first-order desires, second-order desires, and second-order volitions. The object of a first-order desire is to perform or not to perform some action, like smoking or taking unfair advantage of people. The desire that ultimately results in action is thought to be what the person wanted to do, such as smoking a cigarette or cheating people. A person also has second-order desires that have as their objects the first-order desires.<sup>34</sup> To continue the unfair advantage example, a person may have a second-order desire *to have* the desire to treat people fairly which conflicts with a first-order desire to take advantage of them. She can also have a second-order volition *to want* treating people fairly to be her will, and this conflicts with her first-order desire to take advantage of them.

According to Frankfurt, the process of forming a second-order volition starts when a person "cuts off" a certain sequence of desires and makes a decision to *identify*

---

<sup>33</sup> For a recent discussion of moral identity formation and empirical research surrounding it, see Blasi, 2005.

<sup>34</sup> Harry Frankfurt, "Identification and Wholeheartedness." *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. Ed. Ferdinand Schoeman. (New York: Cambridge University Press, 1987), 32.

*with* or commit to a desire (e.g. to be honest).<sup>35</sup> The result is a commitment that partly constitutes her identity (e.g. “I am an honest person”). Either through deliberation or a non-conscious process (e.g. a parent’s example or a desire that has become a habit), the decision to commit to a desire forms an *intention* to actualize the desire.<sup>36</sup> Of course forming an intention does not imply that other desires and intentions will not interfere with her actualizing this particular desire, but deciding to have the intention means that she is not equivocating about which desire she wants to will. In Chapter 1 I equated intentions and second-order volitions with ID-commitments to values.

On my model of items that constitute a self, ID-commitments to particular values represent the specific desires that a person intends to actualize. The objects of these desires are the objects of ID-commitments. In Chapter 2 I described two particular values to which a person can have ID-commitments. I argued that in the context of integrity attributions, people tend to have ID-commitments to personal values (e.g. stamp collecting; seducing others) that are justified as important by personal reasons. They can also have ID-commitments to moral values (e.g. honesty, keeping one’s word, justice) that are justified as important by instrumental and moral reasons. I limited myself to these two values to highlight the difference between substantive integrity and formal integrity in common usage. When considering what constitutes a self-concept, another

---

<sup>35</sup> Frankfurt states, “The etymological meaning of the verb ‘to decide’ is ‘to cut off’. This is apt, since it is characteristically by a decision (though, of course, not necessarily or even most frequently in that way) that a sequence of desires or preferences of increasingly higher orders is terminated.” Frankfurt, 1987, 38.

<sup>36</sup> Frankfurt does not imply that these decisions must be highly articulate and reflective mental acts, but rather he accepts that the decision can occur when a person decisively accepts or “makes up her mind” that a particular desire is her own. Frankfurt, 1987, 38 and 40-42.

value stands above personal values and moral values and often determines which ID-commitments take priority.

I call these values *ultimate goods* and they are the ultimate ends for the sake of which people act. To find out the ultimate good behind a person's action, one need only continually ask *why* she behaved in a certain way. Some ultimate goods may include pleasure, power, happiness, revenge, obligation, or "because it was the right thing to do." A person may believe that one or all of these "goods" may satisfy her needs, desires, or felt obligations and therefore they become a good to her. Because these ultimate goods are ends for many of our actions, they often influence which ID-commitments take priority in a given situation.

For the sake of my analysis, I divide ultimate goods into subjective goods and objective goods. In general, a person desires a *subjective good* because it satisfies a need she has (e.g. hunger, fame, wealth, need to feel good about oneself). Objective goods are those a person desires because they are valuable for their own sake. As discussed above, an objective good also refers to goods that are recognized as not only good for oneself, but are objectively good for others.<sup>37</sup> Objective goods include moral values such as mercy and honesty, artistic expression, beautiful music, and standing up for the truth.

Ultimate goods play an important role in how we structure our identities. For example, having an ultimate subjective good to pleasure could result in an ID-commitment to acquire wealth. The ultimate good motivates the self to select ID-commitments that are coherent with achieving that good. The result could be a person

---

<sup>37</sup> I follow Augusto Blasi's description of objective goods without making any commitments to the metaphysical status of objective goods (Blasi, 2005, 81).



characterized by stinginess and greed. Or consider an ID-commitment to acquire wealth for the *objective* good of relieving suffering. We can see how a similar ID-commitment combined with a different ultimate good can shape a person's identity.

Ultimate goods are not the only values that prioritize ID-commitments and make for a coherent identity. A particular ID-commitment to some value and/or ultimate good may become central to a person's identity. In other words, of all of one's beliefs, desires, and ID-commitments, the self may elevate one in particular that defines one's identity.

Blasi, following Frankfurt, describes one way an identity can be strongly grounded by decisive identifications. I quote his description at length:

A special case of will structuring, particularly important for moral functioning, is when certain desires are not simply ordered on a quantitative scale of practicality, but are totally rejected, for example, as unworthy and bad, and made external to one's will; they are no longer open to volition, under any circumstances, even as they persist as desires. An even more stringent case is when a person cares so deeply about certain desires and about the special order of one's will that he or she wants to be guided by them also in the future. These commitments may be so decisive that they shape the core of one's identity. At this point it becomes unthinkable to intentionally engage in actions and projects that contradict the essence of one's will and identity.<sup>38</sup>

An identity so constructed forms a hierarchy of ID-commitments because the self identifies with some desires to a greater degree than others. These wholehearted or *higher-order ID-commitments* may resolve conflicts such as that between being loyal to a friend and being honest on one's job.<sup>39</sup> In this case, my ID-commitment to be honest may have greater priority within my identity than my ID-commitment to loyalty. Similarly, an ID-commitment to a particular ultimate good can determine which value

---

<sup>38</sup> Blasi, 2005, 80.

<sup>39</sup> Blasi, 2005, 79.

takes priority in a given situation. Higher-order ID-commitments play an important role when the self determines which actions are taken and rejected across situations. Some ID-commitments, when acted on consistently, become virtues or dispositions.

We finally come full circle to where a self, composed of beliefs, desires, and ID-commitments, can be said to have virtues. Consider again the case of More in regards to his virtue of piety. If we accept that a virtue is a stable disposition to act in a certain way across situations, then regular pious actions reveal his disposition to piety. In light of his willingness to die for his piety, it appears that piety is a higher-order ID-commitment that is integrated into his identity. Piety so constitutes his identity that to act against this ID-commitment would be to betray himself. In Chapter 1 I described such people as having Level 3 ID-commitments and explained that certain values are fully integrated into a person's identity.

More's and Socrates' dispositions, however, are not sufficient to explain these particular actions. Based on their own testimonies, we cannot separate their dispositions to justice and piety from their dispositions to be true to themselves. Consider that when More acts from the disposition of piety, he is also acting from the disposition to be true to himself. The two dispositions are indistinguishable in action because both express his identity. We can find the distinction in the objects of these dispositions. The object of More's piety is loyalty to God as represented by Catholic principles regarding divorce. The object of his integrity is his coherent identity constituted by this ID-commitment to piety. I discuss how coherence and the disposition to be true to oneself are related in the next sub-section.

To summarize this discussion, the selection and rejection of desires play an important part in determining what ID-commitments constitute a person's identity. How ID-commitments are prioritized can also determine the structure and stability of a person's identity. A person with ID-commitments to ultimate goods may organize her other ID-commitments to act for the sake of these goods. A person of integrity may also develop higher-order ID-commitments to values and/or ultimate goods thereby establishing a hierarchy of values. These higher-order ID-commitments may be used to resolve conflicts between other ID-commitments. Because the higher-order ID-commitments become a central part of a person's identity, being true to this higher-order ID-commitment can appear the same as being true to oneself. I suggest below that the virtue of integrity is difficult but not impossible to distinguish from these higher-order ID-commitments.

#### **2.4 - Coherence and the Virtue of Integrity**

Now that I have roughly described a few core items that constitute a person's identity, I explain how maintaining a coherent self-concept relates to the virtue of integrity. In Chapter 1 I argued that one necessary condition for being a person of integrity is having coherence among one's ID-commitments, motivations, and actions. I explained how a person of integrity avoids incoherence among commitments and between commitments, motivations, and actions. From common usage we understand that a person of integrity cannot have diametrically opposed ID-commitments (e.g. to justice and injustice) nor can she have an ID-commitment to justice while consistently acting unjustly. Similarly, her motivations must be consistent with her actions. These

requirements can be summarized as maintaining a coherent identity (i.e. beliefs, desires, ID-commitments, dispositions, ends, principles) and acting in accordance with that identity. What is missing from this account is what *motivates* a person to maintain a coherent identity.

In their description of a self presented above, Frimer and Walker suggest that the self is actually a unifying process that attempts to mend inconsistencies.<sup>40</sup> In an extensive review of research on moral motivation, Velleman also suggests that people tend to behave consistently with their self-conceptions and self-attributions.<sup>41</sup> Initially, it appears that the motivation for a self to maintain a coherent identity is a desire or concern for self-consistency and to avoid self-inconsistency. To understand this motivation, one only needs to reflect on times when he or she avoided a behavior simply because, “I am not that kind of person.”

It is true that not everyone has a desire for self-consistency nor does everyone care about maintaining a coherent identity. According to Blasi, however, it appears that a person of integrity *does* have a “serious concern for the unity of his or her subjective sense of self, as manifested in consistency with one’s chosen commitments.”<sup>42</sup> A concern for self-consistency can motivate a self to maintain coherence in the three areas mentioned above: among ID-commitments, between ID-commitments and actions, and between motivations and actions.

---

<sup>40</sup> Frimer and Walker, 2008, 346.

<sup>41</sup> Velleman, J. David. "From Self Psychology to Moral Philosophy." *Philosophical Perspectives* 14. (2000): 368.

<sup>42</sup> Blasi, 2005, 90.

A person of integrity may non-consciously or consciously realize that she has a “serious concern” for self-consistency. In the non-conscious cases, she may not realize it until she is close to compromising one of her ID-commitments. At this point a person may feel great anxiety because she is being forced to go beyond her coherent identity. Other signs that a person has a concern for self-consistency are feelings of deep regret and guilt after acting against her ID-commitments. For example Suzy, the World War II resistance member mentioned in Chapter 1, gave the Gestapo the addresses of the entire organization to save parent’s lives. Later in a prison camp, Suzy at first lied to fellow resistance members about telling the Gestapo, but she eventually admitted it with great grief. The leader of the resistance organization noted that after the war Suzy still suffered from her action and that, “it plays heavy on her conscience. She said that she never forgets it.”<sup>43</sup> At a non-conscious level, a person with a “serious concern” for self-consistency will feel great anxiety when she is pushed to violate ID-commitments that are inconsistent with her identity. She will also feel great guilt and shame if she violates her ID-commitments. At a conscious level, a person of integrity may daily reflect on the coherence of her self-concept and actions.

In paradigmatic cases of integrity, Blasi describes people like Socrates and More who have a heightened sense of identity and a concern for self-consistency. The concern for self-consistency is manifested by actions in accordance with their ID-commitments. In reference to how self-consistency relates to integrity, Blasi states,

---

<sup>43</sup> Kristen Renwick Monroe, *The Hand of Compassion: Portraits of Moral Choice During the Holocaust*. (New Jersey: Princeton University Press, 2004), 126.

This occurs when a person so identifies with her commitments, cherished values and ideals, that he or she constructs around them the sense of a central, essential self. This sort of appropriation determines what ‘really matters’ to the person; it establishes such a hierarchy among the person’s goals and concerns as to create a sense of subjective unity and lifelong direction, and provides one with a sense of depth and necessity in his being.<sup>44</sup>

He goes on to say that compromising one’s identity would be felt as the highest betrayal. Socrates and More come to mind in this context because they would rather suffer death rather than betray what “really matters” to their identities. The motivation of self-consistency appears to drive persons of integrity to maintain a coherent self-concept. It also follows that the stronger the motivation to act in accordance with one’s ID-commitments (and particularly higher-order ID-commitments), the more one develops a disposition to maintain a coherent self-concept. In other words, one develops or at least manifests a disposition to be true to oneself – the virtue of integrity.

Williams could object to this account of integrity by again asking why a separate virtue of integrity is needed if other virtues or even ID-commitments provide sufficient motivation to act across situations. He could argue that if a virtue is a disposition to act, then to include integrity is unnecessary. In fact, it seems like More’s virtue would be less admirable if he were primarily acting for the sake of being himself because integrity would seem like a self-indulgent virtue. Williams states, “If integrity had to be provided with a characteristic thought, there would be nothing for the thought to be about except

---

<sup>44</sup> Blasi, 2005, 92.

oneself – but there is no such characteristic thought, only the thoughts associated with the projects, in carrying out which a man may display his integrity.”<sup>45</sup>

I disagree with this characterization of integrity on two counts. Williams’ first concern is that the motivation behind integrity would be self-indulgence. While this may be true in some cases, it is not the typical motivation of integrity. In the two cases reviewed thus far, the main motivation behind being true to oneself appears to be a commitment to maintain a consistent and coherent identity, not to indulge or aggrandize the self.<sup>46</sup> If a person believes that she is generous, she may tend to act generously for the sake of consistency with her identity. If her identity is so constituted that generosity is a higher-order ID-commitment, then her motivation of self-consistency would be even stronger. But in neither case would we say her motivation is self-indulgent unless we had evidence of that motivation.

Williams’ second objection is that a person of integrity only thinks about her projects and not about herself. On his account we don’t explicitly think about our identities when acting. Socrates may not think about his identity when he goes home instead of completing his errand. A case could be made against this objection when we

---

<sup>45</sup> Williams, 1981, 49. Cheshire Calhoun raises a similar objection to the identity account of integrity presented in Chapter 2. Her objection is that integrity is seen as a personal virtue and this wrongly limits what can be said about the nature and value of integrity. I would respond that she mischaracterizes the identity account as a selfish protection of a self-conception. On my view, the foundation of integrity is a person’s identity and the actions that arise from it. I would agree, however, that the values and actions of the self are often found in the social realm and are not limited to an isolated self. Cheshire Calhoun, "Standing for Something." *The Journal of Philosophy* Vol. 92, No. 5, (May 1995).

<sup>46</sup> George Kateb has suggested that Socrates’ attitude toward justice could be interpreted as a proud statement about his integrity. Even if this were true, it may take away from our evaluation of his character but very little from our attribution of integrity given the sacrifice he was willing to make. George Kateb. "Socratic Integrity." In *Integrity and Conscience*. Eds. Ian Shapiro and Robert Adams. (New York: New York University Press, 1998).

consider that we tend to reflect on who we “really are” and what we stand for when pressured to compromise our core values. The statements by Socrates and More also demonstrate that they considered their identity when deciding whether or not they should risk their lives.

Another response is that Williams does not consider that integrity could be a motivational virtue that disposes one to act on ID-commitments to projects and ends. At the level of thought, Williams may be correct about a person not *thinking* about her identity, but as I have argued earlier he is mistaken in ignoring the possibility that integrity is a motivational virtue. On my account, integrity is a *disposition* to be true to oneself and not necessarily a *thought* about being true to oneself, though thoughts about oneself may come. Because Williams claims that integrity is not a disposition but an ID-commitment to personal projects, he does not consider the distinct motivation of self-consistency and therefore misses an important aspect of integrity.

Where does this leave us with the virtue of integrity? I suggest that the virtue of integrity can be described as a disposition with desires. One desire is a person’s conscious or non-conscious “serious concern” for self-consistency and coherence. In other words, maintaining a coherent self-concept is a high priority for the person. The second desire is the ID-commitment or higher-order ID-commitment that partly constitutes the self. To have an ID-commitment is to have decided that actualizing it is important and to be motivated to do so. It is beyond my analysis to say which desire provides the most motivation for individuals in different situations. My best description



of the disposition of integrity is that it is the self's serious concern for coherence as manifested by consistently acting on one's ID-commitments.

The final unanswered question for the virtue account of integrity is how, or if, it must contain moral content. If integrity is merely a disposition to be true to oneself, tyrants could have this virtue and use it to produce devastating harm. In the next subsection I explain how the virtue of *substantive* integrity requires some kind of moral content.

## **2.5 - Moral Integrity**

The virtue account presented thus far is not an account of substantive integrity, but rather an account of formal integrity. In Chapter 2 I described formal integrity as having ID-commitments that allow a person to violate moral values. Following common usage of the term, I argued that substantive integrity is reserved for people with ID-commitments to moral values and the determination to consistently act on them. The virtue of integrity also refers to these same people, but it captures another central meaning of integrity that I describe below.

Earlier I reiterated that a person can have ID-commitments to personal values (e.g. stamp collecting) and/or moral values (e.g. honesty, justice). I also claimed that a person can have an ID-commitment to ultimate goods. Ultimate goods are the ends at which a person's actions ultimately aim and I distinguished between subjective goods and objective goods. Subjective goods tend to satisfy a need or desire and include such ends as pleasure, advantage, fame, wealth, or power. Objective goods are those a person

desires because they are good in themselves and are often good for others. These goods include moral values, artistic expression, truth, and musical harmony.

Recognizing the broader category of ultimate goods helps explain a third aspect of integrity, especially in its role as a virtue. Before discussing this third aspect, it is important to note that the distinction between formal and substantive integrity does not change. The moral identity account already requires persons of substantive and formal integrity to have a disposition to be true to themselves. A person of formal integrity like Don Giovanni has an ID-commitment to a personal value of seducing women and arguably he has a disposition to be true to himself. However, having a disposition to be true to himself does not mean that he has a *substantive* virtue of integrity especially since his ID-commitment to personal values violates moral values. By consistently violating moral values, Don Giovanni is a person of formal integrity because he is not morally trustworthy. In regards to ultimate goods, it is safe to assume that Don Giovanni has ID-commitments to subjective goods such as pleasure and power. Because these subjective goods may require him to violate moral values in order to achieve them, they also provide evidence that he has formal integrity. I suggest that if Don Giovanni has the virtue of integrity, it is merely a formal virtue.

Next consider a person of substantive integrity like Socrates. He is a person of substantive integrity because he has an ID-commitment to the moral value of justice and the disposition to be true to himself. By “being true to himself” I mean that Socrates has a disposition to maintain a coherent self-concept that is constituted by ID-commitments to the moral value of justice. Socrates also differs from Don Giovanni because he has an

ID-commitment to the objective good of justice. Justice would be considered an objective good because it is good in itself and also a good for others. I conclude that Socrates has a substantive virtue of integrity because he is disposed to be true to a self that is constituted by ID-commitments to moral values that are also objective goods.

In addition to the formal and substantive cases, another type of integrity is needed to account for cases in which the virtue of integrity does not directly involve ID-commitments to moral values. Consider the cases of More and Flew who have ID-commitments to what appear to be non-moral values. More has ID-commitments to the decrees of the Pope and Flew has an ID-commitment to a principle concerning truth. On my account, one could argue that the objects of their ID-commitments are not moral values because they are not justified as important for instrumental and moral reasons relevant to the moral community. In fact, they appear to be justified by personal reasons because they can change their mind about these values without violating any reasonable obligations they have to the moral community (e.g. telling the truth; keeping one's word).

I label this third type of integrity *integrity of conscience* and it is a kind of substantive integrity. Integrity of conscience is a form of substantive integrity for three reasons: 1) a personal value is elevated to the status of a moral value; 2) acting on the ID-commitment does not violate moral values; and 3) we tend to consider the act of not compromising one's conscience an objective good. I take these reasons from a broad understanding of how integrity is commonly attributed to people who will not compromise their consciences even though the values they hold are not specific moral values.

First, a person of integrity of conscience elevates a personal value to the status of a moral value. More, Flew, and others who have ID-commitments to personal values may elevate a particular value to become part of their consciences. By conscience I mean the collection of ID-commitments and beliefs that one considers it morally wrong to violate personally. The process could be one of selecting a personal value and making it a higher-order ID-commitment which centrally constitutes one's identity. Making a personal value part of one's conscience changes it into what I call a *moral-personal value*. My purpose for labeling this value a moral-personal value is to indicate that for the person to violate this value is subjectively experienced as an immoral act. At the same time, the value is personal because unlike moral values, the person often has no reasonable obligation to the moral community to live out the value. Moral-personal values can take many forms and may include life projects (e.g. serve the under-privileged, become an exceptional teacher), personal lifestyle choices (e.g. vegetarianism, home schooling children, recycling), and religious principles (e.g. obey scripture, only eat Kosher food).<sup>47</sup>

The second and third reasons for considering integrity of conscience an aspect of substantive integrity explain why we tend to admire people with integrity of conscience. The second reason is that a person of integrity of conscience rarely violates moral values in order to act on the moral-personal value. To do so would reveal the person as having only formal integrity. Lynne McFall describes a person who violates moral values for a

---

<sup>47</sup> Religious principles can be described as moral values because one's religious community can have a reasonable expectation that a member will follow them. In this discussion, however, I am taking a third-party perspective on religious principles and how they inform integrity attributions.

moral-personal value. McFall imagines a person who decides that the only way to stop people from burning books in a *Fahrenheit 451* situation is to burn the book burners. The activist elevates his personal value of loving literature to a moral-personal value which leads him to kill the book burners lest he be untrue to himself. McFall concludes that while the activist may have some virtue of integrity, it is not integrity with moral content.<sup>48</sup> I agree with her conclusion because when the activist violates moral values in his quest to stop book burners, he becomes a person of formal integrity. If, however, he does not compromise moral values as he works to stop book burners, then he would be a person who has integrity of conscience and therefore substantive integrity.

The third reason for integrity of conscience being an aspect of substantive integrity is that we tend to consider the act of not compromising one's conscience to be an objective good. Integrity of conscience is good in itself partly because it expresses an admirable depth of commitment and resolve under pressure regardless of the consequences. Also, integrity of conscience is good for others because we tend to find their examples just as positive and inspiring as being true a moral value. I cannot fully explain why we admire integrity of conscience and why it seems to be good for others, but it seems to elicit admiration.

As evidence for my claim, consider the continued success of the play and movie *A Man for All Seasons* which is about the life of More. People are drawn to this story of a man who gently resists compromising his integrity of conscience. Immanuel Kant also describes the admiration that comes when people see a person of integrity resist

---

<sup>48</sup> Lynne McFall, 'Integrity.' *Ethics Vol. 98 No. 1*, (October 1987).

temptations. In responding to a person who asked him why teaching virtue accomplishes so little, Kant explains that an act of integrity (German: *Rechtshaffenheit*) naturally demands attention. He states that an act of integrity from duty, as opposed to those for subjective goods, “elevates the soul and awakens a wish to be able to act in like manner oneself. Even children of moderate age feel this impression and one should never represent duties to them in any other way.”<sup>49</sup> An act of integrity in accordance with moral and/or moral-personal values seems to be good for others.

To pull these three reasons together, I relate the story of an opponent of More who also demonstrated integrity of conscience. William Tyndale was a Cambridge scholar who held opinions that contradicted those of the Pope, English Bishops, and More. In particular, Tyndale was convinced that only the Bible should determine what the church does and therefore everyone should have access to a translation. Because Catholic leaders in England banned him from producing English translations, he fled to Germany in 1524 and began translating the Bible into English. He completed the New Testament in 1525 and started translating the Old Testament. More and other Catholic leaders attempted to suppress these copies. More’s dislike for Tyndale’s teachings drove him to write two volumes (his longest book) arguing against Tyndale’s views of the Church. He also sought Tyndale’s capture. In 1535 More was beheaded for his own convictions. As for

---

<sup>49</sup> Immanuel Kant. *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy*. Trans. Mary J. Gregor. (New York: Cambridge University Press. 1999), 65.

Tyndale, he was captured in Antwerp, Belgium and was executed by hanging in 1536. His final words were, "Lord! Open the king of England's eyes!"<sup>50</sup>

Both More and Tyndale appear to have the substantive virtue of integrity, but not specifically because they were morally trustworthy. The attribution of substantive integrity seems justified because their lives and actions demonstrate integrity of conscience. First, they both elevated a personal value to that of a moral value which became part of their identity and conscience. Second, their ID-commitment to these values (e.g. Tyndale's Bible for the common man and More's refusal to approve of the king's divorce) do not appear to violate moral values. In a sense this means that they have not "disqualified" themselves from having substantive integrity. Finally, their integrity of conscience appears to be an objective good that is a good in itself and in general good for others. Note that their moral-personal values need not be objective goods. The attribution of integrity comes from their being true to their *consciences* which many consider an objective good.<sup>51</sup>

My conclusion is that to have the substantive virtue of integrity, which in common usage indicates a person who is morally trustworthy, requires not only being true to oneself but that the "self" have at a minimum ID-commitments to moral values.

But as we have seen with the examples of More and Tyndale, a person's ID-commitment

---

<sup>50</sup> William Tyndale." *Encyclopedia Britannica*. 2010. Encyclopædia Britannica Online. Aug. 12, 2010 <<http://www.britannica.com/EBchecked/topic/611569/William-Tyndale>>. "Sir Thomas More." *Encyclopædia Britannica*. 2010. Encyclopædia Britannica Online. Aug. 12, 2010 <<http://www.britannica.com/EBchecked/topic/392018/Sir-Thomas-More>>. See also John Fox, *Foxes Book of Martyrs*, Ed. William Byron Forbush. Online book: <<http://www.ccel.org/f/foxe/martyrs/fox112.htm>> Aug. 12, 2010.

<sup>51</sup> Another example of two people with opposed integrity of conscience: One person willingly spends time in jail for being a conscientious objector while another person joins the military because her conscience does not allow her to sit by while an invader takes away the freedoms she holds dear.

to moral values may not be a reason why we initially attribute integrity to them.

Sometimes when we attribute substantive integrity to a person our first indication is that she has integrity of conscience. Therefore, in both the moral and conscience cases we can say a person has substantive integrity because their values are “moral” to the agent and they have not violated moral values in actualizing their ID-commitments.

One final note about integrity attributions that rely on integrity of conscience: integrity of conscience is a *prima facie* attribution when the moral trustworthiness of the person is not known. The attribution is based as much on the person having not violated moral values as it is on the person having an ID-commitment to moral-personal values. If a person has integrity of conscience but routinely violates moral values, then most people would disqualify her from having the substantive virtue of integrity. Consider Dr. Ernst Rüdin who was a leading racial hygiene expert and eugenicist in Germany before and during World War II. He was given awards for his work by the National Socialists and was committed to purifying the German race starting with the mentally handicapped. In a 1943 editorial he states that Hitler has furthered the goals of racial hygiene by carrying out the “fight against parasitic alien races such as the Jews and Gypsies.”<sup>52</sup>

If we assume Rüdin’s actions come from a disposition to be true to his conscience, then doesn’t he have the substantive virtue of integrity? I would argue that he has a formal virtue of integrity because even though he may act from a sincere ID-commitment to the moral-personal value of racial hygiene, his moral-personal value is contrary to moral values. Also, as in the Don Giovanni case, his ultimate good is a

---

<sup>52</sup> Quoted in Gershon, Elliot S. “Ernst Rudin, a Nazi Psychiatrist and Geneticist.” *American Journal of Medical Genetics* 74 (1997): 457.



subjective good. Eugenics in this case is good only for Germans and is not good for others.

## **2.5 - Sub-Section Summary**

In this section I have presented an account of what it means for a person to be “true to oneself” and described the items that constitute and guide a self. I have argued that a self is a dynamic process that integrates beliefs, desires, and ID-commitments into a coherent identity. For people with the virtue of integrity, the self has a serious concern and is therefore motivated to maintain self-consistency and to avoid self-inconsistency. The virtue of integrity is the disposition to maintain a coherent identity though it may be difficult to distinguish this motivation from the motivations arising from particular ID-commitments. I also explained how a person with the virtue of integrity may prioritize her ID-commitments to align with ultimate goods and/or higher-order ID-commitments.

Because both tyrants and heroes could have the virtue of integrity if it is limited to a disposition to be true to oneself as describe above, I argued that following common usage the virtue of integrity requires a self-concept constituted by ID-commitments to moral values. I also explained how integrity of conscience is an aspect of substantive integrity because it reveals an ID-commitment to a moral-personal value. I also argued that when a person is true to her conscience but does so in a way that violates moral values, then she has at best the formal virtue of integrity.

I can now revise the moral identity account’s necessary conditions using the insights above. First, to be a person of substantive integrity one must have coherent ID-commitments to moral values and/or moral-personal values and the determination to

consistently act on them. The first necessary condition implies that a person of substantive integrity does not violate moral values when actualizing ID-commitments to moral-personal values. Second, a person of substantive integrity must have a disposition to be true to her self. I described this motivational virtue as the conscious or non-conscious “serious concern” to maintain a coherent self-concept as constituted by her beliefs, desires, and ID-commitments to values. The motivation to be true to oneself can also come from one’s ID-commitments to moral values and/or ultimate goods.

Though I find this account of identity and integrity intuitively correct because of the moral exemplar cases it explains, the account’s accuracy can also be tested by whether or not it can answer concerns about identity, morality, and how a person’s integrity navigates the complexity of our modern age. In section 3 I respond to some of these concerns.

### **Section 3: Concerns and Questions**

Several questions about integrity can now be addressed using the moral identity account as revised above. Four specific questions that need to be addressed are: 1) Do “honest thieves” have substantive integrity? 2) Can a person with substantive integrity ever lie in order to actualize a “higher” moral value? 3) Can a person with substantive integrity alter or change her identity and still have integrity based on identity? 4) Can two people have opposed moral beliefs and both have substantive integrity? I respond to these questions using the moral identity account.

### 3.1 - Specific Virtues and Integrity

In response to the first question, so-called “honest thieves” or “courageous tyrants” do not have substantive integrity if their identity is constituted by ID-commitments to subjective goods even if they have a disposition to be true to themselves. What this means is that they embrace honesty for the sake of the power, pleasure, and wealth that their organization gains. If honesty does not produce these subjective goods then they will most likely violate this moral value. Committing to moral values as a means to pursue subjective goods disqualifies a person from substantive integrity by definition. In these cases, the thief or tyrant has the formal virtue of integrity. My discussion of Scar in Chapter 2 follows this same reasoning.

A similar concern is someone like Flew who demonstrates intellectual integrity or an artist who demonstrates artistic integrity. One may ask if they have substantive integrity. I would say that they may have integrity of conscience but that more information is needed. A person’s disposition to be true to a self-conception constituted by intellectual or artistic values is not evidence that they have ID-commitments to moral values. In Chapter 2 I discussed the case of the doctors who operated the Tuskegee syphilis experiments on African American men. I do not doubt that many of these doctors had identities that included an ID-commitment to intellectual integrity, but I highly doubt that they had an identity that included an ID-commitment to the moral value of respect for all people. Just as with integrity of conscience, intellectual, artistic, and professional integrity may be a sign that a person has an ID-commitment to moral values, but it is not a guarantee. Often more information about their lives is needed, which may

be why it is easier to attribute integrity to historical figures whose entire lives are known than it is to the living.

### 3.2 - Can you lie and still have integrity?

A second question is whether or not a person of substantive integrity can lie in order to actualize another virtue.<sup>53</sup> This concern is important because to say a person has substantive integrity often implies that she is never dishonest. I use the case of moral exemplar John Weidner to discuss this concern.

Case 3: John Weidner grew up in Switzerland, Belgium, and Holland. His family moved around because his father was a Seventh-day Adventist pastor and teacher. Growing up, Weidner hiked in the mountains between Holland, Belgium, France and Switzerland. During World War II he decided that he must help Jews escape the Nazis because he had mountaineering skills and he believed if you can help, you *must* help. He led a resistance operation in Holland and Switzerland.

Weidner was taught never to lie as a child and this became a central part of his identity. However, during the war he lied when the Gestapo questioned him about his rescue activities. When interviewed later about this seeming lack of integrity he states, "It was for me very natural to lie, to say, 'I don't know where are there [Jewish] people.' It was only after the war, did I say, 'Was it right or not?'" When asked if he questioned the rightness of lying at the time, Weidner answers strongly, "*It was right!* It was right. They are human beings . . . I wasn't lying so much to save my life but to save other people." Weidner says that he lied on these occasions because he had to make a choice between higher values. When asked what his highest value was at that time he answered, "Love your neighbor. You have to help." Weidner and his organization saved over 900 people.<sup>54</sup>

One could argue that Weidner lacked integrity during this period because he compromised his ID-commitment to honesty, a moral value. But this objection only

---

<sup>53</sup> McFall brings up this question as a challenge to the contention that integrity only means being honest. (McFall, 1987). Nancy Schaubert also points out problems with this view. Nancy Schaubert, "Integrity, Commitment, and the Concept of a Person." *American Philosophical Quarterly*, Vol. 33 No. 1. (January 1996).

<sup>54</sup> Monroe, 2004, 112.

stands if integrity is defined as automatically following an identity constituted by inflexible ID-commitments.<sup>55</sup> The moral identity account presented above allows flexibility because it recognizes that a person's self-conception can take on a *hierarchical structure*. The structure allows momentary compromises of one moral value to realize a higher moral value or objective good. As discussed above, some ID-commitments matter so much to a person that they create a hierarchy within a self-concept. In other words, the self uses a higher-order ID-commitment to prioritize her other ID-commitments. Prioritizing one's actions according to a higher-order ID-commitment can produce actions that consistently aim at the *highest* objective goods. I referred to the examples of Socrates, More, and Flew to support this description of a hierarchical structure, and now I add the example of Weidner.

While I cannot construct a complete priority list of objective goods, it is uncontroversial to claim that Weidner's ID-commitment to save innocent lives is objectively a higher good than honesty, especially when it requires telling the truth to those who would eventually murder the people one is protecting. What ultimately determines whether or not a person has substantive integrity is not rigidly living by specific moral values, which is rarely possible given the ebb and flow of life. However, substantive integrity can be grounded in structuring one's self-concept so that ID-commitments to the highest objective goods are its organizing principle.

---

<sup>55</sup> Critiques of a form of integrity that would require a rigid and unresponsive character are found in Martin Benjamin, *Splitting the Difference: Compromise and Integrity* (Kansas: University Press of Kansas, 1990). See also Jerry D. Goodstein, "Moral Compromise and Personal Integrity: Exploring the Ethical Issues of Deciding Together in Organizations." *Business Ethics Quarterly* 10, No. 4 (2000).

Interviews with other Holocaust rescuers and contemporary moral exemplars support this account of substantive integrity. In referring to this research, Blasi notes that, “a central theme in the interviews with moral exemplars is how moral commitments originate from and imbue their core identity . . . This is why, as many rescuers and exemplars explained, they felt they had no choice but to do what they had done.”<sup>56</sup> The bottom line is that a person of substantive integrity values moral values as objective goods and to act contrary to these goods is an act of self-betrayal. If a compromise is necessary for a higher objective good, the identity can be betrayed for a moment.<sup>57</sup>

Weidner's example also offers two important insights into the structure of a person's self-conception. First, minor compromises of a self-conception for a higher objective good do not necessarily change a person's identity or our attribution of integrity. Weidner subordinated one objective good for the sake of a higher objective good based on the situation. His ID-commitment to honesty, however, stayed intact because his temporal dishonesty did not fundamentally change the content or structure of his identity.

Weidner's life after the war provides two examples that support my claim. First, after the war Weidner still questioned if compromising his honesty was right or not. His ID-commitment to honesty was not compromised, but only subordinated as needed in

---

<sup>56</sup> Blasi, 2005, 95. He reference studies of Holocaust rescuers by S. P. Oliner and P. M. Oliner, *The Altruistic Personality: Rescuers of Jews in Nazi Germany*, (New York: Free Press, 1988) and also recent studies of moral exemplars by Anne Colby and William Damon, *Some Do Care: Contemporary Lives of Moral Commitment*. (New York: Free Press, 1992).

<sup>57</sup> Someone could argue that a person of integrity should *never* lie. This response is not intuitively correct regarding common notions of integrity. If a person did not lie in Weidner's circumstance, she may be seen as a person of *formal* integrity because she lacks an ID-commitment to the higher objective good of saving innocent lives.

order to realize a noble objective. If his identity had changed, he would not carry this moral concern about his incoherent identity. Second, not long after the war Weidner was offered five million Swiss francs (\$1 million) if he would declare that a German collaborator had actually helped the Allies so that the man could leave Switzerland. Weidner refused because lying during the war was not for gaining power, money or prestige (i.e. subjective goods). It was to save the lives of others (i.e. an objective good).<sup>58</sup>

The second insight about the structure of a person's self-conception is that acting with integrity may require deliberation, practical wisdom, and emotional control. Even when the hierarchical structure of a person's identity makes the person feel like they have no other option, she must still deliberate on how, when, and where to act. The contemporary accounts of integrity presented in Chapter 1 require a person to deliberate on the "reasonable person standard" to determine the *moral content* that should not be compromised.<sup>59</sup> In these accounts, dispositions are expected to fall in line with reason. In actuality it appears that dispositions take the lead and reason follows along to figure out how, when, and where the person can bring about an objective good. Aristotle held a similar view that the virtuous person uses practical wisdom (Gk: *phronesis*) to act in accordance with her character, or as one Holocaust rescuer stated, "The hand of compassion was faster than the calculus of reason."<sup>60</sup> The problem with overly

---

<sup>58</sup> Monroe, 2004, 116.

<sup>59</sup> See McFall (1987) and Gabriel Taylor, *Pride, Shame and Guilt: Emotions of Self-Assessment*. (Oxford: Oxford University Press, 1985).

<sup>60</sup> Monroe, 2004, 55.

emphasizing the “reasonable person” standard for integrity is that it presents reasoning as determining what an objective moral good is rather than something that coordinates the actions of one’s pre-existing ID-commitments to moral values.<sup>61</sup>

In interviews, moral exemplars like Weidner rarely considered *which* objective goods a reasonable person should actualize. Instead, they used practical wisdom to determine *how* to do what was morally right in each situation. Because of the constitution of their self-conceptions, they perceived what was objectively good in such a way that determining “what is right” was glaringly obvious and did not require significant reflection. Williams likewise argues that alienated reason does not make a decision, but the entire person must decide which action to pursue.<sup>62</sup>

A sobering story from Weidner’s life demonstrates how the self, even though motivated by a disposition to be true to himself, still uses practical wisdom to control one’s emotions. Once Weidner was at a train station where many Jews were being deported. A Nazi officer told a woman to quiet her baby, but the baby kept crying. The officer took the baby and stomped it to death on the platform while Weidner watched. Weidner explained to an interviewer why he did not act on his feelings to strike back at the officer. “Our reason has always dominated our emotions, our feelings. My first reaction was to do something. But I won’t save the baby. I won’t save the lady. And I will be out of the picture to help other people. So, to protest will give the satisfaction to

---

<sup>61</sup> Justin Oakley and Dean Cocking argue that the role of *phronesis* is that of a virtue coordinator. “This general regulative idea is what Aristotle calls ‘practical wisdom’ or *phronesis*, and this involves an understanding of the general good for humans, and the capacity to deliberate well such that one realizes virtuous ends in one’s responses to particular situations.” Justin Oakley and Dean Cocking, *Virtue Ethics and Professional Roles*, (New York: Cambridge University Press, 2001). 29.

<sup>62</sup> Williams, 1981, 52.



my own feelings, but [it will give] no result. So I didn't do it."<sup>63</sup> As is commonly reported among moral exemplars, a strong disposition to act on ID-commitments does not justify a lack of practical wisdom.<sup>64</sup>

My answer to the question above is that an *ideal* person of substantive integrity is seriously concerned about having a unified self-concept constituted by ID-commitments to the *highest* objective goods. It does not count against the person's integrity if she momentarily subordinates one objective good or moral value to bring about a higher objective good. Her substantive integrity must be grounded in the hierarchical structure of her self-conception which does not permanently change because of a momentary compromise. This idea person must also use practical wisdom in determining how, where, and when to act on her ID-commitments.

The objection could be made that on this account the syphilis doctors have substantive integrity. As long as they have a hierarchical structure within their self-concepts and ID-commitments toward an objective good (e.g. learning more about a terrible disease), then they can compromise a lesser objective good (i.e. experimenting on non-consenting people) without losing an attribution of substantive integrity. The problem with this objection is that it assumes a hierarchy of objective goods that is disputable. The objection assumes that researching the progression of a disease without the purpose of finding a cure is a higher moral good than deceiving infected adult human

---

<sup>63</sup> Monroe, 2004, 110.

<sup>64</sup> Based on his research and theorizing on emotions, Robert Solomon claims that, "A happy life with emotional integrity is not a life without conflict but a life in which one wisely manages emotional conflicts in conjunction with one's most heartfelt values." Robert Solomon, *True to Our Feelings*. (New York: Oxford University Press, 2007). 268

subjects.<sup>65</sup> On this issue most of society and the medical community has come down firmly on an opposite ordering of objective goods. In fact, if the subjects were not poor African Americans during a time of racial prejudice, true informed consent would have been required from the start. In response to the objection, it seems that the doctors were confused about which objective goods were higher than others. While my claim is not that testing on one non-consenting person to save the lives of millions could not be done by a person of integrity, I am claiming that in this case moral values were violated for a lesser objective good.

### **3.3 - Can a person change the contents of her identity and still have integrity?**

A common objection to identity accounts of integrity is that they rely on a rigid self-concept or identity from which a person acts. Contrary to this view, our intuitions tend to expect a person of integrity to change her ID-commitments based on new evidence about herself and the world. While I have discussed this above regarding practical wisdom, this objection concerns changing one's identity.

In their book *Integrity and the Fragile Self*, Damian Cox, Marguerite La Caze, and Michael P. Levine employ this objection to argue that integrity cannot be a kind of wholeness or merely having a core identity. They argue that, "Understanding integrity involves taking the self to be always in process, rather than static and unchanging or

---

<sup>65</sup> A concise description of the Tuskegee Study is by George E. Pence, "The Tuskegee Study" in *Contemporary Issues in Bioethics*, 7e. Ed. Tom L. Beauchamp, LeRoy Walters, Jeffrey P. Kahn, and Anna C. Mastroianni. (California: Thomson-Wadsworth, 2008).

containing an inner ‘core’ around which reasonably superficial changes are made.”<sup>66</sup>

Instead of integrity requiring a core identity, integrity requires a fragile identity.

Integrity is a virtue located at the mean of various excesses. On the one side we find conditions of capriciousness, wantonness, weakness of will, disintegration, hypocrisy, dishonesty, and an incapacity for reflection or self-understanding. On the other side we find conditions of fanaticism, dogmatism, monomania, sanctimoniousness, hyper-reflexivity and the narrowness and hollowed out character of a life closed off from the multiplicity of human experience.<sup>67</sup>

In other words, “The person of integrity lives in a fragile balance between every one of these all-too-human traits.”<sup>68</sup>

It is beyond the scope of this paper to review each objection Cox et. al. use against generic identity accounts, but the strongest objections attack any reliance on an inflexible central identity around which only superficial changes are made. On their view, the identity account presents us with a person of integrity who protects her identity without adjusting to the outside world. My moral identity account could fall into this category if it presented the self and its identity as inflexible to the external world and new evidence. Martin Benjamin explains the challenge to rigid identity-based integrity that must confront an ever changing world. “Either we lead perfectly consistent but rigid and perhaps fanatical lives or we respond to new circumstances and understanding at the expense of integrity and our identity as persons.”<sup>69</sup> The first way can lead to fanaticism and the second way could lead to disintegration. Cox et. al. argue that the obvious

---

<sup>66</sup> Cox, et. al., 2003, 41.

<sup>67</sup> Cox et. al, 2003, 15.

<sup>68</sup> Cox et. al, 2003, 41.

<sup>69</sup> Benjamin, 1990, 59.

solution to the dilemma is that the virtue of integrity is a mean between the two vices with little need for a stable identity.

I disagree with Cox, et al.'s conclusion concerning a person's psychological structure, but I agree with their conclusion about the self as a process. First, Cox et. al.'s argument that integrity is fundamentally a fragile balance between fanaticism and disintegration does not fit with the empirical research on moral exemplars who appear to have anything but a "fragile" identity. Moral exemplar research confirms that people do have what could be called an identity and evidence for it is found throughout their lives. Also, the existence of an identity does not entail that it is rigid and unchangeable. Flew's example of intellectual integrity and change stands in stark contrast to Cox et. al.'s depiction of an unchangeable, rigid identity. It is true that some beliefs and ID-commitments that constitute a person's identity are not easily changed, but "not easily changed" is different from "fanatically held."

This point leads me to my agreement with Cox et. al.'s notion that the self is a process. The moral identity account recognizes that one's identity is always in process. The self that has a serious concern for self-consistency works to maintain a coherent identity, but this process has both an internal and external role. The self's internal role involves maintaining a coherent self-concept and narrative so that ID-commitments do not conflict and that actions are consistent with who one is. Because generic identity accounts have not sufficiently developed an account of the self, Cox and others fear that too much internal focus will lead to a rigid identity that rejects evidence that would require a change in one's ID-commitments.

The moral identity account addresses the concern by recognizing that the self is also a participant in the external world and must process new situations and evidence so that appropriate responses can be made. Larry May's communitarian account of integrity describes integrity as a *realistic* commitment. He states, "It may be important that one have, at any given time, very strongly held commitments. But it is implausible to think that an integrated or committed person must hold certain unshakable commitments over the course of his or her life."<sup>70</sup> May recognizes an identity that can accommodate ID-commitments to moral values while realistically allowing a person to mature without losing her substantive integrity.

Admittedly, not everyone who has a coherent identity constituted by ID-commitments can easily evaluate her identity, and we may not want someone like Weidner to easily change his moral commitments. One practice that may help a person of integrity open herself to new evidence and change her ID-commitments is to reflect on and be open to new evidence about objective goods and how to reach them within the limits of moral values. My contention is that the difference between a dogmatic person and a person of integrity is a lack of openness to the world and how her ID-commitments impact that world.

When it comes to changing one's ID-commitments, even a tyrant with formal integrity can realize that the subjective good of advantage is blocked if she dogmatically holds on to ineffective ID-commitments. The moral identity account allows her to

---

<sup>70</sup> Larry May, "Integrity, Self, and Value Plurality," *Journal of Social Philosophy*, Vol. 27 No. 1. (Spring 1996): 134. I rely on a stronger notion of identity than May, but I agree with him that maintaining one's identity is a process that requires more than holding unshakable beliefs.

change her ID-commitment from injustice to justice for the sake of effectively attaining subjective goods. Or she could come to realize during a “dark night of the soul” that an identity that pursues advantage and pleasure produces a hollow and lonely life. This type of “conversion” experience may lead her to re-prioritize her ID-commitments and thereby change her self-conception completely. Holocaust rescuer Oscar Schindler appears to have made this change. He exchanged his subjective good of pleasure as a business man and unfaithful husband for the objective good of saving Jews at great cost to himself.<sup>71</sup> I would tentatively suggest that if his new identity was constituted by ID-commitments to moral values for the sake of objective goods for the remainder of his life (which it wasn’t), then he would change from being a person of formal integrity to one of substantive integrity. Or maybe he maintained a certain integrity of conscience in parts of his life, but was not fully a person of substantive integrity.

In response to Cox et. al.’s account, I recognize their concern that identity may breed dogmatism, however that result is not inevitable. I have argued that a self that is properly open to the reality of the outside world can integrate new ID-commitments and reject old ones without completely losing her identity. If Schindler had changed his ID-commitments after the war, he would not have completely lost his identity but instead gained a new identity that pursues a different ultimate good. I discuss the topic of commitment change in Chapter 5 when I address cognitive dissonance and the desire to rationalize one’s behavior.

---

<sup>71</sup> For a discussion of Schindler’s change in character see Lawrence Blum, *Moral Perception and Particularity*. (New York: Cambridge University Press, 1994).

### **3.4 - Can two people have opposed moral beliefs and have substantive integrity?**

Can a person who supports the right to have an abortion and a person who supports defending the life of a fetus both have integrity? How about a person who has moral reasons for supporting slavery and a person who has moral reasons for wanting to abolish slavery? These are important questions because throughout history we find putative people of integrity on different sides of these issues. Consider Robert E. Lee and Abraham Lincoln. Lee defended slave-holding states as commander of the Confederate Army while Lincoln led the Union to defeat slave-holding states. To determine if people with different moral commitments have substantive integrity requires considering our intuitions regarding a variety of cases. To answer this question, I first consider integrity as we use the term when assessing the whole life of historical figures and then how we use it to assess people living today.

#### *3.4.1: Historical Assessment of a Whole Life*

In discussing More and Tyndale I mentioned how integrity attributions apply to a historical person's entire life. To say, "Abraham Lincoln was a man of integrity" indicates that his life *on the whole* expressed a uniform devotion to ID-commitments to some objective goods. We do not attribute integrity to Lincoln because he was President or because of his policies. I have argued that substantive integrity refers to coherent ID-commitments to moral values as well as the objective goods at which they aimed. One could say that Lincoln meets this standard because he consistently lived out his ID-commitments across an entire life. In my review of Lee's life, he also consistently lived out his ID-commitments to objective goods. But how could both of these leaders have

substantive integrity if Lincoln and Lee were fighting against each other much like More and Tyndale?

My suggestion is that both Lincoln and Lee honorably followed their consciences throughout their lives. Lee's conscience required him to protect the sovereignty of Virginia and resist what he thought was an unjustified federal government invasion. When one reads about Lee's life, one learns that he was not fighting to protect the property rights of white slaveholders. Lee is considered a man of integrity because of his consistent honorable behavior that revealed ID-commitments to the good of his state, his troops, his family, and his pursuit of certain objective goods. Note that I am not comparing Lincoln to a person with an ID-commitment to slavery because this would be to compare a person with substantive integrity to a person of formal integrity.

As I mentioned above, sometimes to attribute integrity of conscience to a person requires us to find out more information about her life. After my research on his life, I consider Lee a person of substantive integrity even though he held slaves until he died. I make this claim because his letters reveal that he believed slavery was wrong, but he was tied to a culture and economy that would collapse without it. As a rough comparison, imagine what would happen if people in the United States suddenly had to stop buying any goods made in China. Lee, like many Southerners, felt that it was the white man's burden to ease the South out of slavery. In a letter to his wife Mary regarding freeing over 100 of his late father-in-law's "servants," Lee wrote, "Those who are hired out can soon be settled. They can be furnished with their free papers & hire themselves out. . . . Any who wish to leave can do so. The men could no doubt find homes, but what are the



women and children to do?”<sup>72</sup> His sentiments and actions demonstrated some ID-commitment to the objective good of protecting the innocent. Lee further demonstrated a commitment to respect for all people after the war. For example, one Sunday a black man went up to take communion in Lee’s church. No white parishioners moved. Lee left his pew and knelt down next to the man. The rest of the congregation followed.<sup>73</sup>

Lincoln and Lee both had some type of substantive integrity as evidenced by their ID-commitments to objective goods which happened to include moral values in both cases. Also, they would at least have a *prima facie* substantive virtue of integrity because they lived a life characterized by being true to their consciences. In essence, they did not compromise their ID-commitments during their lifetimes even though their place in the narrative of history unfortunately led them to oppose each other.

#### *3.4.2: Current Attributions of Substantive Integrity*

If we distance ourselves from people who have lived an entire life, we can ask another question about people who currently hold different views on moral issues such as abortion. Can two people who have ID-commitments to opposite objective goods both have substantive integrity? I must first note that it is difficult to attribute integrity, defined as having ID-commitments to objective goods, to a person who is fighting for a cause. The difficulty is that integrity refers to a person’s self-conception and not to a particular cause she takes on some time in life. By “cause” I mean a movement of a group of people to change the behaviors of others or to defend some right. Because

---

<sup>72</sup> Roy Blount, Jr. *Robert E. Lee* (New York: Viking, 2003), 109-110.

<sup>73</sup> Blount, 2003, 153.

causes can be adopted and rejected easily for various reasons, they have only a minimal connection to integrity which rests on a person's ID-commitments to moral values and objective goods more than to the content of a cause.

With these factors in mind, imagine a person Mary who argues that the right to an abortion is morally required while Jane argues that saving the life of an unborn child is morally required. Neither of these stands in themselves would lead to an attribution of substantive integrity. If we move to another level of commitment, we can imagine them both giving of their time and money to the pro-choice or pro-life causes. Mary becomes a Senator to maintain pro-choice legislation while Jane volunteers at a Birth Right center to help teens adopt out their children. Who has substantive integrity? Intuitions do not provide much guidance because attributing substantive integrity to someone who is alive now requires evidence about their ID-commitments to moral values more than their support for a cause.

In these instances, before attributing integrity we would have to know if the person's work for a cause was truly part of her identity and for some objective good. If Mary agrees to vote "No" on an abortion rights bill in exchange for a \$1 million campaign contribution, then her true ultimate good of power is revealed. Compromising her professed ID-commitments to the objective good of protecting women's rights reveals that she does not have substantive integrity. If she turns down the money, then we have evidence by her actions that she is a person of substantive integrity. A similar scenario could be presented for Jane.

I am not implying that a person must be tested before they have integrity, but that a test reveals to them and to outside observers that they actually do or do not have ID-commitments and a coherent identity. The deciding factor in these cases is not the cause, but the person's ID-commitment to moral values and the determination to act consistently. Consider two further cases. A person who never tells a lie may have substantive integrity while a person who lobbies Congress for tougher laws against lying may not have substantive integrity. Or, a person who always keeps his promises may have substantive integrity, while a person who risks his life to bring about civil rights change may not have substantive integrity.

A much clearer way to present the abortion conflict as a case of substantive integrity is from a first-person perspective. Mary becomes pregnant, has an ID-commitment to the right to choose as well as an ID-commitment to being a successful movie star. Jane also becomes pregnant and has an ID-commitment to protect all human life as well as an ID commitment to become a successful movie star. Both of them realize that the effort required in caring for a child will significantly restrict their career plans. If either Mary or Jane has an abortion, they may be a person of formal integrity because they followed their ID-commitment to have successful careers which is most likely a subjective good for advantage and pleasure. However, if neither Mary nor Jane has an abortion, only Jane appears to have substantive integrity because only she followed her ID-commitment to protect all human life, putatively an objective good. Assuming Jane is motivated by something besides an ID-commitment to the child (e.g. likes the idea of being a mom), she does not appear to have substantive integrity in this

case because she did not act from an ID-commitment to actualize an objective good. While these cases are no where near definitive, they are rough sketches of how integrity attributions can be made using the moral identity account.

From both the third-person and first-person perspective, I suggest that we can attribute substantive integrity to people with opposing moral commitments as long as their ID-commitments to moral values or an objective good are consistently acted upon and they do not violate moral values. The unanswered question in the abortion cases is whether or not protecting abortion rights or opposing abortion as murder are objective goods and which is a higher objective good. It is beyond the scope of this paper to answer these questions, however, it appears that people with substantive integrity could hold either position.

## **Conclusion**

In this chapter I have argued that integrity is the virtue of being true to oneself and I have supported my account of integrity with an analysis of how a person of substantive integrity constitutes a self-concept with ID-commitments to moral values and objective goods. I have also answered four questions that have threatened the plausibility of the moral identity account.

At this point I have described the necessary conditions for having substantive integrity. First, to be a person of substantive integrity one must have coherent ID-commitments to moral values and the determination to consistently act on them. This condition assumes that a person of substantive integrity does not violate moral values when actualizing ID-commitments to moral-personal values or objective goods. Second,

a person of substantive integrity must have a disposition to be true to her self-conception. I described this disposition as having a serious concern to maintain a coherent self-conception and the motivation for this concern can be an ID-commitment to self-consistency and/or acting from one's ID-commitments.

In Chapters 4 and 5 I address two challenges to integrity from moral psychology. In Chapter 4 I address the claim that most people do not act consistently across situations. Empirical evidence seems to show that stable mental structures do not exist, which implies that ID-commitments may not exist or may exist but only produce consistent behavior in predictable situations. In Chapter 5 I respond to the argument that the quest for integrity is dangerous because of the human tendency to rationalize unethical conduct. If most people rationalize immoral behavior in their quest for integrity, then the quest can lead people to deceive themselves about their true identities.

## **Chapter 4 - Integrity, Moral Identity, and the Power of Situations**

Most people want to work for a “person of integrity.” They also expect their political leaders to be “persons of integrity.” Others have a personal desire to be persons of integrity themselves and set personal goals to reliably live by moral values, regardless of the situation. But are these expectations of integrity merely stories we tell ourselves about the steadfastness of character? Some social psychologists and philosophers would say that they are just folk psychological ideas that have little empirical support. It is not that they do not want morally trustworthy bosses, leaders, and citizens. The concern is that given the “right” situation, people tend to compromise their moral values. Extensive research seems to indicate that in general people fail to act morally when confronted with ambiguous or unexpected situations.

These empirical findings seriously challenge the moral identity account of integrity. The apparent lack of consistent behavior across situations presented in social psychology experiments has led situationists, those who claim that situational variables primarily determine behavior, to be skeptical about the existence of character traits and dispositions. This skepticism reaches to the existence of integrity. To counter this challenge, I must demonstrate that substantive integrity is possible across a variety of situations.

On my account of substantive integrity, it is a necessary condition that a person have identity-conferring commitments (i.e. ID-commitments) to moral values such as honesty and the determination to consistently actualize these values across diverse situations. These ID-commitments as well as the motivation to be self-consistent appear

as a disposition to be true to oneself and result in honest behavior. Situationism suggests that objective situational factors determine a person's honest behavior more than stable character traits. Hence, a situationist account of integrity would conclude that consistent honest behavior most likely results from living in predictable surroundings. Also, unpredictable situations may lead the person to compromise her ID-commitment to honesty and reveal her disposition as merely a weak desire propped up by situational factors. If this is true, integrity may be more of a situational construct than a stable mental construct grounded on ID-commitments.

In this chapter I respond to this challenge by developing an empirical account of stable mental constructs without discounting the fact that situational factors can influence behavior. In section 1 I review situationist arguments against the existence of character traits and dispositions.<sup>1</sup> I then briefly review two experiments central to situationist arguments against stable mental constructs: the Princeton Seminary experiment and the Milgram Obedience experiments. In section 2, I present recent research in social cognitive psychology that offers a more complete explanation of behavioral consistency while preserving some common notions of stable mental constructs. The research centers on the mental construct of moral identity which I define as a self-conception that is composed of ID-commitments to moral values.<sup>2</sup> In section 3, I apply moral identity

---

<sup>1</sup> I do not defend a general account of character traits in this dissertation. Instead I argue that some people have stable mental constructs and that these are often perceived as character traits. Later I note criticism that situationists have narrowly defined what it means to have a character trait or disposition which has excluded ID-commitments and goals.

<sup>2</sup> The word "identity" may concern or confuse some readers because of the variety of meanings it has across philosophy, psychology, and neuroscience. I use the term "identity" to describe particular aspects of a person's self-conception that contain particular values that the person focuses on and by which she

findings to the Princeton Seminar and Milgram experiments and suggest that some subject behaviors demonstrate the stable mental construct supported by moral identity. I conclude my critique of situationist arguments with a review of what situationist and moral identity findings add to the moral identity account of integrity.

## **Section 1: Situationism and Mental Constructs**

In this section I present the situationist challenge to the existence of character traits and dispositions. I then review two research experiments that provide convincing evidence for the situationist position. I conclude by pointing out the problem the situationists must still answer: situational factors do not explain all behavior variance in these experiments.

### **1.1 - The Situationist Challenge**

John Doris' ground-breaking book *Lack of Character* meticulously presents the situationist case that empirical evidence does not support the existence of stable mental constructs, dispositions, or as I will refer to them, character traits.<sup>3</sup> According to Doris, traditional accounts of virtue and personality often assume the existence of "robust

---

defines herself. In my usage I am following Albert Bandura's social cognitive model because his conceptualizations are central to much moral identity research. Briefly, here is his conceptualization of how a personal identity is constituted by other identities. "Identity formation is an ongoing process not one characterized by fixedness in time. Moreover, the self view is multifaceted rather than monolithic. There are many aspects of the self. They are not equally salient, valued, or functional in different spheres of life or under different circumstances. In a dynamic, multifaceted model, continuity of personal identity requires neither high consistency among different aspects of self nor invariance across different social environments or domains of functioning. For given individuals, their personal identities are likely to be composed of unique amalgams of identities with social, political, ethnic, occupational, and familial aspects of life. Thus for a particular individual, a strong occupational identity may coexist with a moderate ethnic identity and a weak political identity without any felt discordance. Another individual may exhibit a quite different constellation of identities, combining a strong ethnic and political identity with a weak occupational identity." Albert Bandura, "Social Cognitive Theory of Personality," *The Coherence of Personality: Social Cognitive Bases of Consistency, Variability, and Organization*. Ed. Daniel Cervone and Yuichi Shoda. (New York: The Guilford Press, 1999), 231.

<sup>3</sup> John Doris, *Lack of Character* (New York: Cambridge University Press, 2002).



character traits” that produce corresponding behavior. He states, “Virtues are supposed to be *robust* traits; if a person has a robust trait, they can be confidently expected to display trait-relevant behavior across a wide variety of trait-relevant situations, even where some or all of these situations are not optimally conducive to such behavior.”<sup>4</sup> Given these characterizations, situationists make two specific claims concerning character traits. First, empirical evidence does not support the assertion that robust character traits are the cause of consistent behavior. Second, explaining behavior by referencing character traits commits an error because situational factors may fully explain the behavior.

Regarding the first claim, situationists are skeptical about the existence of robust character traits that reliably produce behavior across situations. Doris provides ample evidence from philosophy and personality psychology that people believe that robust character traits exist.<sup>5</sup> He then explains the position as a conditional: “If a person possesses a trait, that person will engage in trait-relevant behaviors in trait-relevant eliciting conditions *with markedly above chance* probability  $p$ ,” which he takes to be substantially above the probability of chance. Doris further qualifies this conditional by noting that trait-relevant behavior is that consistent *individuating behavior* that “is outside the population norm for a situation—that counts as evidence for trait

---

<sup>4</sup> Doris, 2002, 18.

<sup>5</sup> Doris, 2002, 15-27. Doris cites extensive research by Walter Mischel who challenged the established view that global personality traits existed. In one article Mischel states, “Individuals show far less cross-situational consistency in their behavior than has been assumed by trait-state theories. The more dissimilar the evoking situations, the less likely they are to produce similar or consistent responses from the same individual.” Walter Mischel, *Personality and Assessment*. (New York: Wiley, 1968), 177.

attribution.”<sup>6</sup> For situationists, to demonstrate that a trait exists requires the trait-relevant behavior occur in a situation that is unfavorable enough to attribute the behavior to the trait rather than a situational factor—a perfect description of a situation in which a person of integrity should reliably maintain her moral values.

The second claim is that if situationist skepticism about robust character traits is correct, it follows that we often commit the *fundamental attribution error* when we explain the causes of behavior by referring to character traits. For example, we may attribute callousness to Jim when he rushes past a person who fell in the subway, but we didn’t know that Jim must rush to pick up his sick daughter. Gilbert Harman argues that we often commit this error because we overlook situational factors and instead focus on character traits. However, if character traits don’t exist, then our attributions are incorrect and we should consider situational factors before passing judgment.<sup>7</sup>

Situationists do acknowledge that we see behavioral consistency in our interactions with others and in our own lives. Without robust character traits, where does this consistent behavior come from? The situationist explains that we often choose the situations in which we work and live and thereby create a stable environment that supports consistent behaviors. Priests, physicians, and rock stars place themselves in situations that reinforce consistent behaviors, thoughts, and feelings. We move and breathe in predictable environments that support consistent behavior. Another factor that makes us see consistency is that we construe situations in consistent ways. Because

---

<sup>6</sup> Doris, 2002, 19.

<sup>7</sup> Gilbert Harman, “Moral Psychology Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error”, *Proceedings of the Aristotelian Society* (1999): 323.

individuals subjectively interpret and construe the world differently, individual differences exist and situational factors will not produce the same behavior in all people. These individual differences, however, are rarely stable and should not be considered robust character traits.<sup>8</sup>

In summary, the situationist account rejects robust character traits as the primary explanation of consistent behavior. They argue that objective situational factors play a greater role in determining behavior than character traits. The situationists also argue that we often commit the fundamental attribution error when we attribute character traits to a person as a way of explaining their behavior. If robust character traits do not exist, the situational factors must provide the primary explanation for behavior. Finally, situationists argue that consistent everyday behavior arises from living in consistent situations and subjectively construing the meaning of situations in a consistent way.

Situationists support their account by citing hundreds of social psychology experiments that reveal people producing surprisingly inconsistent and uncharacteristic behaviors primarily because of objective situational factors. My concern with the situationist interpretation of these experiments is that it casts doubt on the existence of integrity seeing as it requires consistent moral behavior across diverse situations.<sup>9</sup> In the

---

<sup>8</sup> For an extended discussion of subjective construal, see Chapter 3 in Lee Ross and Richard E. Nisbett, *The Person and the Situation*. (Philadelphia: Temple University Press, 1991).

<sup>9</sup> Doris expands a general expectation of consistent behavior to the quality of integrity, which he defines as a form of “wholeness”. Under this definition, a fanatical Nazi has integrity if he cannot be bribed to spare Jews. Doris’ use of the term integrity as “wholeness” follows Bernard Williams’ and Gabriel Taylor’s use of the term. However, as discussed in Chapters 1 and 2, integrity meaning “wholeness” as opposed to “uncorrupted” only captures one possible meaning of the word and ignores how it is commonly used to indicate a moral virtue. It should be noted that even though Williams defines integrity as maintaining one’s most important life projects, he does allow that it is a complex quality that can have moral meaning. When presenting the example of the fanatical Nazi, he states, “A less fanatical Nazi who was moved not by

next sub-section I consider two of the most cited experiments and consider whether or not they prove the situationist's case. The experiments are the Princeton Seminary Experiment and the Milgram Obedience Experiments.<sup>10</sup> Situationists cite the Princeton Seminary experiment as proof that compassion, which is often thought of as a common moral value, is not resistant to situational factors. They also cite the Milgram experiments to demonstrate that not harming others, another common moral value, can be manipulated by situational factors as well.

After reviewing these experiments from the situationist perspective, I evaluate them from a moral identity perspective. My purpose in evaluating these experiments is not to defeat the situationist thesis regarding robust character traits. My purpose is to suggest that regardless of the unusual situational factors that influenced the behavior of many experimental subjects, some people were true to their moral identity, a characteristic of having integrity.

## **1.2 - Princeton Seminary Experiment**

Moral psychologists John Darley and Daniel Batson created an experiment to determine how situational factors influence helping behaviors.<sup>11</sup> They took as their model

---

bribes, but by the pleas of the Jews would be thought by fanatical Nazis to lack integrity, but probably not by the humane, perhaps because they entertain an idea (they certainly do not want to discourage it) to the effect that this was not a lapse or a weakness, but a rediscovery." Bernard Williams, "Replies", in *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Ed. J.E.J. Altham and Ross Harrison. (New York: Cambridge University Press, 1995), 213. Also Gabriele Taylor, "Integrity." *Proceedings of the Aristotelian Society, Supplementary Vol. 55* (1981) and *Pride, Shame and Guilt: Emotions of Self-Assessment*. (Oxford: Oxford University Press, 1985).

<sup>10</sup> These cases are used by the following situationists to suggest skepticism about robust character traits: Doris, 2002; Harman, 1999; and Ross and Nisbett, 1991.

<sup>11</sup> I understand that this experiment may not meet the highest standards for testing a particular variable given the small sample size (N=40) and reliance on a few subjective judgments made by the researchers

Jesus' parable of the Good Samaritan. Jesus tells the parable after a man asks him, "Who is my neighbor?" in the context of the command to "Love your neighbor as yourself." In the parable a person going from Jerusalem to Jericho is robbed, beaten, and left for dead. Even though a priest and a Levite (i.e. a religious temple worker) see the person, they walk on the other side of the road and offer no help. Then a Samaritan (a person from Samaria with whom the leaders of Jesus' time did not associate) comes upon the injured man. The Samaritan has compassion on him, binds his wounds, sets him on his beast, takes him to an inn and takes care of him through the night. The next day he pays the innkeeper saying, "Take care of him; and whatever more you spend, I will repay you when I come back." Jesus then instructs his questioner to "Go and do likewise."<sup>12</sup>

Darley and Batson were curious about the situational and personality differences of the priest, Levite, and Samaritan regarding helping behavior. The variables they decided to test were the content of one's thinking, the amount of hurry during a journey, and the dispositional variable of different types of religiosity.<sup>13</sup> In the experiment, 40 students at Princeton Theological Seminary were given personality tests that measured different types of religiosity. They were then told that they would give a 3-5 minute speech on a passage. Some were asked to read the parable of the Good Samaritan and others read a discussion of ministering in the professional clergy. After being informed

---

and confederates. I address the study primarily because situationists quote it widely and it does appear to record some effect of situational factors on subjects.

<sup>12</sup> Luke 10:29-37.

<sup>13</sup> John M. Darley and C. Daniel Batson, "'From Jerusalem to Jericho': A Study of Situational and Dispositional Variables in Helping Behavior", *Journal of Personality and Social Psychology* 27, No. 1, 1973.

that the speech would be given in another building, the research assistant put some subjects in the *high-hurry condition* by telling them, “Oh, you’re late. They were expecting you a few minutes ago.” He put others in the *intermediate-hurry condition* by saying that the other assistant is ready so “please go right over.” He put the third group in the *low-hurry condition* by saying that they had a few minutes but they might as well go over and wait there.<sup>14</sup>

All of the subjects had to pass through an alley where a “victim” was slumped in a door way coughing and not moving. He coughed twice and groaned. If subjects stopped and asked him if he was alright, the victim told them he had just taken some pills and that he just needs rest. The victim then rated the subjects on how much they helped on a scale of 0 (failed to notice) to 5 (refuses to leave the victim and insists on getting him help).

The results of the study pointed to situational factors influencing behavior more than the type of passage the subjects read as well as religious dispositions. By situational variables, 63% of those in low-hurry offered help, 45% of those in intermediate hurry offered help, and 10% of those in high-hurry offered help. The conclusion was that the hurry condition was significantly related to helping behavior while the passage read and the measures of religiosity were not statistically significant. The situationist points to the results of this experiment as strong evidence that situational factors explain behavior more than robust character traits. Lee Ross and Richard Nisbett conclude that, “these findings tell us little if anything about the personal dispositions of seminarians but a great

---

<sup>14</sup> Darley and Batson, 1973, 104.

deal about the situational determinants of altruism.”<sup>15</sup> Doris muses that, “in this case the demands of punctuality seem rather slight compared with the ethical demand to at least check on the condition of the confederate.”<sup>16</sup> Harman claims that typical interpretations of the Good Samaritan parable most likely commit the fundamental attribution error of overlooking situational factors for not helping.<sup>17</sup> Situationists conclude that objective situational factors influence helping behavior more than the ubiquitous character trait of compassion.

### 1.3 - Stanley Milgram’s Obedience Experiments

In the early 1960’s, psychologist Stanley Milgram set out to answer an important question regarding authority and obedience: in a laboratory setting, “if an experimenter tells a subject to act with increasing severity against another person, under what conditions will the subject comply and under what conditions will he disobey?”<sup>18</sup> With this question in mind, Milgram conducted elaborate laboratory experiments from 1960 to 1963 at sites in New Haven, CT (Yale University) and Bridgeport, CT (building downtown). These primarily involved male adults between 20 and 50 years old from all walks of life.<sup>19</sup> Each experiment used 40 subjects who answered a newspaper ad promising them \$4.50 to help in a learning study. In the majority of the experiments, the

---

<sup>15</sup> Ross and Nisbett, 1991, 49.

<sup>16</sup> Doris, 2002, 34.

<sup>17</sup> Harman, 1999, 324.

<sup>18</sup> Stanley Milgram, *Obedience to Authority: An Experimental View* (New York: Harper and Row Publishers, Inc, 1974): xii.

<sup>19</sup> The experiment description is found in *Obedience to Authority* and Stanley Milgram, “Some Conditions of Obedience and Disobedience to Authority” in *The Individual in a Social World*. Ed. John Sabini and Maury Silver. (New York: McGraw-Hill, Inc., 1992). 136.

subject was welcomed by an experimenter who wore a white lab coat and another person who was supposedly another paid subject. In actuality, the other person was an accomplice to the experimenter. A rigged drawing was held to divide “teacher” and “learner” roles between the subject and the accomplice; the subject always played the teacher.

The subject was then told that the experiment was to help scientists learn more about the effects of punishment on memory and learning. The experimenter instructed the subject to read a list of words to the learner. Then, the subject would test the learner on what words went together. For each wrong answer, the subject pushed down a lever on a shock generator that supposedly delivered a shock to the learner. The voltage levels went from 15 to 450 volts and progressed upwards in 15-volt increments. The smallest voltage was labeled “Slight Shock” and the highest voltage “Danger: Severe Shock.” The experimenter instructed the subject to progressively increase the voltage one increment for each wrong answer.

Milgram and his team conducted at least 18 separate experiments and varied the experimental conditions in each of them. The first four experiments are worth noting as they are the most cited. In the *Remote Feedback* condition, the experimenter attaches an electrode to the victim’s arm and straps him into a chair in an adjacent room. The subject has no verbal contact with the victim. At the shock level of 300 volts, however, the victim pounds on the wall in protest and is no longer heard from after 315 volts. The *Voice Feedback* condition is like the first except the victim protests verbally during the experiment and can be heard through the walls of the laboratory. To maintain



consistency, the victim's protests were played off of a tape player and different protests were played at different shocks. At 75 volts the victim grunted. At 150 volts he demanded that they let him out. At 180 volts he insisted that he cannot stand the pain, and at 300 volts he refuses to answer anymore of the questions and says that they must let him out. After 345 volts, the victim is not heard from again.

The third and fourth experimental conditions are similar to the Voice Feedback condition, but they move the victim even closer to the subject. In the *Proximity* condition, the victim and the subject are seated only 1 ½ feet apart in the same room. The victim is both visible and his protests are audible. The *Touch-Proximity* condition is identical to the third condition except that the victim receives a shock only when he places his own hand on a shock plate. At 150 volts the victim demands to be freed and refuses to place his hand on the shock plate. The experimenter then tells the subject to hold the victim's hand on the shock plate which involves physical contact between the subject and victim.

Not surprisingly, the change in proximity increased the salience of the victim in the eyes of the subject. The number of *obedient subjects*, defined as those who continued to shock until told to stop, declined the more salient the victim's pain became, and conversely the number of *defiant subjects* increased. Out of 40 adults studied in each condition, "34 percent of the subjects defied the experimenter in the Remote condition, 37.5 percent in Voice Feedback, 60 percent in Proximity, and 70 percent in Touch-

Proximity.”<sup>20</sup> Milgram suggests that the increased proximity and salience of the victim’s presence, two situational factors, could be a factor in the increase in defiance.

In additions to the settings, a consistent situational variable across all four experiments is the experimenter who insists that the subject continue the experiment no matter how much the learner protests. Whenever a subject argues that the experiment should stop, the experimenter demands that the subject continue. He uses phrases such as, “The experiment requires that you continue, teacher,” “You have no other choice, you must go on,” and “The shocks are painful but not dangerous.” Using these and similar phrases, the experimenter pushes the subject to continue asking questions and applying punishments regardless of the victim’s screams and frantic requests to be let out.

Two other versions of the experiment examined the effect groups have on behavior. Group effects are important to acknowledge because situationists point to them as evidence against robust character traits. I note them because an ideal person of integrity would be fairly immune to group effects when pressured to compromise their ID-commitments to moral values. To test the effect others have on subjects, Milgram designed two Voice Feedback experiments where the subject was part of a group. In one experiment, the subject was paired with two confederates. Confederate 1 would read the question, confederate 2 told the subject if the answer was right or wrong, and the subject was required to administer any shocks. When the learner gives his first strong protest (150 volts), confederate 1 refuses to continue in the face of the experimenter’s pressure and sits away from the shock generator. The subject then reads the question and applies

---

<sup>20</sup> Ibid., 141.

shocks. At 210 volts the learner says he does not want to continue which results in confederate 2 refusing to continue and disobeying the experimenter. In this experiment 90% of the subjects defied the experimenter.<sup>21</sup> In another experiment in which the subject only read the questions and a confederate applied the shocks, only 8% of the subjects defied the experimenter.

Given the wide variety of behavior seemingly determined by situational variables, Doris claims that, “The variation in obedience across experimental conditions – from near negligible to near total – is powerful evidence that situational variation can swamp individual differences.”<sup>22</sup> He does leave the door open for some individual subject differences by allowing that situational manipulations may not produce uniform behavior. Ross and Nisbett disagree and suggest that each subject may experience different situational pressures or construe the situation differently which would produce non-uniform behavior.<sup>23</sup> Doris does not place much confidence in this explanation because it seems that most subjects experienced a relevantly similar situation given the careful construction of the experiment.<sup>24</sup> Regardless of their views on the personal construal of situations, the situationist points to the Seminarian and Milgram results as evidence against claims that robust character traits produce consistent behavior across a variety of situations.

---

<sup>21</sup> Milgram, 1974, 116-122.

<sup>22</sup> Doris, 2002, 46.

<sup>23</sup> Ross and Nisbett, 1991, 11-13.

<sup>24</sup> Doris, 2002, 47.

#### 1.4 - The Problem of Consistent Behavior

A problem still exists. The situationist cannot rest easy because even in the highest incidents of obedience in the Milgram experiments, some people *did* resist the situational factors and little evidence exists that supports the hypothesis that the subjects construed the situations differently which resulted in different behavior. Even if personal construal was a key factor, the helpful or defiant subjects' mental constructs of beliefs, commitments, and goals most likely formed the basis for this construal. Overall, 16 (40%) of the seminary students offered some form of help and 24 (60%) did not. It is not known if those who helped construed the situation the same way, but it would be surprising if situational factors or personal construal alone could exclusively account for why 40% of them helped.

Situationists must account for why some people behave consistently across a variety of situations, particularly in situations in which many other people act "out of character." Because they reject robust character traits, Ross and Nisbett acknowledge that non-trait factors may explain why people behave consistently. To explain consistent behavior they use an "idiographic" approach which seeks to learn different things about different people in order to "appreciate the distinctiveness and coherence in their behavior."<sup>25</sup> Instead of trying to discover character traits such as compassion or honesty that can be attributed across different people, this approach looks for what makes an individual's behavior coherent with who she is.

---

<sup>25</sup> Ross and Nisbett, 1991, 163.

Ross and Nisbett consider five personal factors that may explain consistent behavior and personal coherence: goals and preferences, competencies and capacities, subjective representations of situations, attributional styles and perceptions of personal efficacy, and conceptions of the self.<sup>26</sup> These factors are drawn from social cognitive research which examines the relationship between the individual and her social environments. In line with my discussion of moral identity and the self in Chapter 3, I am also drawn to the social cognitive findings regarding consistent behavior. My eventual response to situationist doubts regarding stable mental structures also focuses on the power of self-conceptions and moral identity, which to varying degrees include the five factors noted by Ross and Nisbett.

The moral identity account that I develop and use to reinterpret the experiments above does not directly attack situationist claims regarding the influence of situational factors. I do, however, argue that stable mental constructs exist and can produce consistent behavior across diverse situations. In section 2, I support the following two claims. First, I claim and provide evidence that some people have a moral identity and that it enables them to better regulate their behaviors across diverse situations. Second, I claim and provide evidence that some subjects in the experiments above appear to demonstrate behavioral consistency in line with their moral identities. I concede the situationist point that robust character traits, as narrowly defined by the situationists, do not explain consistent behavior in these cases. However, I argue that other mental

---

<sup>26</sup> Ross and Nisbett, 1991, 162-168.

constructs are fairly situation resistant.<sup>27</sup> I provide the evidential support for the first claim in section 2 and support the second claim in section 3.

## **Section 2: Moral Identity and Consistency**

Psychologist Augusto Blasi initiated the call to research moral identity when he asserted that moral reasoning and self-control may not fully explain nor motivate moral behavior. He argued that a person's desire for a consistent moral identity motivates some people to behave morally.<sup>28</sup> I argued in Chapter 3 that a person of substantive integrity has a self that can be described as a process that is motivated to maintain a consistent self-conception constituted by ID-commitments. I also suggested that the more she values her ID-commitments and self-consistency, the more she will behave in accordance with those ID-commitments. The social-cognitive framework behind these claims posits at least three mental structures that are important for understanding identity and behavior. The first mental structure is *self-concepts or identities*. Researchers explain that an overall self-conception is composed of multiple self-concepts or identities that are formed by ID-commitments to values in different domains. Different identities may include mother, husband, professor, achiever, Libertarian, Hispanic, and sports-lover. Each of these identities have different commitments and possibly ID-commitments.

---

<sup>27</sup> Robert Merrihew Adams has argued that factors such as goals and commitments should be considered character traits. I do not take up his argument, but I agree that the situationist conception of character traits as only referring to behavioral dispositions is too narrow, though understandably so if their target is personality psychology. See Chapter 8 of Robert Merrihew Adams, *A Theory of Virtue* (New York: Oxford University Press, 2006).

<sup>28</sup> Augusto Blasi, "Moral Cognition and Moral Action: A Theoretical Perspective." *Developmental Review*, 3. (1983).

The second structure is *moral scripts or schemas*. Having a moral identity constituted by ID-commitments also means that individuals may perceive situations through different moral scripts or schemas. By developing a moral schema, a person interprets situations primarily from a moral perspective. For example, imagine that two real estate agents encounter a person who wants to buy a property she cannot afford. Agent 1 has a strong achievement identity and, lacking a moral schema, perceives the buyer as a means to making his monthly sales quota. Agent 2, on the other hand, has a strong moral identity and perceives the buyer as a person who needs honest advice about the risks of borrowing too much money. What each sales person perceives is mitigated by the schema provided by his or her identity.

But which identity is used if the agents have several? The agents use the identities that are most accessible to their *working self concept*, the third mental structure. Social cognitivists posit that the self has more access to some identities than others as it makes coherent and self-consistent decisions throughout the day. We could speculate that Agent 1 has continual access to his achievement identity while Agent 2 has continual access to her moral identity. For people like Agent 2, their ID-commitments to moral values are highly accessible to their working self concept which results in interpreting situations using a moral schema and acting from their ID-commitments.

Social cognitivists and personality psychologists have conducted numerous experiments to determine if this conceptual framework exists and if so, how moral identity influences behavior. Recent research programs have revealed individual cognitive structures that they have called *moral identity*. These findings demonstrate that

some people have internalized moral principles to a great degree and that their working self-concept (i.e. identities accessible to memory at a particular time) can easily access them across situations.<sup>29</sup> Researchers have also found that these people typically demonstrate more pro-social behaviors and social responsibility than those with a less accessible moral identity. Below I describe the conceptual framework for moral identity in more detail and the evidence for how it influences moral decisions and conduct across situations.

### **2.1 - Internalizing and Accessing Moral Principles**

In this sub-section I consider conceptual frameworks for moral identity and how experiments based on these frameworks suggest that some people have *high moral identity centrality* or HMID and some have *low moral identity centrality* or LMID. Moral identity centrality refers to how easily a person can access her moral identity when making decisions in each situation. It is hypothesized that a person with HMID will act according to moral values more consistently across situations than a person with LMID.

Moral identity researchers have different conceptualizations of how a person develops a moral identity and how it becomes central to her self-conception. Barry R. Schlenker and his colleagues conceptualize that most people have an ethical ideology which is, “an integrated system of beliefs, values, standards and self-assessments that define an individual’s orientation toward matters of right and wrong.” On Schlenker’s

---

<sup>29</sup> Karl Aquino, Americus Reed II, Stefan Thau, and Dan Freeman. “A Grotesque and Dark Beauty: How Moral Identity and Mechanisms of Moral Disengagement Influence Cognitive and Emotional Reactions to War.” *Journal of Experimental Social Psychology* 43, (2007).



account, an ideology can create moral schema for interpreting events and a moral identity that “describes one’s ethical character and provides a basis for self-regulation.”<sup>30</sup>

Schlenker et. al. posit at least two ethical ideologies: principled ideologies and expedient ideologies. These ideologies differ by the strength of a person’s commitment to moral principles. A person with a *principled ideology* believes moral principles exist, that they should guide her behavior, that they apply across situations, and that one may need to actualize them regardless of personal consequences. She defines her identity by her steadfast commitment to principles, which becomes her substantive integrity. A person with an *expedient ideology* considers moral principles as helpful and flexible guides. She does not want to be too rigid and miss opportunities that may further her advantage and she rationalizes deviations from principles when necessary and does not consider substantive integrity as part of her identity. One can imagine these ideologies on opposite ends of a continuum that ranges from a strong commitment to moral principles to a pragmatic commitment to moral principles.

A person of integrity on Schlenker et. al.’s model *commits* to the principles that form her principled ideology and this changes her identity. Much like the ID-commitment model presented in Chapter 1, Schlenker notes that, “Commitment crystallizes and strengthens corresponding attitudes, making them more accessible in memory, more resistant to subsequent change, and more likely to guide future behavior.”<sup>31</sup>

Commitments can be to goals, principles, people, organizations, or a set of ideas. For

---

<sup>30</sup> Barry R. Schlenker, Marisa L. Miller, and Ryan M. Johnson, “Moral Identity, Integrity, and Personal Responsibility” in *Personality, Identity, and Character: Explorations in Moral Psychology*. Eds. Darcia Narvaez and Daniel K. Lapsley. (New York: Cambridge University Press, 2009), 316.

<sup>31</sup> Schlenker et. al., 2009, 318.

example, a person who commits to be an honest person can access this standard easily from her memory, is more likely to endure difficulty rather than act dishonestly, and will behave honestly across a variety of situations. Commitments also make alternative actions less appealing and unjustifiable while limiting the behavior options available.

The upshot of this conceptualization is that when a self commits to moral principles, she is motivated to act on those principles, a conclusion I defended in Chapter 3. In Schlenker et. al.'s words, "When there is a strong linkage between self and principles, represented by feelings of duty or personal obligation to follow the principles, the principles have been both internalized and appropriated as part of one's identity."<sup>32</sup> The hypothesis put forward is that consistent moral behavior results from having a moral identity based on ID-commitments to moral principles (i.e. moral values).

To test these conceptual frameworks, researchers have designed measurement scales and experiments to determine if some people have HMID that influence their evaluations, decisions, and behaviors. Schlenker developed an Integrity Scale to assess the strengths of principled or expedient ideologies.<sup>33</sup> The items evaluate three characteristics of a person's ideology and commitments: the inherent value of principled conduct, the strength of their commitment to principles, and their unwillingness to rationalize unprincipled behavior. Instrument items included, "Integrity is more important than financial gain;" "The true test of character is a willingness to stand by

---

<sup>32</sup> Schlenker et. al., 2009, 319.

<sup>33</sup> Barry R. Schlenker, "Integrity and Character: Implications of Principled and Expedient Ethical Ideologies." *Journal of Social and Clinical Psychology*, 27, No. 10. (December 2008).

one's principles, no matter what one has to pay;" and "Some actions are wrong no matter what the consequences or justification."<sup>34</sup>

Karl Aquino and Americus Reed developed a similar measure of the centrality of moral identity that relies heavily on social cognitive frameworks. They base their notion of moral identity on a "self-schema that is organized around a set of moral trait associations."<sup>35</sup> They conceptualize moral traits that are linked in a person's memory to an entire network of similar moral traits. The activation of a subset of moral traits will then activate others within the network of moral traits. Moral traits presented in their instrument include being caring, compassionate, fair, friendly, hardworking, generous, helpful, and honest. Their instrument then questions the respondent on the degree these traits are rooted in her core self-concept (i.e. internalization dimension) and the extent they display their morality outwardly in their actions (i.e. symbolization dimension).<sup>36</sup> If the person has a strong internalization dimension, which I have called a serious concern or ID-commitment to acting on her moral identity, then her moral self-schema is highly central to her when she makes decisions across situations.<sup>37</sup>

---

<sup>34</sup> Schlenker et. al., 2009, 322-323.

<sup>35</sup> Karl Aquino and Dan Freeman, "Moral Identity in Business Situations: A Social Cognitive Framework for Understanding Moral Functioning," in *Personality, Identity, and Character: Explorations in Moral Psychology*. Eds. Darcia Narvaez and Daniel K. Lapsley. (New York: Cambridge University Press, 2009): 377.

<sup>36</sup> For a full description of their instrument, see Karl Aquino and A. Reed, "The Self-Importance of Moral Identity." *Journal of Personality and Social Psychology*, 83, No. 6. (2002).

<sup>37</sup> Aquino and Reed use the term *self-importance* when referring to how easily accessible a person's moral identity is to the self across situations. Another term used to describe a moral identity centrality is *chronic accessibility*. To reduce confusion, I use the acronyms HMID and LMID to describe the degree of a person's moral identity centrality across situations.

Both Schlenker's and Aquino and Reed's instruments have good internal consistency and test-retest reliability, and going forward I assume these instruments accurately identify high and low moral identity centrality. Even though these two conceptualizations approach moral identity from different directions and with different instruments, a strong correlation has been found between what they measure. In a study that correlated Aquino and Reed's moral identity measures with Schlenker's ethical ideology measure, researchers concluded that the internalization dimension of moral identity (i.e. HMID) was more strongly related to the commitment to a principled ideology than the symbolization dimension. Researchers stated that, "It makes sense given these distinctions that internalization would be more closely related to another internalized construct—the commitment to a principled ideology—than symbolization."<sup>38</sup> The study suggests that that both measurements track a person's serious concern and ID-commitment to a moral identity.

After determining that the moral identity construct exists, the researchers must determine if it is easily accessible to the person's working self concept. This step is critical because if situational factors consistently make a person's moral identity less accessible for decision making, then moral identity does not answer situationist skepticism about stable mental constructs. In other words, when the seminarian sees the victim slumped in the doorway and must decide what to do, can he easily access his moral identity or do the situational factors make his moral identity less accessible? Determining the resiliency of moral identity accessibility is a critical step because if

---

<sup>38</sup> Brent McFerran, Karl Aquino, and Michelle Duffy. "How Personality and Moral Identity Relate to Individuals' Ethical Ideology." *Business Ethics Quarterly* 20, 1. (January 2010).

situational factors consistently make a person's moral identity less accessible for decision making, then moral identity does not respond to the situationist challenge to integrity.

The ease of access to one's moral identity and its commitments is explained as high or low centrality (HMID/LMID). Social-cognitive theory claims that moral functioning involves cognition, memory, identity, and how information is processed using the knowledge that is accessible to the self at a particular time. When a person acts, the theory predicts that a moral identity that is readily accessible and central will influence behavior more than those identities that are less accessible.<sup>39</sup> To determine if the conceptualizations above are accurate, several studies have tested whether or not people with HMID have reliable accessibility to their moral identities.

In one experiment, Schlenker et. al. found that those with HMID preferred characters in a case who behaved ethically when trying to advance their careers. "More principled people strongly preferred characters who made ethical career decisions over those who made unethical ones, and this preference was largely unaffected by whether the character was successful or unsuccessful in the career moves."<sup>40</sup> In another experiment, they found that people with HMID tended to behave consistently with their principles in a decision-making task even when others tried to convince them to make a decision that most would consider unethical. People with LMID were more quickly and easily convinced to make the unethical decision.<sup>41</sup>

---

<sup>39</sup> A full description of centrality and accessibility can be found in Aquino and Freeman, "Moral Identity in Business".

<sup>40</sup> Schlenker et. al., 2009, 325.

<sup>41</sup> Schlenker et. al., 2009, 328.

These initial results do not defeat situationist claims. Instead, social-cognitive psychology integrates situational factors in to their research. Unlike the situationists, they determine how situational influences can activate or deactivate a moral identity. Like the situationists, researchers acknowledge that situational factors influence decisions and they conceptualize the interaction as a person's moral identity becoming more or less accessible because of these factors. Several experiments have tested the power of situational influences to moderate moral identity accessibility and I briefly review three of these experiments below. The purpose of this review is to demonstrate that situational factors influence people differently according to their moral identity accessibility.

In the first experiment, Aquino et. al. measured if a moral priming task affected the intention to behave in a pro-social manner. Subjects were tested on their moral identity centrality at least a day before the experiment. On the day of the experiment, some of the subjects were morally primed by reading and recalling the Ten Commandments before the experiment while the control group worked with general knowledge items. The subjects were then asked to consider donating to a pro-social cause. The researchers found that moral priming increased moral identity accessibility for those with LMID and had a minimal effect on those with HMID.<sup>42</sup> The results suggest that people with HMID can access their moral identity when deciding to donate to a cause and priming them with a moral stimulus does not significantly increase their intention to donate.

---

<sup>42</sup> Karl Aquino, Dan Freeman, Americus Reed II, Vivien K. G. Lim, and Will Felps. "Testing a Social-Cognitive Model of Moral Behavior: The Interactive Influence of Situations and Moral Identity Centrality." *Journal of Personality and Social Psychology*, 97, No. 1. (2009): 129.

In the second experiment, Aquino et. al. found that situational factors can *reduce* the accessibility of moral identity in people who have HMID. At least two days before the experiment, subjects completed a questionnaire that included measures of moral identity centrality. On the day of the experiment, subjects were told that they would role play an employer who must negotiate the lowest possible salary with an employee while not revealing information about terminating the job in six months. Some of the subjects were told they would receive \$100 if they negotiated the lowest salary compared to other subjects. The control group was told that regardless of the outcome, they would be entered into a random drawing for \$100. In a nutshell, the experiment offered some subjects an incentive to negotiate the lowest salary and others no incentive. All subjects were monitored to see if they lied, concealed the truth, did not answer questions, or told the truth.

An analysis of the results suggests that, “participants high in moral identity centrality were more likely to lie in the performance incentive condition [\$100 contest] compared with the random condition . . . . However, the incentive manipulation had no effect on lying for participants who were low in moral identity centrality.”<sup>43</sup> In other words, situational factors like incentives can make one’s HMID less accessible. Also, subjects with LMID were unaffected by situational factors like incentives when it came to lying. The main finding is that moral identity centrality *moderates* how a person acts given situational factors. It is also worth considering the actual data for individuals since our focus is on individual rather than broad character-trait differences. Subjects with

---

<sup>43</sup> Aquino et. al., 2009, 134.

HMID who had the performance incentive told the truth more than any other group. Also, those with HMID in the random incentive group lied the least of all the groups.

Given the moderating role of moral identity centrality, Aquino et. al. wanted to see if subjects with HMID will increase their cooperation over time if primed with moral situational cues. In the third experiment, subjects played a game in which they must sacrifice personal gain and cooperate in order to advance the collective good. Some subjects were “morally primed” by completed a handwriting task of reading, copying, and using words with moral meanings in a story. Subjects were then placed in groups and asked to make “investment decisions” by deciding whether or not to allocate points to either a joint account or their personal account. The experimenters structured the exercise so that individual self-interest conflicted with the social good. They also manipulated each trial so that it appeared that most other participants were defecting for personal gain.

The experiment results revealed that across 20 allocation decisions for each subject, moral priming correctly predicted increased or level cooperation for those with HMID and predicted no increase in cooperation for those with LMID. Aquino et. al. concluded that, “These results support our hypothesis that priming the moral self-schema would motivate participants to sustain cooperation overtime despite the defection of others, but only if they were high rather than low in moral identity centrality.”<sup>44</sup> They also found that non-primed subjects with HMID began cooperating but eventually defected to the level of those with LMID.

---

<sup>44</sup> Aquino et al., 2009, 137.



The findings in these experiments support the situationist claims that situational factors influence behavior. But these situational factors may operate by making moral identity more or less accessible which can result in different moral outcomes. These experiments reveal that moral identity accessibility moderates the effectiveness of these situational factors which suggests that situational factors alone do not account for how most people make decisions. The findings clearly reveal that people with HMID can easily access moral standards when making decisions and that their moral identity moderates the situational factors that influence their behavior.

In addition to moral identity accessibility experiments, other research suggests that people who have HMID have a higher likelihood of producing pro-social behavior and demonstrate greater social responsibility than those with LMID. These findings are significant for establishing the consistent, real-world behavior that moral identity accessibility produces.

## **2.2 - Pro-Social Behavior and Social Responsibility**

A consistent finding in moral identity research is that a positive relationship exists between HMID and moral behavior. In the moral identity literature, researchers have found a positive relationship between moral identity and three behaviors: 1) moral behavior as helping others, 2) helping out-group members more than in-group members, and 3) taking responsibility for one's own actions.

Aquino and Reed found a positive relationship between HMID and helping behavior in two studies. In one they found that HMID was associated with an increased probability that people volunteered to help in the community (e.g. homeless shelter,

organize food drive, mentor troubled youth, or visit patients in a nursing home) during the last two years.<sup>45</sup> In the second study they found that HMID was positively associated with the likelihood that participants donated to a food drive and also the amount of food they decided to give.<sup>46</sup> These findings match research on moral exemplars which found that they are distinguished in part by their commitment to helping.<sup>47</sup> Schlenker et. al. also found that HMID was negatively related to antisocial behavior including self-reports of “telling self-serving lies, cheating in high school and college, stealing, breaking promises, infidelity, and alcohol and drug use.” He also reported that, “The relationships remained equally strong even after controlling for social desirability bias.”<sup>48</sup>

In addition to pro-social behavior, Reed and Aquino cite evidence for a positive relationship between HMID and moral behavior to out-groups. It has been shown that different standards of morality can sometimes lead members of an in-group to act intolerantly toward members of out-groups.<sup>49</sup> Using American subjects, Reed and Aquino compared giving to the New York Police and Fire Widows and Children’s Benefit fund (in-group) and the United Nations Children’s Fund (UNICEF) Emergency Effort for Afghan Children and Families (out-group). Based on this study Aquino and

---

<sup>45</sup> Aquino and Reed, 2002, 1433. Aquino and Reed consider volunteering a higher measure of moral commitment than simply giving money to similar causes which makes their finding more significant.

<sup>46</sup> Americus Reed II and Karl Aquino, “Moral Identity and the Expanding Circles of Moral Regard Toward Out-groups.” *Journal of Personality and Social Psychology*, 84, No. 6. (2003).

<sup>47</sup> See Lawrence J. Walker and Jeremy A. Frimer, “Moral Personality Exemplified,” in *Personality, Identity, and Character: Explorations in Moral Psychology*. Eds. Darcia Narvaez and Daniel K. Lapsley. (New York: Cambridge University Press, 2009).

<sup>48</sup> Schlenker et. al., 2009, 331.

<sup>49</sup> T. R. Cohen, R. M. Montoya, and C. A. Inkso. “Group Morality and Intergroup Relations: Cross Cultural and Experimental Evidence,” *Personality and Social Psychology Bulletin*, 32. (2006).

Freeman conclude that, “Consistent with the notion that moral identity is associated with an expanded ‘circle of moral regard’ toward out-groups, greater self-importance of moral identity was associated with an increase in the amount donated to UNICEF.”<sup>50</sup> In another study, Aquino and his colleagues found that concern for the needs of socially distant others is uncharacteristic of those with LMID.<sup>51</sup> Schlenker et. al. also concluded that HMID is related to greater benevolence and overall helping behaviors regardless of in-group or out-group status.<sup>52</sup>

A third positive relationship exists between HMID and taking responsibility for one’s actions. This relationship relates to having a wider circle of moral concern because being concerned with more people can result in a person becoming more responsible for her actions and inactions towards them. Two measures related to moral behavior and responsibility are moral disengagement and moral justification. Moral disengagement describes the tendency of people to distance themselves from immoral acts by rationalizing. It has been found that those who rationalize and psychologically disengage from antisocial behavior are more likely to act in antisocial ways. Moral justification describes a willingness to justify or excuse antisocial behavior so that they are less condemned by themselves and others. Schlenker et. al. present findings that HMID is related to lower scores on moral disengagement (i.e. tendencies to distance self and

---

<sup>50</sup> Aquino and Freeman, 2009, 386.

<sup>51</sup> Aquino, et. al., 2007.

<sup>52</sup> Schlenker et. al., 2009, 330.

rationalize antisocial behavior) and moral justification (i.e. ease of justifying illegal or immoral behaviors).<sup>53</sup>

Schlenker et. al. also claim that HMID increases a person's felt responsibility for her actions. Schlenker et. al. developed a triangle model of responsibility to determine how people become engaged by tasks and then seek to disengage from them using excuses and justifications. A person is more personally responsible for outcomes when she meets three conditions. First, she knows clear *prescriptions* that apply to a situation (e.g. clear principles). Second, she is *bound and obligated* to follow those prescriptions (e.g. commitment to principles). Third, she appears to have *control* over the outcomes. A person's beliefs about these three conditions directly relates to her engagement and performance.

However, when a person fails to perform in these areas, she may seek to reduce her responsibility by rationalizing away one or more of the conditions. She may argue that the standards were not clear or conflicted, that she wasn't really obligated to perform, and/or that outside events are the real cause of the outcome. Another way to avoid responsibility is to justify her performance by changing the assessment of prescriptions (e.g. "Following orders is important as well."), the outcome of event (e.g. "It wasn't all that bad.") and/or her identity (e.g. "You can't expect me to meet these standards.").<sup>54</sup>

Schlenker et. al. argue indirectly that a person with HMID has stronger linkages to principles, obligations, and outcome control than a person with LMID. First, a person

---

<sup>53</sup> Schlenker et. al., 2009, 330.

<sup>54</sup> For a full account see Barry R. Schlenker, T. W. Britt, J. W. Pennington, R. Murphy, and K. J. Doherty. "The Triangle Model of Responsibility." *Psychological Review*, 101, (1994).

with HMID has a stronger commitment to principles. Second, because she expresses a clearer purpose in life and has less alienation, Schlenker et. al. suggest that the moral prescriptions are clearer to her. Finally, HMID subjects express greater internal control which suggests they have higher feelings of personal control. Schlenker realizes that more research is needed between responsibility and HMID, but he suggests that the link is more than conceptual at this point.<sup>55</sup> I revisit the connection between HMID and responsibility in Section 3.

### 2.3 - Summary of Moral Identity Research

The experimental findings presented above are summarized in the following table.

	<b>Moral Prime and Intent to Donate</b>	<b>Performance Incentive (PI) vs. Random Incentive (RI) in Negotiation</b>	<b>Moral Prime and Cooperation Over Time</b>	<b>Pro-social Behavior</b>	<b>Responsibility for Self and Others</b>
<b>HMID</b>	Minimal effect	PI – Increased lying over RI; same percentage told the truth. RI – Lying percentage lowest of all groups.	Moral prime sustained cooperation over time. Non-primed group initially cooperated then defected over time.	Positive relationship	Greater circle of moral concern and pro-social behavior to out-groups.
<b>LMID</b>	Strong effect	PI – No effect. RI – Highest percentage lying of random groups.	Moral prime influenced initial cooperation, then defected over time. Non-primed defected consistently over time.	No reported relationship	Smaller circle of moral consideration than HMID.

I draw three conclusions about people with HMID from these findings.

---

<sup>55</sup> Schlenker et. al., 2009, 335.

1. *In daily life, people with HMID engage in more pro-social behaviors, fewer anti-social behaviors, and have a wider circle of moral concern than people with LMID.* This conclusion is significant for the moral identity account of integrity. Substantive integrity requires an ID-commitment to moral values as evidenced by few if any anti-social behaviors and a wide circle of moral concern that does not arbitrarily limit who is morally considerable. While substantive integrity does not require a person to consistently perform pro-social and altruistic behaviors, evidence of consistent pro-social behavior does reveal some level of commitment to moral values compared to those with LMID. Also, while the evidence does *not* directly support an expectation of consistent moral behavior across diverse situations, it does support the generalization that some people with HMID are more likely to have ID-commitments to moral values and to what is objectively good (e.g. help *any* human in need).
2. *HMID and its accessibility moderate how people make judgments, decisions, and behave.* The evidence reveals that people with HMID tend to have their moral identity constantly accessible. A working self concept's accessibility to one's moral identity informs the self about which moral values should be used when making decisions about behavior. According to the moral identity model, moral identity centrality and accessibility are important mental structures for being a person of substantive integrity.
3. *Situational factors and deliberate primes can affect the accessibility of a person's moral identity.* The evidence is clear that situational factors that encourage self-

interest and non-cooperation affect some people with HMID. One explanation for this effect is that situational factors make a person's HMID less accessible and possibly make other identities more accessible. For example, a person's occupational identity may be more salient than her moral identity in a competitive sales situation. The evidence is also clear that priming some HMID people with moral stimuli tends to make their moral identity accessible over time while it has a minimal effect on people with LMID. In the next section I argue that people can and do deliberately "prime" and re-enforce their moral identity accessibility with daily reminders and devotions.

In the next section I use these three conclusions to reinterpret the situationist account of the Seminarian and Milgram experiments.

### **Section 3: Revisiting the Experiments**

In this section I consider alternative explanations of some subject behavior in the Seminarian and Milgram experiments. My purpose is to support the claim that consistent moral behavior has some basis in moral identity and that some helpful and defiant subjects had access to their moral identities. If I successfully make my case, I conclude that the situationist conclusions about the experiments are incomplete. Throughout my review, I argue that moral identity accessibility is evident in some subjects and that objective situational factors and the subjective construal of situations do not accurately explain their actions.

### 3.1 - Revisiting the Seminararians

As a reminder, Darley and Batson tested three variables to determine which ones most impacted helping behavior. These variables were the dispositional variable of different types of religiosity, the content of one's thinking, and the amount of hurry during a journey.<sup>56</sup> All subjects completed personality tests that measured different types of religiosity. Afterward, some read the parable of the Good Samaritan and others read a discussion of ministering in the professional clergy. After being informed that they must give a speech in another building, subjects were put in a *high-hurry condition*, *intermediate-hurry condition*, or *low-hurry condition*.<sup>57</sup> The results analysis pointed to situational factors influencing helping behavior more than the type of passage the subjects read and religious dispositions. Listed by situational variables, 63% of those in low-hurry offered help, 45% of those in intermediate hurry offered help, and 10% of those in high-hurry offered help. The conclusion was that the hurry condition was significantly related to helping behavior while the passage and the measures of religiosity were not statistically significance.

It appears that objective situational factors significantly influenced behavior more than character traits. But that is not the whole explanation by any means. Moral identity research points to at least three questions that need to be answered: 1) What happened to the subjects' moral identity accessibility during the experiment? 2) What situational and

---

<sup>56</sup> Darley and Batson, 1973.

<sup>57</sup> Darley and Batson, 1973, 104.



internal factors produced the results? and 3) What role if any did religious identity play in helping behavior?

*1) What happened to the subjects' moral identity during the experiment?* Even though we cannot know which of the seminarian's had HMID or LMID, we can interpret the results using moral identity constructs. On a moral identity account, it could be argued that situational factors made the moral identity of some HMID subjects less accessible by activating another identity or creating confusion. The situational factors would have little effect on those with LMID. Of the 10% of the high-hurry subjects who did help, their behavior could be interpreted as their being people with HMID who maintained a constantly accessible moral identity. While these social-cognitive interpretations are speculative, moral identity findings support them as an alternative explanation of how some subjects reacted during the experiment. I take this analysis further in question 3 below.

*2) What situational and internal factors were responsible for the results?* If we assume for the moment that objective situational factors had the most influence on helping (or non-helping) behavior, we can still ask if the hurry condition alone made the difference. Contrary to this view, some have argued that the situational hurry factor does not fully explain the difference in helping behavior. Instead, the subjects in the high and intermediate hurry condition may have had an ID-commitment toward promptness or they may have felt conflict between obligations to the victim and obligations to the people waiting for them. Darley and Batson report that some subjects who passed the victim appeared aroused and anxious upon arriving to the second location. They

interpreted this anxiousness as a conflict between helping the victim and helping the experimenter; both commitments to moral values that they could not simultaneously actualize.<sup>58</sup> From the moral identity perspective, these anxious subjects most likely have access to their moral identity but are unsure about which behavior would fulfill their moral commitments. While situational factors influence the behavior, the subject's moral identity attempts to moderate the overall effects on their behavior.

In support of the interpretation that moral identity moderated behaviors, consider a similar experiment that Batson et al. conducted five years later.<sup>59</sup> In this experiment they wanted to determine which factor most accounts for reducing helping behavior: a hurry condition or conflict over whom to help. In the experiment, subjects were told that a researcher needed their data in another building. Some were told that the data was important for the researcher to complete the project while others were told it was not important. Half were then told that they were late and must hurry and the other half that they had plenty of time. All of the subjects had to go down a stairway where a confederate posed as a victim who was coughing and groaning.

The experiment revealed that 8 of 10 subjects in a no-hurry condition helped the victim when a researcher was *not counting* on the subject. Surprisingly, 7 of 10 subjects in a *hurry condition* helped the victim when a researcher was *not counting* on the subject. This suggests that the hurry situational factor does not completely account for a lack of helping behavior. When the experimenters added a researcher who *was counting on*

---

<sup>58</sup> Ibid., 1973, 108.

<sup>59</sup> C. Daniel Batson, P.J. Cochran, M.F. Biederman, J. L. Blosser, M. J. Ryan, and B. Vogt. "Failure to Help When in a Hurry: Callousness or Conflict", *Personality and Social Psychology Bulletin* 4 No. 1, 1978.

subjects to deliver the data, 5 out of 10 in a no-hurry condition helped the victim and only 1 of 10 subjects in a hurry condition helped. The results suggest that situational factors like hurry conditions do influence helping behavior, but factors such as conflicting moral commitments may be moderating helping behavior behind the scenes.

The results of both experiments reveal the power of situational factors, but they also allow for interpretations that include moral identity accessibility as a moderating force. I suggest that the anxious and confused non-helping subjects in the Seminarian experiment showed signs of anxiety because their moral identity was accessible but gave them little guidance in choosing between two mutually exclusive moral behaviors. When they pass the victim, they are acting both for and against their ID-commitments.

A final interesting data point needs to be made regarding moral identity and the priming effect. Based on the moral identity findings above, a person with HMID who is morally primed should have a more accessible moral identity and would be expected to demonstrate more pro-social behavior across situations. A loose interpretation of the Seminarian results could indicate a correlation between moral priming and helping behavior. Darley and Batson found that no statistically significant relation existed between helping behavior and which passage the person read. The raw data, however, reveal that 53% of those who read the Samaritan story helped the victim while only 29% of those who read the neutral passage helped. I cannot draw any strong conclusions from these numbers, but they do suggest that for some individuals the Samaritan prime may have made their moral identity more accessible during the experiment.

3) *What role did religious identity play in helping behavior?* Darley and Batson tracked a personality/disposition factor that captured different forms of religiosity. Religiosity as a personality measure did not predict helping behavior, but types of religiosity did predict the *kind of help* subjects offered. The religiosity measures that produced significant results were religion as a quest (i.e. “religiosity emerging out of an individual’s search for meaning in his personal and social world”) and doctrinal orthodoxy (i.e. “agreement with classic doctrines of Protestant theology”).<sup>60</sup> At the time of the experiment, the orthodoxy measure was considered a measure of religious identity.<sup>61</sup> Darley and Batson compared these measures to a “degree of help” measure. They created this measure because they found “helping behavior” ranged from subjects asking if the victim was OK and leaving to subjects helping to such a degree that the victim had to insist that they leave before the next subject arrived.

Comparing these two measures surprised Darley and Batson. They found that a high score on the religion as quest measure predicted tentative and incomplete helping behavior while those scoring low on the measure offered more complete or “added” help. These “super helpers” did not leave until the victim insisted that he would be fine and encouraged them to leave. Darley and Batson originally categorized these subjects as “rigid” because their behavior was highly likely “among doctrinal orthodox subjects.” They later revised their interpretation of super helper actions as *different* rather than

---

<sup>60</sup> Darley and Batson, 1973, 102.

<sup>61</sup> The orthodoxy items seek to measure a subject’s certainty in believing traditional doctrines such as, “There is life after death.”, “Jesus was born of a virgin.”, “Jesus is the Divine son of God.”, and “Jesus walked on water.” The level of certainty regarding these items varies greatly depending on denomination. See Charles Y. Glock and Rodney Stark, *Christian Beliefs and Anti-Semitism*. (New York: Harper and Row Publishers, 1966).

*inappropriate*. They describe a person with a specific schema from which they view the situation: "This kind of helper seems quickly to place a particular interpretation on the situation, and the helping response seems to follow naturally from this interpretation."<sup>62</sup>

These subjects wanted to serve the underlying needs of the victim while the high religion-as-quest subjects were more tentative and responsive to victim requests to leave.

A moral identity interpretation of these findings is that super helpers had a moral schema and moral identity that were highly accessible and they acted enthusiastically to help the victim. It also appears that their strong *religious identity* may have increased their moral identity accessibility. It would not be a stretch to suggest that those who are certain about orthodox religious beliefs are fairly clear about the standards by which they live. Meanwhile, those who helped and had a strong religion-as-quest disposition seem to have only minimally accessed their moral identity or at least valued the victim's opinion of the situation differently. My interpretations about the relationship between religious devotion and altruistic behavior are more than speculative.

The relationship between *religious devotion* and helping behavior is well supported by findings in religious psychology and Holocaust rescuer research. One consistent result in religious psychology is that irreligious people (i.e. those who profess they are not religious) and very religious people (i.e. those who attend church regularly and are active members) are less prejudiced and more open to out-group members than moderately religious people. Researchers call this the "curvilinear relation" between religion and prejudice because prejudice does not increase linearly from irreligious

---

<sup>62</sup> Darley and Batson, 1973, 107.

people to very religious people. In fact, the positive correlation between religion and prejudice becomes negative the more "pious" a person is.

A few studies that reflect the body of work in this area will suffice to establish this relationship.<sup>63</sup> Glenn Wilson and Francis Lillie surveyed two groups at the extremes of religiosity, the officer cadets of the Salvation Army and the members of the Young Humanist Association. They found that both groups showed exceptionally low levels of racial prejudice.<sup>64</sup> Another study compared a group of seminarians and nuns who took communion daily with laypersons at the same Catholic university who did not take it daily. The researchers found that the daily communicants, on average, had significantly lower ethnocentrism scores than the lay members.<sup>65</sup> A review of similar studies by Richard Gorsuch and Daniel Aleshire found that the 20 studies were consistent with the curvilinear relation.<sup>66</sup>

Most recently, Pearl M. Oliner confirmed the same finding in her surveys and interviews of European rescuers and non-rescuers of Jews. Rescuers who identified themselves as very religious, moderately religious, and irreligious did not differ much in their stereotypic thinking, but very religious and irreligious non-rescuers were

---

<sup>63</sup> These studies are primarily done with white Christians as that group was of most interest for the researchers in measuring religious devotion and prejudice in the 1960's and 1970's. Similar findings have been found in other countries toward minorities.

<sup>64</sup> The study is quoted in David M. Wulff, *Psychology of Religion: Classic and Contemporary (2<sup>nd</sup> Edition)*. (New York: John Wiley and Sons, 1997), 224-225. The study is Glenn Wilson and Francis Lillie, "Social Attitudes of Salvationists and Humanists." *British Journal of Social and Clinical Psychology*, 11, (1972).

<sup>65</sup>Wulff, 1997, 224. Gregory Shinert and Charles Ford, "The Relation of Ethnocentric Attitudes to Intensity of Religious Practice." *Journal of Educational Sociology*, 32, (1958).

<sup>66</sup> Wulff, 1997, 226. Richard Gorsuch and Danile Aleshire, "Christian Faith and Ethnic Prejudice: A Review and Interpretation of Research." *Journal for the Scientific Study of Religion*, 13, (1974).

*significantly* less oriented toward negative stereotypes of out-group members than the moderately religious.<sup>67</sup> These findings match the moral identity studies that found a positive relationship between HMID and having a wider circle of moral concern.

In summary, a strictly situationist explanation of the Seminarian experiment is not complete. In fact it is quite dismissive of the role mental constructs like moral identity may have played in the experiment. The situational hurry factors did seem to make a significant difference in behavior, but this conclusion does not tell the whole story of the variation in helping behavior. Considering moral identity adds a new perspective on what moderated helping behavior, the anxiousness experienced by some subjects, and the degree of helping behavior. The moral identity interpretation is particularly clear in explaining the degree of helping behavior given the association between moral identity, religious identity, and devotion. Devotion can also be to secular causes and moral codes as demonstrated by the Young Humanists. At least in the case of religious devotion, a fairly consistent association exists between it and having a wider circle of moral concern which is a characteristic of people with HMID.

I am not claiming that a moral identity interpretation of the Seminarian experiment provides a complete account of the helping behavior given during the experiment. However, I do consider the moral identity interpretation a more accurate explanation than a stand-alone situationist account. On my analysis, the helping behavior of individual seminarians cannot be completely explained by consistent situations and

---

<sup>67</sup> Pearl M. Oliner, *Saving the Forsaken: Religious Culture and the Rescue of Jews in Nazi Europe*, (New Haven: Yale University Press, 2004), 32-33.

subjective construals. I suggest that for some seminarians, their behavior came from their HMID that remained accessible regardless of the hurry condition.

### **3.2 - Revisiting the Milgram Experiments**

Situationists claim that the Milgram experiments suggest that situational factors and not robust character traits explain why a high percentage of ordinary people continued to “shock” an innocent person, sometimes to his death. If individual differences such as robust character traits existed, then situational factors trounced them. As Doris bluntly states, “Or is it to be supposed that 39 virtuous subjects and one vicious subject were assigned to the three percent obedient ‘subject chooses shock level’ condition, while 37 vicious subjects and three virtuous subjects were assigned to the 93 percent obedient ‘peer administers shocks’ condition?”<sup>68</sup> Doris’ point is well taken. Situational variations seem to produce different behaviors and it is inappropriate to fully explain subject behavior by attributing virtuous or vicious character traits to subjects.

In reviewing the Milgram results, Doris does observe that different situational manipulations did not produce completely uniform behavior. Individual “dispositions” or some other personal factor must explain why all subjects did not obey or defy the experimenter. Even though he undertakes an exhaustive review of related literature, he does not find any personal factor that explains different individual responses. Gender, age, traditional personality measures, and perception of authority are all considered and discounted as lacking the influence needed to explain different responses. Doris also considers that the defiant subjects may have had character traits that enabled them to

---

<sup>68</sup> Doris, 2002, 46.



stand-up to the experimenter. He eventually downplays this option because the subjects were only observed in one trial and he has little confidence that defiant subjects would act compassionately across other situations or that obedient subjects were in a habit of shocking people.<sup>69</sup>

The Milgram results lead me to conclude that the best explanation for most of the *variation in overall subject behavior* can be explained by a complex combination of situational factors. However, I do not conclude that situational factors come close to explaining individual behavior differences. While situational factors set-up a slippery slope that pressured subjects to obey the experimenter, moral identity findings indicate that mental constructs and alternative identities may have played some part in moderating behavior. While I do not claim that moral identity predicts all of the individual differences in behavior, I do present a case for two claims.

First, the variance between obedient and defiant behaviors can be explained partially by the degree that situational factors make a person's moral identity or other identities more or less accessible. To prove this hypothesis I consider how defiance increases or decreases when a person's moral identity is made more or less accessible during the situation. My second claim is based on the plausibility of the first. If defiant subjects tend to have a more central and accessible moral identity as evidenced by their

---

<sup>69</sup> Doris, 2002, 48. I argue below that Doris' skepticism about defiant subjects is not warranted. When we consider the occupational and social identities of the subjects, we find that in real life some defiant subjects were more responsible for their actions while some obedient subjects were more subordinate to authority figures. Milgram found a similar trend which I address below.

behavior, then they will take more responsibility for their behaviors.<sup>70</sup> From Schlenker et. al.'s work above on responsibility, I take it as a given that a person's locus of responsibility is her identity or self-concept and a person who takes responsibility for her actions knows what moral values she holds and believes that she was in control at the time of her actions and could have acted otherwise.<sup>71</sup>

To prove these two claims, I take as given three findings from moral identity research. First, moral identity accessibility tends to moderate moral behavior. If a person's moral identity is highly accessible, then she will act on it. If her moral identity is non-existent or inaccessible because of another identity, then she will not act on her moral identity. Second, the *type* of situational factors that confront a subject can make her moral identity more or less accessible. In line with the incentive and cooperation experiments above, situational factors can activate other identities and/or make moral identity more or less accessible. Third, people with more access to their moral identities tend to have a wider circle of moral concern which means they tend to take responsibility for the plight of others.

I support my first claim by presenting several Milgram experiments that reveal how the variance between obedient and defiant subjects can be linked to the degree that

---

<sup>70</sup> Much like my interpretation of the Seminarian experiment, I do not have data that divides subjects into groups of HMID and LMID nor do I have information on any primes the subjects experienced before entering the laboratory. In place of this information, I assume, as situationists do, that the majority of subjects hold some moral value or principle that harming innocent people is morally wrong. I focus on those whose behavior reflects an accessible moral identity without claiming that they have HMID or LMID. Note that I am *not* using defiant or obedient behavior to prove the existence of moral identity centrality. Rather, I am considering how moral identity inaccessibility or accessibility as revealed in the experiments and post-interviews appears to have mediated situational factors and produced different behaviors.

<sup>71</sup> I leave aside any discussion of determinism as it is beyond the scope of this dissertation.

situational factors make a person's moral identity more or less accessible. I first consider experiments in which situational factors minimally interfered with subjects' access to moral identity. On a moral identity interpretation, these experiments produced more moral behavior (i.e. lower shock levels and more defiance) because subjects more easily accessed their moral identity. I then present experiments in which situational factors either make the person's moral identity less accessible or activate another identity which results in more obedience. I conclude my review of experiments by reviewing subject comments that reveal the influence of particular identities on their behaviors. I support my second claim by reviewing Milgram's evidence that defiant subjects took more responsibility than obedient subjects for their actions, an indication of moral identity centrality and accessibility.

### *3.2.1 - High Moral Identity Accessibility Situations*

According to moral identity theory and findings, we would expect the majority of subjects in less ambiguous or low pressure situations to easily access their moral identities and act on them. Reasons for this expectation are that other identities will not be activated and/or the self will maintain access to its moral identity. This idea is not new as Milgram also notes that access to moral principles during the high-pressure experiments seems to be reduced.<sup>72</sup> I briefly describe the results of five experiments to

---

<sup>72</sup> Milgram offers his own explanation of moral identity access. He states, "Morality does not disappear, but acquires a radically different focus: the subordinate person feels shame or pride depending on how adequately he has performed the actions called for by the authority." Milgram goes on to speculate that moral concerns are not lost but focus on other moral principles such as loyalty, duty and discipline. My account comes to a similar conclusion, but from an identity accessibility perspective. Milgram, 1974, 8 and 146.

support my claim that moral identity access moderates subject behavior and the interpretation of situational factors.<sup>73</sup>

Experiment 11 is an experiment that sets a baseline for moral identity accessibility. In this experiment, subjects are free to choose the shock level and the experimenter does not require them to increase the voltage for each missed answer. Over the course of thirty “memory” trials, the mean shock level was between 40 and 60 volts. It is worth noting that the learner does not protest about the shocks in all experiments until 75 volts. Of 40 subjects, only 2 went above the “Danger: Severe Shock” line. Milgram notes that this experiment tells us, “how men act toward others when they are on their own.”<sup>74</sup>

Moral identity findings are in line with these results because in the absence of the experimenter insisting on administering higher shocks, most subjects maintained access to their moral identity. Also, this experiment did not use the “slippery-slope” condition that led subjects on a step-by-step path to inflict progressively greater shocks. The step-by-step application of minimally increasing shocks is noted by Ross and Nisbett as a critical factor that “trapped” some subjects into continuing down the slippery slope of obedience.<sup>75</sup> Note that Doris cites this experiment (see quote above) to show how character traits do not explain subject behavior. But on a moral identity interpretation, when the subjects are allowed to choose the shock level themselves, moral identity

---

<sup>73</sup> Moral identity findings would not support a strong claim regarding whether or not defiant subjects have HMID. In low pressure situations, even LMID individuals may act humanely.

<sup>74</sup> Milgram, 1974, 72.

<sup>75</sup> Ross and Nisbett, 1991, 56.

remains accessible and situational factors play little or no part in affecting that accessibility. While it says nothing of robust character traits, it does tell us something about behaviors when moral identity is accessible.

Another experiment that supports my claim regarding moral identity accessibility is Experiment 15. In this experiment two experimenters are in the room during the experiment. At the 150 volt mark the learner yells that he does not want to continue. The experimenters do not argue but give conflicting orders: Experimenter I says, “We’ll have to stop,” and Experimenter II says, “The experiment requires that we go on. Please continue teacher.” One-hundred percent of the subjects defected in this experiment. My interpretation is that without a clear authority figure driving them on, subjects had easy access to their moral identity and decided to stop the experiment.<sup>76</sup>

Experiment 7 also presented a situation that allowed subjects to maintain access to their moral identity. The subject is told to conduct the experiment alone and to call the experimenter on the phone with any questions. In this experiment 79.5% of the subjects were defiant compared to 35% when the experimenter stayed in the room (Experiment 5). The lack of the experimenter’s physical presence can explain some of the variation in this experiment, but moral identity accessibility most likely played a role as well. My interpretation of Experiment 7 is that subjects maintained access to their moral identities in the absence of the experimenter which resulted in more defiance.

The influence of the experimenter’s presence and prodding cannot be minimized when considering moral identity accessibility. He is more than an objective situational

---

<sup>76</sup> Milgram, 1974, 105-107.

factor like the shock generator. His presence continually reinforces the subject's role and temporal identity as a less-knowledgeable subordinate who signed a contract to participate in the experiment. It is common for us to take on roles and sometimes identities that are directly related to the presence of other people. For example, when my son walks into a room I cannot help but access my identity as a father. When I am attending a lecture, I sit and listen according to my identity as an audience member.

In experiments such as Experiment 5 in which the experimenter stays in the room and acts as the expert and authority, the subject's identity as an "experimental subject" is continually reinforced. The power relationship is also reinforced by the experimenter's often callous insistence that, "The experiment requires you to continue" and "You have no choice." Obedience in Experiment 5 was 65%. However when the experimenter is not in the room as in Experiment 7, it seems that the subject's moral identity becomes more salient than his identity as a subject. The result of this increased accessibility is that obedience dropped to 20%.

In addition to the absence of the experimenter, another factor that seems to make subject moral identity more accessible is bringing the learner into the same room. I suggest that moral identity becomes more accessible because the subject's "harming" behavior becomes more salient when the learner is being shocked in the same room. For example, in Experiment 3 (Proximity) 60% of the subjects were defiant and 70% were defiant in Experiment 4 (Touch-Proximity). From a moral identity perspective, the proximity of the learner may have allowed more subjects to maintain access to their moral identities which includes their direct understanding of their responsibility for their

actions. The proximity of the learner also offset the role-reinforcing presence of the experimenter.

What of those who were obedient in Experiments 3, 4, and 7? From a moral identity perspective, I offer three possible explanations: 1) obedient subjects had LMID from the start which means they started with no access to their moral identity; 2) their identity as a subject overwhelmed any other identity; and/or 3) the self was stuck between acting on a commitment to complete the task for the experimenter and a commitment not to harm another person. Because I have no way to test the first explanation, I take up the other two explanations in the next sub-section.

Two final experiments that support my claim that situational factors influenced moral identity accessibility are Experiments 13 and 13a. In Experiment 13 Milgram wanted to see what would happen if an ordinary man gave commands rather than the experimenter. The experimenter sets up the learner as in the other experiments. He then assigns a confederate to record times at the experimenter's desk and the subject to read words and apply shocks. The experimenter does not tell the subject the order for administering the shocks. The experimenter is then called out of the room and the confederate recommends that a good system for administering the shocks is to increase the shock level step-wise when the learner makes a mistake. He insists on this procedure. Of 20 subjects, 16 (80%) were defiant despite the confederate's persuasive arguments and insistence. I suggest that the orders coming from an ordinary person did not carry enough force to distance most subjects from their moral identities nor to reinforce their

role as a experimental subject. Note again the power of the experimenter's presence on moral identity accessibility.<sup>77</sup> But wait, this experiment continued.

Milgram set up Experiment 13a within Experiment 13 to determine what would happen if the ordinary man insisted that *he* would apply the shocks himself when a subject refused to continue the experiment. In this "bystander" condition, Milgram describes an almost comical situation:

Of the sixteen subjects exposed to this situation, virtually all protested the action of the co-participant; five took physical action against him, or the shock generator, to terminate the administration of shocks. (Several attempted to disconnect the generator from the electrical source; four physically restrained the co-participant.) One, a large man, lifted the zealous shocker from his chair, threw him in a corner of the laboratory, and did not allow him to move until he had promised not to administer further shocks. However passive subjects may have seemed when facing authority, in the present situation five of them rose heroically to the protection of the victim.<sup>78</sup>

What situational factors would have caused these five men, and only these five men, to take physical action against the confederate? I can think of none because the *lack* of the experimenter is not an objective situational factor to consider. Based on moral identity findings, however, I suggest the individual difference of moral identity accessibility resulted in this heroic behavior. What may explain the actions of these five individuals is that their moral identity was highly accessible after having stopped shocking the subject and this accessibility as well as their own individual ID-commitments led them to take drastic action to stop the sadistic confederate.

---

<sup>77</sup> Milgram, 1974, 93-99.

<sup>78</sup> Milgram, 1974, 97.



The other subjects, however, did not follow the actions of the five defiant heroes. Eleven subjects who also refused to continue shocking the learner handed the generator over to the confederate who shocked the learner to the maximum voltage. While virtually all of them vehemently protested, unlike the five heroic subjects they could not find a way to stop the confederate. Milgram was also interested in the effects of groups on subjects and I explain this part of Experiment 13a in the next section.

I have provided several examples of experiments in which situational factors minimally interfered with subject access to their moral identity. I have suggested that moral identity accessibility can explain some of the variance in individual behavior. While my explanations do not explain all of the variance found in the experiments, they do provide a plausible explanation of how moral identity accessibility moderates individual behavior in these experiments. I continue to support my claim below as I provide possible explanations of what happened in experiments with high obedience.

### *3.2.2 – Low Moral Identity Accessibility Situations*

According to my claim about how moral identity accessibility moderates behaviors, the majority of subjects in conditions of strong authority influence and/or distance from the learner should have *less* moral identity accessibility and therefore will tend to obey. Separating the learner behind a wall factors into moral identity accessibility because the results of their actions are less salient from a distance. Unlike the experiments discussed above, in these experiments the situational constraints and the experimenter exert significant pressure on the subjects. The experiments that I would classify as “strong authority/learner distance experiments” are Experiments 2, 5, 6, 8, 10,

16.<sup>79</sup> All of these experiments had an experimenter in the room applying consistent pressure on the subjects to continue his or her duties and the subject could hear the learner's responses from the other side of a wall. The factors that Milgram varied in these experiments were such things as all the subjects were women, the experimenter was easy-going and the learner strong and stern, the location changed from Yale University to an office in Bridgeport, Connecticut, and an additional experimenter acted as the learner. Across these different situational changes, the average obedience was 59% in a range from 65% to 47.5%.

How does a moral identity account explain the obedience and defiance in these cases? Above I argued that low situational pressures allowed subjects to maintain access to their moral identity. In reviewing the data and post-experiment interviews of all experiments, it appears that these experiments revealed three kinds of subjects. The first group is the obedient subjects who had pre-established identities that matched the subject identity. These identities became more salient than their moral identities and they calmly followed the experimenter. The second group is constituted by those subjects who felt significant anxiety and tension between their commitment to the experimenter and their moral identity. The result was anxious obedience or exhausted defiance. The third group of subjects had clear access to their moral identities throughout the experiment and calmly refused to continue. I describe the first and third groups below. While the second "anxiety" group is interesting, they are difficult to interpret for situationists and the moral

---

<sup>79</sup> Experiments 2 and 5 are exactly alike except Experiment 5 is the first time learner complains of a "heart problem." In Experiments 1-4, the learner does not complain of heart problems before or during the experiment. Obedience levels were about the same at 62% and 65% respectively. All experiments after Experiment 5 included the heart condition as a factor.

identity account because they are obviously conflicted between different identities and moral obligations. I want to reiterate that I do not claim that moral identity accessibility explains all variance in subject behavior, but only that its role in moderating behavior is evident in enough cases to explain differences in individual behavior.

The first group of subjects were obedient and I suggest that in addition to the relative inaccessibility of their moral identity in these experiments, the experimenter's presence and consistent pressure may have made pre-existing subject identities more accessible. The role of "subject" in the experiment appears to have triggered similar subservient or obedient identities to which subjects are consciously or non-consciously committed. Evidence for the activation of what I call a "subordinate identity" comes from interview comments recorded by Milgram. While interview comments are few in number, they reveal that some individuals took on the subject identity completely because it was similar to their own subordinate identities.<sup>80</sup>

For example, Karen Dontz was a 40-year old housewife who worked part-time as a registered nurse. Throughout the experiment she dutifully carried out the experimenter's requests. At 225 volts she stated, "I hesitate to press these." When told to continue she obeyed until hesitating at the 345-volt switch. The experimenter assured her that there is no "permanent tissue damage" and she continued. When she continued

---

<sup>80</sup> Milgram collected background information on subjects from the first four experimental conditions and notes some general tendencies between professions and subject behavior. He states, "Those in the moral professions of law, medicine and teaching showed greater defiance than those in the more technical professions such as engineering and physical science. The longer one's military service, the more obedience – except that former officers were less obedient than those who served only as enlisted men, regardless of length of services." He did not make any claims about the predictability of these general observations and behavior. Milgram, 1974, 205.

to push the 450-volt switch, she asked, “What if there’s something wrong with the man, sir?”

After the experiment, she talked about how she was not clear about her rights like at the hospital where she works. She explains that nurses have a right to question a doctor’s orders if they are going to harm a patient. She states, “If I question the dose of a drug, I can ask the doctor three times: ‘Is this the order you want? Is this the order you want? And, if he keeps on saying, ‘Go ahead,’ and I know this is above the average dose, I may call his attention to the fact that it’s too much. It’s not that you are better than he is, but you can say, ‘Did you want her to have so much, doctor,’ and then you repeat it. Then you still have the right to bring the question up to the supervisor.”<sup>81</sup>

Regarding Dontz’s obedience, I suggest that her identity as a subservient person in general and an obedient nurse in particular was activated and guided her through the experiment. According to Milgram, she did not show much anxiety when carrying out the experiment, though she was concerned for the learner. It would appear that Dontz is often distanced from her moral identity in these types of settings and her obedience is produced by the activation of a familiar subservient identity from work. It appears that during the experiment a non-moral identity moderated behavior and produced calm obedience.

Pasqual Gino was another obedient subject whose responses indicate that he accessed another identity during the experiment. He was a 43-year old water inspector and participated in the experimenter-on-the-phone experiment (Experiment 7) that I

---

<sup>81</sup> Milgram, 1974, 77-78.

classified as allowing subjects more accessibility to their moral identity. In this experiment he was alone in the room and called the experimenter when he had questions. While Gino called at different voltage levels because he was concerned with the learner's pleas, he continued past 450-volts until the experimenter returned. After the experiment he stated that he felt tension during the experiment because he was alone. Several months later in a group discussion he described giving the learner shocks until he would not respond anymore. "And then I don't hear no more answer from him, not a whimper or anything. I said, 'Good God, he's dead; well, here we go, we'll finish him.' And I just continued all the way through to 450-volts."<sup>82</sup>

Gino then describes the identity that guided him during the experiment. First of all, he is deferential to Yale because he believes they know what they are doing and he is just a subject. He then identifies his subject role with a similar identity he held as a soldier. In explaining why he obeyed to the end, he relates his training as a soldier. He states, "If the lieutenant says, 'We're going to go on the firing range, you're going to crawl on your gut,' you're going to crawl on your gut. And if you come across a snake, which I've seen a lot of fellows come across, copperheads, and guys were told not to get up and they got up. And they got killed. So I think it's all based on the way a man was brought up in his background." He prided himself on successfully doing his job.<sup>83</sup>

My reading of Gino's comments is that his identity as an obedient soldier or worker moderated and guided his behavior as a subject in the face of the learner's pleas.

---

<sup>82</sup> Milgram, 1974, 86-87.

<sup>83</sup> Milgram, 1974, 88.

I suggest that he had glimpses of his moral identity when the learner said he would not participate anymore. These pleas led him to call the experimenter. But his behavior overall followed his soldier identity with its principles of obedience even against one's instincts.

The second group of subjects that I want to consider are those who appear to have had constant access to their moral identity. Defiant subjects also found themselves reverting to identities, but ones that were true to their moral identity. In these cases, situational factors did not interfere with their moral identity access and may have reinforced its salience. Three subjects in particular represent this group.

Jan Rensaleer was a 32-year old industrial engineer who emigrated from Holland after World War II and was a member of the Dutch Reformed Church. In Experiment 2 he questioned what he should do at 150-volts. It is at this shock level that the subject first insists that he does not want to continue and demands to be let out. Rensaleer is told to continue and he does until 255-volts (Learner response: "Ugh!! Get me *out* of here!"). Rensaleer then pushed himself away from the shock generator and tells the experimenter that he cannot continue because the learner seems to be getting hurt. When the experimenter tells him there is no permanent tissue damage Rensaleer responds based on his identity as an electrical engineer. He said, "Yes, but I know what shocks do to you. I'm an electrical engineer, and I have had shocks . . . and you get real shook up by them—especially if you know the next one is coming. I'm sorry."<sup>84</sup>

---

<sup>84</sup> Milgram, 1974, 51.

The experimenter applied more pressure and told Rensaleer that he “has no other choice.” At this Rensaleer balked. He stated, “I *do* have a choice. Why don’t I have a choice? I came here on my own free will. I thought I could help in a research project. But if I have to hurt somebody to do that, or if I was in his place, too, I wouldn’t stay there. I can’t continue.”<sup>85</sup>

It is interesting to contrast Rensaleer and Gino because they acted differently in the experiment but both referred to serving in an army. The difference between the two is that Rensaleer does not take his soldier identity too far into the experiment. He insists that he should have stopped the first time the learner complained but he was trying to follow the authority like in the army. He explains, “Say, if you’re serving in the army, and you have to do something you don’t like to do, but your superior tells you to do it. That sort of thing, you know what I mean?”<sup>86</sup> He equates his initial obedience in the experiment with being a subordinate under authority, but eventually access to his engineer identity and moral identity lead him to be defiant. Milgram also notes that Rensaleer’s past experience in Nazi-occupied Europe informs his non-compliance to orders and Rensaleer correctly predicts that more people will comply in the experiments than will be defiant.

Another subject who defied the experimenter was Gretchen Brandt, a 31-year old medical technician at the University Medical Center. She emigrated from Germany five years before the experiment. In Experiment 8, which tested all women subjects in the

---

<sup>85</sup> Milgram, 1974, 51.

<sup>86</sup> Milgram, 1974, 51.

standard experiment, Brandt asks the experimenter if she should continue when the learner complains. He says yes and she continues until 210-volts when the learner says, “Ugh!! Experimenter! Get me out of here. I’ve had enough. I won’t be in the experiment anymore.”<sup>87</sup> She informs the experimenter that she will not go on because of the learner’s heart condition. When told to continue she says, “Well, I’m sorry, I think when shocks continue like this, they *are* dangerous. You ask him if he wants to get out. It’s his free will.” She continued to respond that the learner’s free will is important and that she would not like to be the one receiving the shocks. When told she has no other choice, she responded, “I think we here [sic] on our own free will. I don’t want to be responsible if he has a heart condition if anything happens to him.”<sup>88</sup>

Brandt was described as firm and resolute throughout the experiment and showing no signs of tension. In the interview afterwards she reiterated that she did not want to be responsible for hurting the learner. While Brandt’s behavior did not obviously link to an alternative identity, it did demonstrate continuous accessibility to a moral identity that respected the free will of others. She had little or no tension during the experiment and only realized at 210-volts that the learner could be seriously harmed. She took responsibility for his care and respected his free will choice to be part of the experiment in the face of a recalcitrant experimenter. Milgram even commented that her lack of tension and control of her behavior made her defiance appear rational. “Her behavior is the very embodiment of what I had initially envisioned would be true for all subjects.”

---

<sup>87</sup> Milgram, 1974, 56.

<sup>88</sup> Milgram, 1974, 85.



On a final note, Brandt was asked if her adolescent years in Germany may have influenced her behavior. She remarks, “Perhaps we have seen too much pain.”<sup>89</sup> Again we see how moral identity accessibility and possibly a hint of national identity moderated Brandt’s behavior.

I present one final interview that reveals the strength of moral identity, whether combined with a similar identity or not, in moderating individual behavior during Experiment 3 (Proximity). The subject in question is only identified as an Old Testament professor at a major divinity school. He looks at the learner while giving the shocks. The professor draws back his lips and bares his teeth whenever he administers shocks. At 150-volts, the learner says that he does not want to continue with the experiment and insists on being let go. The professor will not go any further. When told that it is essential that the experiment continue, he explains that he does not understand why the experiment is placed above the learner’s life. He states, “If he doesn’t want to continue, I’m taking orders from him.”<sup>90</sup> At this point the professor is completely in control of the experiment. He has complete access to his moral identity and is clear about what behavior is moral and immoral. When he is told he has no choice, he references his identity as an American and replies, “If this were Russia maybe, but not in America.”

Throughout the experiment, Milgram notes that the professor connects with his profession (i.e. occupational identity) as a teacher of religion. The professor even questions the *ethics* of the experimenter. “Surely you’ve considered the ethics of this

---

<sup>89</sup> Milgram, 1974, 85.

<sup>90</sup> Milgram, 1974, 48.

thing. Here he doesn't want to go on and you think that the experiment is more important? Have you examined him? Do you know what his physical state is?"<sup>91</sup> The professor is seeking to know if the experimenter has forgotten his moral commitments because the professor has access to his. The professor also explains that he must obey God (i.e. religious identity). After the experiment, the experimenter asks, "What in your opinion is the most effective way of strengthening resistance to inhumane authority?" The professor responds, "If one had as one's ultimate authority God, then it trivializes human authority."<sup>92</sup>

The professor provides a clear example of how other identities and especially an accessible moral identity can moderate behavior. For him, the situational factors did not distance him from his moral identity and may have made it even more accessible. His identity as a "servant of God" may have also played an important part in moderating his behavior. The professor's moral confidence appears to be similar to the "super-helper" seminary students who the victim had to entreat to leave. These subjects had a constantly accessible moral identity that situational factors did not touch.

The situationist may object at this point and argue that situational factors still explain the triggering of moral identity and alternative identities. The situationist may argue that the learner's response is a *situational* cue that produces subject behavior regardless of his or her moral identity accessibility. Ross and Nisbett consider the learner's withdrawal of his implied consent at 150-volts to be a situational factor that

---

<sup>91</sup> Milgram, 1974, 48.

<sup>92</sup> Milgram, 1974, 49.

opened a channel for defiant behavior.<sup>93</sup> This is the point at which the subject, who presumably is caught in the routine of shocking the learner in incremental steps, confronts the following plea, “Ugh!!! Experimenter! That’s all. Get me out of here. I told you I had heart trouble. My heart’s starting to bother me now. Get me out of here, please. My heart’s starting to bother me. I refuse to go on. Let me out.”<sup>94</sup> At this point the subject has a rationale for stopping the step-wise progression of shocks and 32% of all defiance in the first ten experiments occurred at this point (66 of 205 defiant subjects). If I include four other points of direct learner refusal to participate further (180-, 300-, and 315-volts), defiance reaches 61% for the first ten experiments (125 of 205 defiant subjects).

Ross and Nisbett describe these points as “channel factors”, a term coined by Kurt Lewin to indicate “small but critical facilitators or barriers” within situational details that can constrain or free behavior. They hypothesize that if the experimenter told the subjects that they could push a button if they wanted to terminate the experiment, defiance would have increased significantly. The button would be the channel factor that would allow the subjects to free themselves from obedience. Milgram, however, denied subjects this channel which led to greater obedience.<sup>95</sup> Like Ross and Nisbett, Doris considers channel factors as situational factors that would change behavior. He states, “Milgram’s lesson is not simply that situational pressures may induce particular

---

<sup>93</sup> Ross and Nisbett, 1991, 56.

<sup>94</sup> Milgram, 1974, 56.

<sup>95</sup> Ross and Nisbett, 1991, 57.

*undesirable* behaviors, but that situational pressures may induce particular behaviors, *period.*”<sup>96</sup>

While I agree with Ross and Nisbett’s thought experiment about channel factors, characterizing the learner’s complaints as merely another situational factor seems shallow in light of the mental constructs active during the experiments. When a person pleads to be set free and not shocked, this is a person-to-person interaction that seems to go beyond the classification of “objective situational factor.” The learner’s complaints have a different moral status than a situational factor like a hurry-condition or a shock generator button sequence. It is a human-to-human communication that “You are harming me.”

When the learner yells that he does not want to continue, I would suggest that the moral identity of most subjects became readily accessible and this explanation is well supported by the high number of defiant subjects at these particular protest points and the testimony of Rensaleer, Brandt, and the Professor. At these points in particular, a person with an accessible moral schema would perceive that her actions were contrary to her moral values. The proximity experiments in particular support a moral identity explanation of defiance triggered by direct learner complaints. For example, in Experiment 4 (touch proximity), 16 of 40 subjects were defiant at the 150 volts point. As in the earlier examples, subject behavior can plausibly be explained by moral identity accessibility and not merely a situational factor inducing a response.

A situationist could also object that pre-established roles and identities may have some influence on behavior, but it seems ad hoc to suggest that identities can moderate

---

<sup>96</sup> Doris, 2002, 51.

behavior in the Milgram experiments. The experimenter's presence and comments most likely influenced behavior more than a subject's subservient identity, especially for such a short period of time. My response is that even temporary identities can have a powerful influence on moderating behavior depending on how deeply they are accepted and reinforced. If the situation is total and the agent completely accepts a new identity, a person can even forget who she is. Consider the Stanford Prison Experiment, another common situationist example. In 1971, psychologist Phillip Zimbardo took several college students and set-up a simulated prison in a building at Stanford University. Students were randomly assigned to be guards and prisoners. Because of constant abuse by the guards as well as Zimbardo's own desire to see the prison work, some prisoners began to have emotional breakdowns after only 3 days.

Zimbardo describes prisoner 819 (no names were allowed during the experiment) who was starting to have a breakdown. After 819 ripped up his pillow and mattress, the guards put him in solitary confinement where he was hysterical. Zimbardo agreed to release him and brought him to a recreation room. Zimbardo said, "Okay, 819, look, time is up, we're going to pay you for the whole time."<sup>97</sup> At the same time, the guards have the prisoners chant that 819 is a bad prisoner and the others are being punished for it. The prisoner said, "I've got to go back! I've got to go back and prove that I am not a bad prisoner." Zimbardo recalls that at that point Zimbardo himself realized that the prisoner did not know who he was. He said, "Wait a minute, you're not a prisoner, you're not 819,

---

<sup>97</sup> Interview with Phillip Zimbardo in Tamler Sommers, *A Very Bad Wizard: Morality Behind the Curtain*. (Canada: Believer Books, 2009). 38.

this is an experiment, you're a student, your name is Stewart . . . and I'm Phil Zimbardo."<sup>98</sup> The Stanford Prison experiment demonstrates not only the power of the situation, but also the power of the identities a person can quickly embody and how they can distance a person from knowing who she is. I would conclude that it is more than plausible that in the Milgram experiments some subjects could moderate behavior using an occupational identity, a moral identity, or the subservient subject identity.

To summarize, I have argued that situational factors alone do not fully account for variations in individual behavior in the Milgram experiments. By considering how a person's moral identity moderates situational factors when determining behavior, we have a more complete account of how subjects behaved across different situations. My re-interpretation of the experiment results are not intended to refute situationism, but rather to support my claim that the mental construct of moral identity accessibility played a moderating role given situational factors. I have provided a description of the experiment results that support this hypothesis as well as direct comments from subjects that describe how they relied on access to occupational, subject, and moral identities to inform their behaviors.

### *3.2.3 – Moral Identity Accessibility and Responsibility*

My second claim regarding the Milgram experiments and moral identity is that if defiant subjects tend to have a more accessible moral identity, then they will take more responsibility for their behaviors. My claim assumes that subjects with more access to their moral identity will tend to defy the experimenter as I argued in 3.2.1 above. I base

---

<sup>98</sup> Sommers, 2009, 38.

my claim on the findings presented in section 2 that people with more accessible moral identity have a more expansive circle of moral concern. In other words, they take more personal responsibility for the welfare of others. My basis is tentative because more research needs to be done in this area. To establish a connection between defiance, moral identity, and responsibility, I first consider the role responsibility played in the experiments overall and then look at additional evidence that defiant subjects tended to take more responsibility for their actions than obedient subjects.

The connection between moral identity accessibility and responsibility comes primarily from Schlenker's extensive work on responsibility and excuses as presented above. On Schlenker's model, we tend to consider ourselves and others more personally responsible for outcomes when we know clear prescriptions that apply to a situation (e.g. clear principles), we are bound and obligated to follow those prescriptions (e.g. commitment to principles), and we appear to have control over the outcomes. For example, a person with constant moral identity accessibility would know that moral principles apply in a particular situation, would consider herself to be obligated to follow them, and would consider herself to be in control of her actions and the outcomes of the situation.

The Milgram experiments created significant ambiguity for most subjects in the areas of principles, commitments, and control of outcomes. For some subjects, the moral principles that applied in the experiment were not clear and their commitments were divided between the learner and the experimenter. Other subjects who had constant access to their moral identity understood the principles that applied regarding the learner

and understood their commitment not to harm another person. The experimenter's presence and intransigence also gave the impression that the subject's did not control the outcomes and that few behavioral options were available. Some accepted his control and others acknowledged that they were the ones doing the shocking. The learner's inability to learn the words was also out of the subject's control.

Attempts by many subjects to clarify what moral principles applied and who was responsible for harming the learner reveal the ambiguity of the experimental conditions. Several subjects continued to ask the experimenter some form of the question, "Do you take full responsibility for what happens?" Milgram suggests that people in the experiment are locked into a subordinate position and many lose their sense of responsibility because the experimenter takes full responsibility for what happens. He notes, "In the post-experimental interview, when subjects were asked why they had gone on, a typical reply was: 'I wouldn't have done it by myself. I was just doing what I was told.' Unable to defy the authority of the experimenter, they attribute all responsibility to him."<sup>99</sup>

In a position friendly to mine, Milgram suggests that the loss of responsibility comes from a loss of "the self." For some subjects, the situation manipulates the connection between the subject's self and her action and instead emphasizes the connection between the experimenter's will and the subject's action. The self/action disconnect reduces the subject's responsibility for his or her actions. Instead of action

---

<sup>99</sup> Milgram, 1974, 8.



flowing from the motives of the “self” who often wants to stop the experiment, action flows from the motives of the experimenter.

Milgram’s explanation of responsibility is similar to the moral identity accessibility account. He suggests that situational factors pressure the subjects to replace their intentions with the experimenter’s intentions, while along a similar line I suggest that the situational factors make the moral identity of many subjects less accessible. In either case, subjects were unsure of what principles applied, what commitments they had, and who controlled the outcomes. The result was subjects attributing responsibility to the experimenter, the learner, and themselves to different degrees.

In what follows I use the same evidence Milgram uses in his interpretation of responsibility and make my case that people who maintained access to their moral identity tended to defy the experimenter more often and attributed more responsibility to themselves. On my account, taking responsibility is a sign of moral identity accessibility. My evidence takes the form of comments from three subjects and also information from post-experiment interviews.

First consider how responsibility is handled by an obedient subject. Fred Ponzi was an obedient subject who experienced significant tension and conflict during the experiment. One of the keys to encouraging his obedience was the experimenter taking full responsibility for what happened to the learner. At 180-volts Ponzi stops because he is very concerned with the learner’s health. The experimenter states, “Whether the learner likes it or not, we must go on, through all the word pairs.” Ponzi then plays his responsibility card: “I refuse to take the responsibility. He’s in there hollering! . . . I mean

who's going to take responsibility if anything happens to that gentleman?" The experimenter exploits this question and states, "I'm responsible for anything that happens to him. Continue please." This answer seems to satisfy Ponzi and he continues. Later in the experiment he again asks, "You accept all responsibility?" and the experimenter says, "The responsibility is all mine. Correct. Please go on."<sup>100</sup> Ponzi continues until told to stop. When discussing the experiment he said he continued because the experimenter told him to. His actions were primarily the experimenter's responsibility and the experimenter said as much during the experiment.

On my account, Ponzi seems to have access to moral principles because he pleads with the experimenter to check on the learner and to stop the experiment. But why didn't he stop? On Milgram's account, a determining factor was the experimenter relieving Ponzi of his responsibility for his actions. The experimenter's authority and Ponzi's subject identity carried enough weight to make Ponzi continue against his will.

In contrast to Ponzi, consider two defiant subjects described above and how they reacted when it came to responsibility. Brandt stopped the experiment specifically because she did not want to be responsible for harming the learner. When told that it is "absolutely essential that we continue," she interrupts and says, "I don't want to be responsible for anything happening to him. I wouldn't like it for me either." When the experimenter tells her that she has no choice she responds, "I don't want to be responsible

---

<sup>100</sup> Milgram, 1974, 75-76.

if he has a heart condition if anything happens to him.”<sup>101</sup> In the interview she reiterated not wanting to be responsible for harming the learner.

Rensaleer, who hesitated at 150-volts and finally stopped at 255-volts, also discusses responsibility at length after the experiment. In fact, he took full responsibility for shocking the learner against his will. He states, “One of the things I think is very cowardly is to try to shove the responsibility onto someone else. See, if I now turned around and said, ‘It’s your fault . . . it’s not mine,’ I would call that cowardly.”<sup>102</sup> Milgram comments that Rensaleer is hard on himself and does not allow the authority structure of the experiment to reduce his personal responsibility.

By comparing these brief interviews and Milgram’s comments on responsibility, I suggest that one difference between some obedient and defiant subjects in the area of responsibility is that the defiant subjects had continual access to their moral identities, which means they understood the principles that applied, they were committed to them, and they took control of the results. In other words, taking responsibility is an *effect* of moral identity accessibility. In Milgram’s terms, subjects who did not lose “self” knew that they were responsible and those who lost “self” knew that the experimenter was responsible. Those subjects like Ponzi who struggled with who was responsible knew that they should not harm the screaming learner, and yet they could not access (or did not have) any ID-commitments to not harm others. Meanwhile, calm defiant subjects had a

---

<sup>101</sup> Milgram, 1974, 85.

<sup>102</sup> Milgram, 1974, 52.

moral identity, accessed it during the experiment, and took responsibility for their actions.

Milgram's responsibility research and additional responsibility studies also support my interpretation of the link between defiance, moral identity accessibility, and responsibility. Because Milgram identified the issue of responsibility and obedience as a possible pattern, he had the interviewer ask subjects questions after the first four experimental conditions. Subjects were asked, "How much is each of us responsible for the fact that this person was given electric shocks against his will?" They proportioned responsibility using a pie-chart model that represented the responsibility of the experimenter, the learner, and the subject. In reviewing the results, Milgram states, "The major finding is that the defiant subjects see *themselves* as principally responsible for the suffering of the learner, assigning 48% of the total responsibility to themselves and 39% to the experimenter." The obedient subjects assigned about the same amount of responsibility to the experimenter (38.4%), but less to themselves (36%) and much more to the learner (25% compared to 13%). They reasoned he was more responsible because he "volunteered and did not learn very efficiently." Overall, Milgram found that defiant subjects took more personal responsibility for their actions and attributed less responsibility to the learner.<sup>103</sup>

Other obedience studies have found a similar connection between defiance and responsibility for one's actions toward others. In a Milgram type obedience experiment in Austria, Thomas Blass notes that defiant subjects had a "greater tendency to accept

---

<sup>103</sup> Milgram, 1974, 204.

responsibility for their actions than the obedient subjects.”<sup>104</sup> Milgram followed up his own experiments with an analysis of personality and obedience. He and a fellow researcher tested 40 former participants who had “gone against the tide” in the first four experimental conditions. Twenty participants had been defiant in the remote or voice-feedback experiments and 20 had been obedient in the proximity or touch-proximity conditions. They found that defiant subjects scored significantly higher on social responsibility measures.<sup>105</sup>

In summary, subject responses and other research findings support my claim that moral identity accessibility, responsibility, and defiance appear to be linked. In some cases defiance appears to arise from moral identity access as indicated by subjects taking more responsibility for their actions. On my account, unlike many obedient subjects, the defiant subjects did not lose moral identity accessibility and therefore they understood moral principles and that they were responsible for the outcomes of their actions.

#### *3.2.4 – What About Bystander Cases?*

Before concluding this section, I must briefly address so called “bystander” cases and how they relate to moral identity. Because bystander experiments resulted in some of the lowest and highest obedience percentages of all 18 experiments, it is important that I not ignore them in regards to moral identity and situationist skepticism. Below I review the three experiments that tested group effects. While it is beyond the scope of this

---

<sup>104</sup> Thomas Blass, “Understanding Behavior in the Milgram Obedience Experiment: The Role of Personality, Situations, and Their Interaction.” *Journal of Personality and Social Psychology*, Vol. 60, No. 3, (1991): 404.

<sup>105</sup> Blass, 1991, 402. See, A. C. Elms and Stanley Milgram, “Personality Characteristics Associated with Obedience and Defiance Toward Authoritative Command.” *Journal of Experimental Research in Personality*, 1 (1966).

chapter to review the extensive literature on group and bystander effects on helping behavior, I do layout an account of moral identity accessibility that is in line with two findings in particular. First, the actions or inactions of others can influence a person to act in line with others. Second, the presence of other people can diffuse the responsibility a person takes for the actions of the group.<sup>106</sup> I refer back to these findings as I review the three experiments.

In Experiment 17, Milgram paired the subject with two confederates. As a reminder, one confederate reads the list of word pairs, the second tells the subject if the answer is correct or incorrect, and the subject administers shocks. At 150-volts, the first confederate refuses to continue against the insistence of the experimenter. The subject must now read the word pairs and administer shocks. At 210-volts the second confederate refuses to continue even though the experimenter orders him to continue. He says, "I'm not willing to shock that man against his will." The results of the experiment are that 63% of the subjects were defiant after the 210-volt level and 90% were defiant overall.

Milgram suggests that confederate defiance may have been effective because confederate refusal to continue opened the door for subject defiance and as the confederates refused to participate it became obvious that responsibility rested solely on

---

<sup>106</sup> The seminal discussion of group effects such as conformity are found in Solomon Asch, *Social Psychology*. (New Jersey: Prentice Hall, 1952) and summarized in Ross and Nisbett, 1991, pp. 30-35. Discussions of diffusion of responsibility are covered in John M. Darley and Bibb Latané, "Bystander Intervention in Emergencies: Diffusion of Responsibility." *Journal of Personality & Social Psychology* 8. (1968). A review of many of these studies with updated research is Robert P. Abelson, Kurt P. Frey, and Aiden P. Gregg. *Experiments with People: Revelations from Social Psychology*. (New Jersey: Lawrence Erlbaum Associates, 2003).

the subject.<sup>107</sup> I would agree with Milgram that defiance initiated by the confederates clarified the situation so that the responsibility for harm done became more salient to the subject. From a moral identity perspective, confederate defiance may trigger access to the subject's moral identity and emphasize the subject's responsibility for harming the learner. These are only speculations because several factors are at play.

For example, subjects may see confederate defiance as a way out of an uncomfortable situation, or they may bow to the pressure of two onlookers as opposed to the experimenter, or defiant subjects may only wish to conform with others. Because so many variables are acting in this scenario, I can only speculate that moral identity access increased when the confederates defied the experimenter and this resulted in a heightened sense of responsibility which produced defiance.<sup>108</sup> Situationists, however, are in the same uncomfortable boat as I am in explaining the results. They would need to show that situational factors led to such high defiance rather than internal desires to conform, heightened responsibility, moral identity accessibility, and/or group pressure.

Experiment 18 had quite different results. In this experiment the subject performs subsidiary duties while a confederate operates the shock generator. While Milgram gives few details of this experiment, he notes that 92% of the subjects were obedient and only 3 subjects refused to participate in the experiment to the end. He explains the results in terms of the subject being doubly absolved from responsibility. "First, legitimate authority has given full warrant for their actions. Second, they have not themselves

---

<sup>107</sup> Milgram, 1974, 118.

<sup>108</sup> From a moral identity perspective, the most interesting subjects are the four who obeyed the experimenter even after the confederates refused to continue. Possibly an alternative identity was triggered and they calmly obeyed to the end.

committed brutal physical acts.”<sup>109</sup> On the moral identity account, it appears that the defiant 8% had access to their moral identity and acted accordingly. But what about the other 92%? I follow Milgram in speculating that bystander situations diffuse responsibility across different people and therefore the subject who is only performing minor tasks does not take responsibility for harming the learner because he is not actually performing the shocks.

If we compare Experiment 18 findings with experiment 13 and 13a, we get a clearer picture of how moral identity accessibility may work in bystander situations. As a reminder, Experiment 13 produced 20% (4 of 20) obedience when an ordinary man (i.e. confederate) sets up the step-wise shock progression and gave orders to increase the shocks. In Experiment 13a, the ordinary man takes over the shock generator from the defiant subjects and 68% of the subjects vehemently protest but do not stop the confederate while 32% of them physically stop the confederate from continuing.

The common thread in these experiments does not seem to be that moral identity is less accessible in bystander situations. The common thread is that the principles and commitments that constitute some people’s moral identity involve *not personally harming others*. They do not appear to involve *physically stopping someone else from harming others*. Principles and commitments to not harm others are those by which most people monitor themselves everyday. Common expressions such as “I’m sorry,” “Excuse me,” and “Please forgive me,” indicate that we are taking responsibility for personal actions that may have harmed someone. Rarely will someone apologize for not stopping

---

<sup>109</sup> Milgram, 1974, 122.



a third party much less two people (experimenter and confederate) who do something wrong. Another plausible explanation for high obedience in bystander experiments is that many people have developed a commitment to not harm others but have not developed commitments to physically stop someone from harming others. Moral identity may be strongly averse to personally harming someone but only weakly averse to stopping a bystander from harming someone. This may not be an ID-commitment and therefore not part of a person's self-concept.

When we consider Experiment 13a we see the distinction between personally harming someone (20% obedience) and stopping someone else from harming someone (68% obedience). These appear to be distinct orientations that draw from different commitments. At least some subjects, possibly because of having moral identity accessibility, physically stopped the confederate. Or maybe they had an identity that required them to stop an ordinary person from harming another person. More research needs to be done in this area to discover how the commitment to stop someone from harming others differs from the commitment to not personally harm someone. As mentioned above, the former seems to be encouraged in society while most of us rarely have the opportunity to develop the latter.

My conclusion from the bystander experiments is that moral identity may be most accessible when a person is confronted with personally harming a person. When the subject's moral identity is accessible, she tends to take more responsibility for her actions. The subject who must watch as another person shocks the learner may still have moral identity accessibility, but her action options are limited. She can quit, protest,

and/or fight the confederate, and we saw each of these options played out in the experiments. Because limited options are available and the experimenter takes responsibility for the experiment and the confederate's actions, I would speculate that the subject actually has less responsibility for what happens and therefore is more obedient, if we can call watching someone shock another person obedience.

### **3.3 - Situationists, Moral Identity, and Integrity**

In this section I have argued and supported my claim that situational factors were mediated by moral identity accessibility in the Seminarian and Milgram experiments. Using the Seminarian experiment, I presented a moral identity interpretation that offers a more complete explanation of subject behavior than the situationist account, in particular the behavior of super-helpers. In my review of the Milgram experiments, I supported my claim that moral identity accessibility can explain individual subject behaviors by presenting an analysis of subject responses and interviews, specific experimental conditions and different individual reactions, and the link between moral identity accessibility and responsibility. In summary, my response to situationists is that the same experiments that generate skepticism about robust character traits also generates evidence that individual moral identity accessibility may mediate behavior across situations.

My response to situationists defends the notion of integrity from unwarranted skepticism about the existence of stable mental constructs that are necessary for integrity. Invoking situational factors to describe behavior does not fully explain behavior across situations given the evidence for moral identity accessibility and how it appears to moderate situational factors and behavior. Having defended this position and developed

a fuller account of how the self uses moral identity to mediate responses to situations, I propose two additional necessary conditions for a person to have substantive integrity.

*A person of substantive integrity is a person with 1) HMID and 2) whose moral identity is constantly accessible (even when confronted with situational factors).*

The first necessary condition is that a person of substantive integrity must have HMID. This condition is necessary because it recognizes that the person must have a moral identity composed of ID-commitments to moral values and the determination to consistently act on them. A person with LMID lacks the ID-commitments to be a person of substantive integrity, though in a morally supportive environment she may act like a person of substantive integrity. The second necessary condition is that she must maintain access to her moral identity across situations. Here the situationist may argue that this is highly improbable, but I disagree. The necessary condition does not require the person to never make a mistake across all situations. It does require her moral identity accessibility to be situationally immune in moral matters most of the time and in some cases, all the time (e.g. fidelity to others). This condition allows us to learn from our mistakes and correct our behavior “next time.”

One concern about this condition is that it may be overly strict or too ideal for “mere mortals” to ever have integrity. I disagree because some mortals meet this condition more often than they fail to meet it. Consider the actual behavioral consistency of holocaust rescuers, corporate whistleblowers, moral exemplars, and some of the

HMID, non-primed subjects in the experiments above.<sup>110</sup> Their lives and laboratory performances reveal that situational factors do not make moral identity less accessible for everyone. With the right moral schema in place, situational factors may make it more accessible for some people. Even though the investment decision experiment above reveals that situational factors generally do have a positive and/or negative effect on moral identity accessibility for thirty-three undergraduate business students, these results can only be generalized so far. They do not warrant a situationist interpretation of most human behavior nor does it invalidate the life-long histories of moral exemplars. The second condition may be an ideal, but it is within reach of those who strive to develop and maintain a moral identity.

Another concern about the second condition is that we may not be able to protect ourselves from situational factors. Life throws us difficult situations that may reduce our moral identity accessibility to a point where a person of substantive integrity will slip. I agree that if situational pressures can be ratcheted up so high (e.g. a gun to my head) that a normal person may do a range of unthinkable actions. But in everyday life, living a life of integrity is a process of bringing one's behaviors into alignment with one's moral identity rather than achieving an unchangeable character. One way of maintaining this

---

<sup>110</sup> One only needs to think of whistleblowers who consistently act on their moral values in the face of career ruin. One of the most famous examples is engineer Roger Boisjoly who, after the space shuttle Challenger exploded, believed that the truth behind the controversial decision to launch the shuttle should be known. His ID-commitment to honesty led him to tell the truth to investigators about pre-launch safety discussions and how safety was sacrificed for meeting a deadline. He did this against advice of his managers. Another example is internal auditor Cynthia Cooper who was told by one of the nation's most respected Chief Financial Officers, Scott Sullivan, to stop auditing the accounting records of Worldcom. She refused and eventually uncovered accounting tricks that resulted in the largest bankruptcy in United States history at the time. For the Boisjoly story see Russell Boisjoly, Ellen Foster Curtis, and Eugene Mellican. "Roger Boisjoly and the Challenger Disaster: Ethical Dimensions." *Journal of Business Ethics* 8, (1989). For Cooper's story see Richard Lacayo and Amanda Ripley, "The Whistleblowers." *Time*. (December 30, 2002).

alignment is to take one or two “situation vaccinations” that can protect as revealed by moral identity research.

A person on the first vaccination can maintain her integrity by denying conflicting identities access to her working self concept. For example, a lawyer may establish at the start of her career that she will never lie for a client. The more she respects this ID-commitment to honesty, the more likely her occupational identity will drop lying as a possible behavior. Other people do not need to deny other identities because their childhood education and role models make some immoral actions unthinkable. Lawrence Blum describes Magda Trocmé, a courageous French rescuer of Jews in WWII, as a person who did not see other alternatives to rescuing others. She stated, “I do not hunt around to find people to help. But I never close my door, never refuse to help somebody who comes to me and asks for something.”<sup>111</sup> Her mother always helped people in need and she naturally followed this path.

A second situation vaccination a person can use is to reinforce her moral identity accessibility by regularly devoting herself to causes, relationships, and belief systems that espouse moral values. The moral identity evidence supports moral priming as a reinforcement mechanism for moral identity accessibility, and it seems from research presented above that a person can maintain her *own* moral identity accessibility through self-initiated devotions such as regular religious attendance or becoming grounded in a moral world-view.

---

<sup>111</sup> Lawrence Blum, *Moral Perception and Particularity*. (New York: Cambridge University Press, 1994), 86.

The situationist may argue that the second vaccination is just situationism dressed up like an internal mental construct. Even Aquino et. al. hint at situational dominance when they attempt to reconcile the situational instability of moral identity with extensive research on moral exemplars. Like situationists, they suggest that a person with HMID may live and work in environments where her moral identity is continually activated and reinforced. Over time she gains a sustained commitment to moral action. In contrast, a person with HMID who works in a highly competitive industry may have an occupational identity reinforced daily thereby reducing her moral identity accessibility. Therefore, she may be more susceptible to situational factors that result in compromising her moral identity.<sup>112</sup>

I agree with this description, but the evidence does not suggest that situational factors are the *primary* cause of consistent moral behavior. It is common for people to purposely integrate non-situational reinforcements into their lives. Situationists like Ross and Nisbett may object and argue that consistent behavior is primarily a result of people choosing the situations in which they want to live and they alter these situations to fit their preferences and competencies. As mentioned earlier, they argue that clergy, physicians, and rock stars made choices that continue to place them in situations that support consistent behavior. Their choices continually shape their situations and dictate responses from others. They comment that individual perceptions and assumptions may have been based initially in “deep personal convictions or on relatively incidental

---

<sup>112</sup> Aquino et. al., 2009, 139.

differences in situational construal,” but now consistent behavior is situation-based and self-reinforcing.<sup>113</sup>

Ross and Nisbett may be correct that most people behave consistently because they set up predictable lives with predictable friends, jobs, houses, and pets. But this characterization ignores the deep personal convictions and passions that not only launch some lives but maintain their resilience through hardships and failures. Situational factors do not fully account for the devotion some people develop to live in accordance with high moral standards or an over-arching purpose. They also fail to consider how daily devotion can emerge from a self that desires to keep its moral identity accessible. One doesn't need to look too hard to find people with HMID who daily devote themselves to a cause, a relationship, and/or a belief system which reinforces their moral identity accessibility. These non-situational factors can result in consistent behavior across diverse situations over the long run.

I reiterate that I agree that situational factors can support consistent moral behavior. That is why I do not claim that a person has substantive integrity if she acts on moral values only within carefully selected relationships and environments. While it is not required that she be tempted to compromise her ID-commitments for us to know she is a person of integrity, she does need to demonstrate during her lifetime that she has ID-commitments to moral values that produce consistent behavior across situations in the long run, especially in situations that may reduce her moral identity accessibility.

---

<sup>113</sup> Ross and Nisbett, 1991, 155-156.

For example, if situational factors influence her to lie at work in order to get a promotion but have no affect at home, she is not a person of substantive integrity.<sup>114</sup> Her moral identity is not continually accessible. But if she recognizes that corporate pressures make her moral identity less accessible, she can change over time, possibly change jobs, and progress toward substantive integrity. On the moral identity account, becoming a person of substantive integrity may be a process of stumbling, learning, and getting back up to act in a way that does not violate her ID-commitment to moral values. Substantive integrity is not a constant state of unbroken righteousness but rather a journey to greater clarity of principles and control of behaviors.

### **Conclusion**

In this chapter I have argued that situationist evidence against stable mental constructs miss important moral identity considerations. I have not attempted to reject all situationist conclusions because I acknowledge the power of situational factors on moral identity accessibility and behavior. Instead I have attempted to show that recent findings in moral identity research provide additional tools for interpreting experiments often cited by situationists. I have also demonstrated that moral identity constructs have sufficient empirical support to integrate situationist insights without denying the existence of integrity.

I conclude that substantive integrity is possible and that situationist skepticism about stable mental constructs is not universally applicable. The existence of stable

---

<sup>114</sup> In Chapter 3 I described how holocaust rescuer John Weidner lied to the Gestapo to protect innocent people in his care. I explained that he remained a person of integrity because he only lied to achieve a higher objective good (i.e. save innocent lives) and dishonesty in these rare cases did not fundamentally alter his ID-commitment to honesty.



mental structures such as HMID and constantly accessible ID-commitments to moral values make situational factors only one input into the description of human behavior. In addition to situational factors, the moderating role of moral identity may perform an important role in how a person acts across a variety of situations. These mental structures also form the ground from which a person of integrity can consistently, though not perfectly, act on her ID-commitment to moral values and reinforce her accessibility to these ID-commitments.

The issue of occupational identities took a prominent place in my discussion about conflicting identities because these identities often subtly distance a person from her moral identity. In the final chapter I address a concern that the *quest* for integrity may cause people to compromise their moral values and deceive themselves regarding their conduct.

## Chapter 5 - The Quest for Integrity: Of Self-Deception and Self-Defense

Few people enjoy admitting when they have done or said something wrong. A preferred method of covering up bad behavior and restoring a coherent and upright self-image is to rationalize our behavior to ourselves and others. Carol Tavris and Elliot Aronson explain that, “Our convictions about who we are carry us through the day, and we are constantly interpreting the things that happen to us through the filter of those core beliefs. When they are violated, even by a good experience, it causes discomfort.”<sup>1</sup> In other words, events and actions that do not match our self-conception create tension within us. Imagine a CEO whose latest product strategy is failing, but she believes that she is an excellent strategist. In an effort to reduce the dissonance between her self-conception and the failed strategy, she may blame the economy, the employees who are executing the strategy, or customer habits.

But what happens when a person genuinely believes she is an ethical person even though she performs unethical actions as part of her profession. These cases are of interest to integrity theorists because the tension between conduct and self-conceptions could lead people to change their beliefs and moral principles. David Luban expresses great concern about the quest for integrity itself becoming a morally dangerous enterprise. He argues that our desire for coherence and harmony between our beliefs and conduct can drive us to unconsciously change our moral values to match our unethical conduct. To make the quest for integrity more harmful, our desire to appear upright to ourselves and others may motivate us to rationalize our bad conduct and deceive

---

<sup>1</sup> Carol Tavris and Elliot Aronson, *Mistakes Were Made (but not by Me)*. (New York: Harcourt, 2007). 13.

ourselves. Because the quest for integrity may strengthen these tendencies, integrity can be a harmful and dangerous pursuit. Luban's concern is that the more we change our beliefs to match our unethical conduct in the quest for integrity, the more damage we will inflict upon others and ourselves.

If we look at Luban's concerns from a moral identity perspective, his claims are no less frightening. He is arguing that a person may start with substantial integrity but because of her desire for self-consistency and uprightness, she will unconsciously move to a state of formal integrity with ID-commitments to immoral values. If the quest for substantial integrity produces changed moral beliefs, rationalization, and self-deception, then integrity should not be pursued but rather avoided.

In this chapter I consider Luban's arguments for why the quest for integrity is dangerous. In section 1 I present Luban's two descriptions of integrity and his evidence that cognitive dissonance reduction and other factors can turn the quest for integrity into a recipe for self-deception. While I acknowledge that Luban has found a serious problem in human character, I defend the moral identity account by questioning our ability to change our moral commitments and presenting an alternative explanation for the rationalization and self-deception that resulting from the quest for integrity. I also briefly discuss the life of essayist and French government official Michel de Montaigne. Philosophers such as Luban use his supposed ability to separate himself from his position as Mayor of Bordeaux to raise concerns about pursuing integrity. In section 2 I conclude the dissertation by discussing how the moral identity account can give us insight into how

we can maintain integrity given situational and professional pressures to compromise moral values.

Following my usage in Chapter 4 and in line with the literature on the subject, I continue to use the term “moral identity” to describe an aspect of a person’s total self-conception that is characterized by identity-conferring commitments (i.e. ID-commitments) to moral values. As I have argued in Chapter 4, a person of substantive integrity must have a highly centralized moral identity (i.e. HMID) that is constantly accessible across situations. I am following Albert Bandura’s social cognitive model in my usage because his conceptualizations are central to the moral identity research being conducted today.<sup>2</sup> His description of identity also fits well with the discussions of cognitive dissonance and the demands of professional roles discussed below. In line with these constraints, when I say a person has a “moral identity” I am saying that she has a moral self-concept that is one of several which constitute her self-conception.

### **Section 1: Integrity as a Dangerous Quest**

Many internal and external factors influence how we interpret the world and also how we interpret our own actions. In the previous chapter I argued that a person’s moral identity *centrality* and *accessibility* moderates how she interprets situations and which actions she takes. I also acknowledged that a person who daily interacts with others in an environment that emphasizes other non- or immoral identities may find her moral identity less accessible to her working self-concept. The result of a person’s moral identity

---

<sup>2</sup> Albert Bandura, “Social Cognitive Theory of Personality,” *The Coherence of Personality: Social Cognitive Bases of Consistency, Variability, and Organization*. Ed. Daniel Cervone and Yuichi Shoda. (New York: The Guilford Press, 1999).

becoming less accessible could be “good people doing bad things.” According to Luban, the motivation for changing one’s identity and rationalizing the resulting unethical conduct is the quest for integrity.

In this section I review Luban’s arguments for why the quest for integrity can be dangerous. I respond that Luban applies his notion of integrity too narrowly and does not consider how ID-commitments can be difficult to change and that the quest for integrity is not necessary for a person to rationalize her unethical conduct. Instead, his concerns about the quest for integrity mostly apply to people who are more committed to pleasure and advancement than to moral values. I also use one of Luban’s main examples of self-deception to argue that it actually demonstrates how the quest for integrity, properly understood, is actually a quest to be true to an authentic moral self in the face of temptation.

Luban begins by defining integrity as inner consistency and harmony, using the notion of integrity as wholeness (Latin *integrare* - “to make whole”). On his account, harmony is specifically between one’s beliefs and ones’ actions - “a textbook definition of integrity.”<sup>3</sup> Luban correctly characterizes this definition because many philosophers and business ethicists also take integrity to mean mere consistency between commitments, words, and actions (see Chapter 1). The moral identity account of integrity also captures this aspect in the first necessary condition that requires coherence between one’s commitments and actions. Luban also argues, much as I do, that the motivation for self-consistency moves some people to act in accordance with their beliefs and

---

<sup>3</sup> David Luban, *Legal Ethics and Human Dignity* (New York: Cambridge University Press, 2007), 269.

commitments. He does not, however, posit a general desire for self-consistency but focuses specifically on the desire for cognitive dissonance reduction.

Tavris and Aronson define cognitive dissonance as the “state of tension that occurs whenever a person holds two cognitions (ideas, attitudes, beliefs, opinions) that are psychologically inconsistent, such as ‘Smoking is a dumb thing to do because it could kill me’ and ‘I smoke two packs a day.’”<sup>4</sup> Multiple studies suggest that dissonance between our beliefs and conduct can create great discomfort which we can reduce by either changing our beliefs (e.g. “I need to smoke to help me relax.”) or our conduct (e.g. flush *all* cigarettes down the toilet). When it comes to immoral conduct, Luban designates two “roads” of integrity that match these two cognitive dissonance reduction methods. The *low road* is taken by those who change their beliefs to rationalize immoral conduct and the *high road* is taken by those who change their conduct to match their beliefs.

From the moral identity perspective, the low road could be a serious concern for people pursuing substantial integrity because it implies that the motivation of self-consistency is more than the disposition to be true to oneself or to act on one’s ID-commitments. On the low road, a person is motivated to reduce cognitive dissonance which results in changing her beliefs and commitments to match her conduct. Another product of this change in beliefs is rationalization, self-deception, and an increase in unethical conduct. The high road may also be a concern for substantial integrity because Luban argues that we become confused about what moral values actually apply to

---

<sup>4</sup> Tavris and Aronson, 2007, 13.

particular situations, especially when we are immersed in an unethical profession or workplace. In response to these concerns, I first discuss the low road of integrity and then the high road.

### **1.1 - The Low Road of Integrity**

Luban describes the low road as the one we take when fall prey to cognitive dissonance and unconsciously change our beliefs to conform to our conduct. Thousands of experiments have demonstrated the effect of cognitive dissonance and how it can be resolved by changing one's beliefs. In the classic cognitive dissonance experiment, subjects complete a tedious and boring task. Afterwards, some subjects are paid a \$1 to advocate for a belief they did not hold (e.g. "Tell the next subject that the tedious task was enjoyable."). By the time they leave the experiment, these subjects tend to believe that the experiment was enjoyable. Other subjects who are offered \$20 to advocate the same belief tend to leave the experiment believing that the task was tedious.<sup>5</sup> The most common explanation for the first group's change in belief is that the \$1 inducement was not enough to justify their advocating behavior so they unconsciously changed their beliefs to match their behaviors. The subjects who were promised \$20 did not change their beliefs because the large payment was enough to justify their behavior.

Luban associates the quest for integrity along the low road with cognitive dissonance reduction. Reducing dissonance by changing beliefs, however, is not enough to make the quest for integrity dangerous. We also desire to appear morally upright to ourselves and others. Because immoral conduct in particular threatens to undermine our

---

<sup>5</sup> The classic experiment is found in Leon Festinger and J. M. Carlsmith, "Cognitive Consequences of Forced Compliance," *Journal of Abnormal and Social Psychology* 58, (1959).

upright self-image, we may unconsciously change our beliefs to protect our moral self-assessment. In support of this claim, Luban quotes psychologists who describe humans as *intuitive lawyers* who are constantly defending their righteous self-image.<sup>6</sup> He concludes that the unconscious desire to reduce cognitive dissonance and maintain a positive view of ourselves drives us to compromise our beliefs to fit our immoral conduct.

Luban argues further that the quest for integrity can produce a recursive cycle of rationalization and commitment. He states, “One consequence of dissonance theory is that once I act, my beliefs will rationalize the action and therefore impel me to further action of the same sort — which, in turn, calls for renewed rationalization, and further action. Action, we might say, breeds *commitment*, and commitment breeds further action in an ever steeper slippery slope.”<sup>7</sup> The quest for integrity eventually results in a progressive moral downfall. Described this way, a person pursuing substantial integrity may unknowingly accelerate her immoral behavior.

According to Luban, the cyclical nature of dissonance reduction is especially common among professionals whose role expectations and organizational goals conflict with moral principles. Instead of rejecting moral principles outright, the person deceives herself into thinking she can live by two conflicting moral codes and still be morally upright. Luban explains that, “dissonance theory predicts that I will preserve my conception of myself as a morally upright individual in the only way left: by abandoning

---

<sup>6</sup> Luban, 2007, 270.

<sup>7</sup> Luban, 2007, 273.



the belief that my other beliefs should be consistent.”<sup>8</sup> The quest for integrity drives us to subordinate or change our moral principles while not acknowledging that these changes are making us *less* upright. The result is self-deception. We hold inconsistent beliefs but deceive ourselves into thinking they are consistent so as to reduce cognitive dissonance and justify immoral conduct. Luban concludes that, “The quest for integrity kills, and in killing it leaves the survivors with their own sense of rectitude intact, like a tattered flag flapping in the wind over the fallen.”<sup>9</sup>

Luban’s cognitive dissonance account of integrity, if correct, would mean that some people who pursue substantial integrity are in danger of changing or subordinating their moral values in pursuit of inner harmony and an upright self-image. While Luban’s account describes common psychological processes such as cognitive dissonance reduction and a desire to justify ones’ actions, I argue that he fails to pin these processes to the quest for integrity and especially the quest for substantial integrity. In evaluating his arguments, I explain why the moral identity account can also explain Luban’s concerns and how the quest for integrity is not necessary for belief change, rationalization, and self-deception.

Before evaluating Luban’s actual arguments, I need to highlight two concerns I have with how Luban has constructed his overall argument. First, Luban assumes that the person pursuing integrity has already behaved immorally and will continue to do so. As presented, the cognitive dissonance process begins with immoral conduct that does not fit

---

<sup>8</sup> Luban, 2007, 285.

<sup>9</sup> Luban, 2007, 285.

one's moral beliefs. By introducing the quest for integrity using a person who has a tendency to act immorally, Luban tilts the discussion on a downward trajectory which makes the quest for integrity appear like a slippery slope. If cognitive dissonance reduction is a strong motivation, then it seems that a person could also do good deeds and reconcile her immoral beliefs to match her conduct or make a mistake and return to her moral commitments. In this case the quest for integrity could produce a wealth of moral behavior and moral commitments—a moral ski lift perhaps. Even though Luban's main objective is to explain the slippery slope that lawyers often confront, his account needs to consider those who have a different moral trajectory and who avoid the slippery slope from the start.

My second concern is more central to Luban's claims. He is unclear about *what* we actually change to reduce cognitive dissonance. He says that integrity is a quest for inner harmony, but he is not clear on what is harmonized. In different places he states that harmony can be between one's beliefs and actions, principles and actions, and/or commitments and actions.<sup>10</sup> In general, Luban focuses on beliefs and how internal mechanisms encourage us to alter them. He further argues that continuous exposure to certain environments can change one's beliefs into permanent commitments and principles. For example, a person may develop a "professional morality" that justifies immoral actions at work and eventually becomes a fixed part of her moral identity.

---

<sup>10</sup> Contrast, "If your conduct conflicts with your principles, modify your principles." (268); "But the net effect is a happy harmony between what I do and what I believe -- the textbook definition of integrity." (269); "Action, we might say, breeds commitment, and commitment breeds further action in an ever steeper slipper slope." (273).

My concern with Luban's handling of these concepts is that he does not consider the marked differences between having beliefs, principles, and commitments. He argues that we alter beliefs over time and they become commitments in our quest for integrity, and in the early development of one's ID-commitments this may be true. What he ignores is that people may already have deep ID-commitments to moral values. For example, Luban highlights the obedient subjects who compromise their moral principles during the Milgram obedience experiments (see Chapter 4), but he does not discuss the defiant subjects. He depicts those on a quest for integrity as wanting to change their already malleable beliefs, but this is not the case especially for moral exemplars. As I have shown with numerous examples in earlier chapters, people vary in their level of commitment to moral principles.

Luban may not need a strict distinction between beliefs and commitments because he is not giving an account of integrity, but rather an account of the *quest for integrity*. His main purpose is to explain how the *quest* or motivation for integrity produces problems, not how integrity is a problem itself. His chapter is even titled, "Integrity: Its Causes and Cures" which emphasizes that the causes of integrity are his main concern. But by focusing on the main causes of cognitive dissonance reduction and rationalization, he ignores a common claim that integrity itself requires ID-commitments or unconditional commitments. In examining some very real causes for rationalization and self-deception, he does not consider the level of commitment with which we enter each situation.

Because Luban does not consider different levels of commitment, he does not address those who have a high level of commitment to moral values such as John Weidner or Socrates. In Chapter 1, I described different levels of commitment. As a reminder, a Level 1 commitment may be changed easily, but Level 2 or 3 commitments are held so deeply that some may die before they compromise them. It appears that Luban is focusing his critique on people with Level 1 and possibly a weak Level 2 commitment to moral values. If we follow his description of the low road of integrity, his main concern should be those with a low level of commitment and who easily change their beliefs or who have their commitments co-opted for the sake of internal harmony. Meanwhile, his worries about the quest for integrity may not apply to those who have Level 2 and 3 ID-commitments to moral values.

Another possibility is that Luban is targeting people who do have high level ID-commitments but to something besides moral values. Barry Schlenker et. al.'s research is helpful at this point in distinguishing between people who have ID-commitments to different ultimate ends. As discussed in Chapter 4, Schlenker et. al. found that some people have a *principled ideology* which means they have strong commitments to moral principles and constantly self-regulate their conduct. Luban may be concerned for these people of substantial integrity because they may change their beliefs about morality to match their unethical actions. But some research shows that these people tend to take the high road of integrity by changing their actions to fit their ID-commitments to moral values. If this is the case, then Luban's target is those people who do immoral things and then justify their actions. On the moral identity account, these people act like they do

because they either have little access to their moral identity (if they have one) and/or they live in accordance a non- or immoral identity. They may live by what Schlenker et. al. call an *expedient ideology* which means they recognize moral principles but discard them when they seem too rigid or when acting on them would cause them to miss opportunities that may further their advantage. In line with Luban's description, Schlenker et. al. have found that those who are committed to an expedient ideology *do* tend to justify and rationalize deviations from principles more than those with a principled ideology.<sup>11</sup>

By reframing Luban's project in terms of commitments and moral identity, I can better evaluate his claims regarding the quest for integrity and develop a more accurate picture of the topic. I contend that if Luban is primarily concerned with people who have expedient ideologies, then his arguments may fit his target audience of lawyers but not the moral exemplars discussed in earlier chapters. Luban may insist that his description of the quest for integrity still applies to people of substantial integrity who have a principled ideology, but I disagree. In light of the moral identity account, the quest for substantial integrity on the low road is not the best explanation for why people rationalize and perpetuate immoral conduct. I have two reasons for contending that the dangers Luban describes along the low road of integrity should not be a great concern for a person of substantial integrity.

The first reason is the implausibility of people in general changing core moral beliefs and commitments to match their conduct. Luban uses as an example the obedient

---

<sup>11</sup> Barry R. Schlenker, Marisa L. Miller, and Ryan M. Johnson, "Moral Identity, Integrity, and Personal Responsibility" in *Personality, Identity, and Character: Explorations in Moral Psychology*. Eds. Darcia Narvaez and Daniel K. Lapsley. (New York: Cambridge University Press, 2009).

Milgram subjects who supposedly justified each progressive shock by changing their beliefs about the morality of the previous shock. But did they change their core beliefs and principles? The after-experiment interviews with subjects indicate that most of them did not leave the experiment with different moral beliefs and commitments. Even Luban concludes that dissonance-induced belief can be ephemeral because when people leave the situation they take on their “genuine long-term moral and personal commitments.”<sup>12</sup> If this is the case, then we must consider another explanation besides a change in belief or commitment that causes us to rationalize and escalate immoral conduct.<sup>13</sup>

If beliefs and commitments regarding morality do not change easily, a change in conduct could be explained by a lack of moral identity accessibility in certain contexts and/or an alternative identity guided by an expedient ideology. Luban comes close to drawing the same conclusion regarding professional roles. He states that over time we can develop a professional identity with its own professional morality that is “distinct from the morality of your extra-professional life, justifies what you do—and this belief will be no transitory thing, but rather a fixed part of your moral personality.”<sup>14</sup> But the creation of another morality and changing one’s commitments does not require the quest

---

<sup>12</sup> Luban, 2007, 288.

<sup>13</sup> Luban could note the apparent change in the moral values of German citizens that seemed to accompany Hitler’s rise to power. I would respond that like the Milgram bystander experiments in which obedience reached over 90%, an ID-commitment to not personally harm others appears to be quite different from an ID-commitment to stop someone else from harming others or to sacrifice one’s life so that one will not harm another person. Also, a change of moral values such as honesty and keeping one’s word did not necessarily change, but the scope of *to whom* these values applied did change for many. A final note on this particular example. According to historian Paul Johnson, Hitler was wildly popular with his people until he invaded Czechoslovakia in March 1939. After this time he ruled not by consent but by force and fear, a macrocosm of the Milgram experiments. Paul Johnson, *Churchill*. (New York: Penguin Group, 2009).

<sup>14</sup> Luban, 2007, 289.

for integrity or the reduction of cognitive dissonance. Instead, it seems that a person can learn to live according to a professional identity and a moral identity because both are useful when living by an expedient ideology. As a reminder, the expedient ideology respects moral principles but does not require them when they are impractical or impede one's advantage. Luban again tends to agree with this description: "Nor will this dualistic view of morality bother you [as the agent]. You will effortlessly negotiate the transition from one form of life to the other with no sense of tension or contradiction."<sup>15</sup> The reason no contradiction exists is because morality is always subordinate to one's pursuit of subjective goods.

If we can account for Luban's concerns about changing one's moral beliefs and principles without referring to a quest for integrity, then how does the quest become a danger? The answer may be rationalization and self-deception. Luban says that the quest for integrity can *deceive* us into thinking that our lives are harmonious and upright even though we are living through two or more identities with conflicting moral principles.

In a famous example, Montaigne appears to claim that he is able to keep his position as Mayor of Bordeaux separate from his identity. He states, "I have been able to take part in public office without departing one nail's breadth from myself and to give myself to others without taking myself from myself."<sup>16</sup> Luban argues that Montaigne tries to avoid the reality that he is actually himself *and* the mayor. Montaigne, like many lawyers, is fooling himself because he does not recognize that the quest for integrity and

---

<sup>15</sup> Luban, 2007, 289.

<sup>16</sup> Michel de Montaigne, *The Complete Essays of Montaigne*, Trans. Donald M. Frame. (California: Stanford University Press, 1958), 770.

its “evil twin” of rationalizing immoral conduct are almost impossible to distinguish. The quest for integrity and with the drive to appear morally upright deceive us into thinking we have integrity when we are actually living with conflicting identities.

The dangers of rationalization and self-deception bring me to the second reason why the dangers of the low road should not greatly concern a person of substantial integrity. Against Luban, I claim that the quest for integrity defined as coherence is *not necessary* to produce the rationalization and self-deception that make integrity dangerous. All that is necessary is the desire to avoid the discomfort of cognitive dissonance and the desire to appear upright. If one consciously or unconsciously aims at these all-too-common subjective goods, one will rationalize immoral conduct, deceive oneself about being a moral person, and may even be encouraged to repeat this process. If this is correct, the quest for substantive integrity is not necessary to produce rationalization and self-deception.

Consider an example from Luban’s target audience. When a lawyer decides to make a jury believe something that is false in order to win a case, the quest for integrity may not cause her to change her moral principles and rationalize her conduct. According to the moral identity account, she may be bending moral principles because she has an ID-commitment to a professional identity that is more central and accessible than her moral identity. Also, she may justify her conduct to reduce the discomfort of cognitive dissonance, but the quest for integrity is not necessary for her to do this. Following Schlenker’s research, I suggest that a primary motivation for the lawyer discounting her moral principles is her expedient ideology. If this is correct, her actions and



rationalizations may be chosen for the sake of receiving subjective goods such as pleasure and advantage. These could take the form of making partner, benefiting from large settlements, fitting in with other lawyers, or merely achieving a career goal.

Regarding the problem of self-deception, Luban may be correct in claiming that the lawyer deceives herself into thinking that she can lie at work and be honest at home, but the quest for integrity is not necessary to create this self-deception. Self-deception is most likely caused by what Luban calls integrity's "evil twin"; the desire to appear upright and avoid responsibility for her actions. I suggest and argue below that this type of self-deception can be produced by the quest for moral impunity, and does not require the quest for a harmony between professional and moral identities. Even Luban admits that the quest for integrity and rationalization are difficult to distinguish. I suggest that the reason they are indistinguishable is because the quest for integrity is not necessary in these "low road" cases.

Rationalization, unlike the quest for integrity, works primarily at the level of construing one's actions in a way that justifies them as moral. Luban may be correct that the result of rationalizing is that the lawyer deceives herself. She truly believes that she did not do anything wrong and she is not responsible for any harm done. What is necessary for rationalization in this case is an expedient ideology that seeks to avoid responsibility and shame. In Chapter 4 I explained that rationalization is a form of *moral disengagement* and that initial research suggests that people with a principled ideology tend to use it less than those with an expedient ideology.<sup>17</sup> Moral disengagement

---

<sup>17</sup> Schlenker et. al., 2009, 330.

describes the tendency for some people to distance themselves from their own immoral behavior and to avoid responsibility for their actions or inactions. In what follows I describe a few moral disengagement practices and show that they clearly match the deception process Luban associates with the quest for integrity.

Bandura identifies several disengagement practices for “reconciling perturbing disparities between personal standards and conduct” for the purposes of reducing a loss of self-respect.<sup>18</sup> On Luban’s account each moral disengagement practice is driven by the quest for integrity on the low road. While cognitive dissonance reduction may motivate a person to deceive herself in an effort to appear upright, it does not follow that the person of substantial integrity will be similarly motivated. Instead, a person with an expedient ideology may be thus motivated or they may be motivated to reduce a loss of self-respect as Bandura describes it. His description fits well with my claim that rationalization and self-deception most likely come from an expedient ideology that prioritizes pleasure and personal advantage over moral values.

Bandura classifies moral disengagement practices by those that center on the conduct itself, the consequences of the actions, and the victims. Below I describe three common moral disengagement practices that reveal a quest for pleasure and advantage more than a quest for integrity as harmony. The practices are moral justification, misconstruing the consequences, and blaming the victim. All of these practices describe the actions of the self-deceived people who concern Luban.

---

<sup>18</sup> Bandura, 1999, 216.

*Moral justification* describes a willingness to justify or excuse antisocial behavior so that a person is less condemned by herself and others. It is often used to portray actions so that they appear to be serving some moral purpose. For example, mortgage brokers who sell high-interest rate mortgages to unqualified buyers sometimes justify their actions by saying, “We are helping our customers achieve the American dream of home ownership.” In actuality they are condemning customers to future foreclosure. In this case it would be strange to say that the quest for integrity caused the brokers to rationalize their behavior and deceive themselves. A more accurate description is that they are following an expedient ideology (i.e. “Moral rules don’t apply in business”) and use moral justification to avoid taking responsibility or feeling shame for their actions.

A disengagement practice often used by people in organizations is *diffusion of responsibility* (see Chapter 4). This practice centers on “obscuring or minimizing the agentive role in the harm one causes.”<sup>19</sup> Blaming an authority figure or saying that “all mortgage brokers in my firm do this” are excuses that morally disengage a person from taking responsibility for her actions. Some people morally disengage by disregarding or misrepresenting the harm they caused. They minimize, explain away, hide, or simply disbelieve the consequences of their conduct which allows them to be free of responsibility. Diffusion of responsibility is a practice that does not require a quest for integrity because its purpose is to avoid feeling responsible for one’s immoral actions, not necessarily the harmonizing of one’s beliefs and conduct.

---

<sup>19</sup> Bandura, 1999, 218.

A final moral disengagement practice of note is *blaming the victim* for one's actions and *dehumanizing* them. For example, some obedient subjects in the Milgram experiments blamed the learner for not doing his part.<sup>20</sup> People may also say that they were provoked into doing an immoral action and thereby are not responsible for reacting with an immoral action. Dehumanizing those who are harmed is another way to use the victim to morally disengage from one's actions. Bandura states, "Once dehumanized, they are no longer viewed as persons with feelings, hopes and concerns but as subhuman objects."<sup>21</sup> In his work on ordinary people who commit genocide, James Waller states that, "Regarding victims as outside our universe of moral obligation and, therefore, not deserving of compassionate treatment removes normal moral restraints against aggression."<sup>22</sup> Why would a person use this moral disengagement practice? Luban may argue that she would employ it in her quest for integrity. I would disagree because a more accurate description is that the perpetrator wants to avoid feelings of guilt and shame for treating others poorly. It is true that the result of her quest to avoid pain and shame may be a feeling of harmony and dissonance reduction, but we need not posit the quest for integrity as the primary driver behind her moral disengagement.

To different degrees we all use practices like these to morally disengage from our conduct, its consequences, and its victims. Even children, who I would not consider people who are on a quest for integrity, use these practices because they want to avoid

---

<sup>20</sup> Stanley Milgram, *Obedience to Authority: An Experimental View* (New York: Harper and Row Publishers, Inc, 1974): 204.

<sup>21</sup> Bandura, 1999, 218.

<sup>22</sup> James Waller, *Becoming Evil: How Ordinary People Commit Genocide and Mass Killing*. (New York: Oxford University Press, 2007). 207.

responsibility and shame for their conduct. Adults can be the same in this regard because we want to avoid sanctioning ourselves and losing our self respect.

To summarize, the quest for integrity is not necessary for changing beliefs, rationalizing behavior, and deceiving oneself. If my argument is correct, people of substantial integrity need to be greatly concerned with Luban's arguments. When we consider moral disengagement practices, the motivation for rationalization and self-deception can be the subjective good of pleasure (e.g. feel no cognitive dissonance, shame, or responsibility) and advantage (e.g. feel morally upright, career success) and not necessarily the quest for integrity as coherence. In other words, the people with which Luban is concerned seem to be those who have limited access to their moral identity and/or have constant access to an identity that lives by an expedient ideology. Even Luban implies that the people taking the low road may take on an alternative (e.g. professional) identity with its own morality.<sup>23</sup> I have argued that all a person needs to rationalize her immoral conduct is a desire for subjective goods such as pleasure and advantage which includes avoiding responsibility and self-sanctions. The quest for integrity is also not necessary to deceive ourselves about holding inconsistent principles. Often there is no deception because either the person has embraced the expedient ideology and is convinced that moral principles simply don't apply in certain domains, or she morally disengages from her conduct to avoid responsibility and shame.

Even though I conclude that the quest for integrity as harmony and coherence is not necessary for the belief change, rationalization, and self-deception that concerns

---

<sup>23</sup> Luban, 2007, 291.

Luban, I do acknowledge that cognitive dissonance reduction along with a desire for moral justification could produce these results. I have also argued that a person of substantive integrity should not be *greatly concerned* about the quest for integrity as coherence on the low road. My reason for making this claim is that to be a person of substantive integrity like Weidner and Socrates, one must have coherent *ID-commitments* to moral values. This means that the person is not likely to change these ID-commitments easily or based on situational factors. Also, their moral identity is continually accessible which means that moral values are not easily ignored. But note that I say that a person of substantive integrity should not be *greatly concerned*. I believe they should be somewhat concerned because I recognize the power of situational factors as well as cognitive dissonance reduction to shake a person's ID-commitments at times, and in section 2 I address strategies to maintain one's integrity.

Luban could object and argue that the quest for integrity may not be *necessary* but it is *sufficient* to produce the rationalization and self-deception that he describes. His objection could take two forms. The first form of the objection is that we see obedient subjects in the Milgram experiments rationalizing their behaviors and deceiving themselves about their self-conceptions. The quest for integrity on the low road, defined as harmonizing one's beliefs and conduct, is sufficient for this rationalization and deception process. The second form of the objection is that a person may actually have dual moralities that she believes are consistent when they actually aren't. She may even say that she is a "person of integrity" because she sees no inconsistency in lying at work

and telling the truth at home. I address the first form of the objection at this time and the second form after I discuss Luban's account of the high road of integrity.

Luban could contend that the quest for integrity is sufficient to motivate a person to rationalize her behavior and deceive herself. Recalling how obedient Milgram subjects resolved their cognitive dissonance by harmonizing their beliefs with their conduct, Luban argues that their rationalizations should not surprise us. "After all, if some of my roles impose inconsistent moral demands . . . and my daily life leads me to occupy all these roles and if, further, the actions I take in each role lead me to adopt beliefs that vindicate these actions, then dissonance theory predicts that I will preserve my *conception of myself* as the morally upright individual in the only way left: by abandoning the belief that my other beliefs should be consistent."<sup>24</sup> He concludes that seeking integrity produces self-deception and the denial of inconsistencies in our beliefs.

I can agree with Luban that the quest for integrity narrowly defined as coherence between beliefs and actions may be sufficient to produce self-deception. But I would point out that Luban's narrow definition of integrity is best described as mere cognitive dissonance reduction. The subjects to which Luban refers in the Milgram experiments probably desired internal consistency and not what is commonly called integrity. To illustrate my point, all one needs to do is think of a subject who defied the experimenter's commands and a subject who shocked the learner to the end during the same experiment. Then answer the question, "Which of these subjects is most likely a person of integrity?" It is my experience that most people would easily choose the defiant subject because

---

<sup>24</sup> Luban, 2007, 285. Emphasis mine.

integrity strongly implies moral trustworthiness. Anyone who says “both” subjects will have to tell a complex story about why they both have integrity. But if I accept Luban’s narrow definition of integrity, then it could be sufficient to produce rationalization and self-deception in some people.

In light of this definition, however, Luban must answer a critical question about integrity. If the quest for integrity is sufficient to explain why people change their beliefs and deceive themselves, how would he describe the quest of defiant Milgram subjects who did not change their beliefs and did not deceive themselves? I argue later that in these subjects we find examples of people who are on an *actual* quest for integrity (i.e. substantial integrity) and not a quest for internal harmony. They understood that they were responsible for the actions they took and realized that they had done something wrong. Several of them acknowledged that their initial conduct did not fit with their values and that they learned that they must not blindly follow orders. If Luban claims that the quest for integrity is sufficient to produce rationalization and self-deception, then he needs to explain why it was also sufficient for defiant subjects to avoid rationalizing their conduct and take full responsibility for their actions. If only the obedient subjects were on a quest for integrity, then it appears that the quest for integrity as Luban defines it is best described as rationalizing bad conduct to reduce the discomfort of cognitive dissonance – not the quest for integrity.

The second and more significant form of Luban’s objection could be that the quest for integrity is sufficient in the cases of professionals who develop separate professional and personal moralities and deceive themselves about their inconsistent



moral standards. They may even claim that they have integrity while maintaining these two moralities. I take up this objection after considering Luban's high road of integrity argument.

### **1.2 - The High Road of Integrity**

The moral identity account of integrity is far different from Luban's account of integrity as harmony and coherence because having substantive integrity requires ID-commitments to moral values. Luban acknowledges that common usage of "integrity" does not come close to fitting his low road distinction and he imagines an objector telling him, "You are not talking about genuine integrity." In response to this objection he develops the high road account of integrity which involves conforming one's conduct to one's beliefs and principles, not the other way around. Luban admits that integrity can mean being uncorrupted or untouched by sticking to one's moral commitments even when under pressure to compromise.

Luban's concern with the high road is that it is plagued by subjective perspectives about what is morally right and wrong. He notes that dogmatically sticking to one's moral standards may not reveal integrity because we all change our perspectives on right and wrong as we mature and gain experience. Given our changing moral principles, how can we tell from within ourselves which principles are actually right and wrong? Luban argues that situational and psychological uncertainty affects our judgment regarding which moral principles apply in different situations. We only know what is "right and reasonable" to ourselves and this may have been corrupted by our desire to appear morally upright. The result is a moral compass with no true north.

I would agree with Luban's general argument if it were not for the actions of moral exemplars and even some defiant Milgram subjects who seem to have constant access to their moral identities. Luban's contention may be correct concerning the majority of people who self-select into professions and roles that tend to justify immoral conduct, but I have argued earlier that those who have an expedient ideology as opposed to a principled ideology lack access to their moral identities anyway. They are not confused about morality but quite clear about how to abandon it for the sake of subjective goods. Another possibility is that these people have not formed a moral identity in the first place and are those with low moral identity centrality (LMID). As demonstrated in the moral identity experiments in Chapter 4, people with LMID tended to compromise moral values more often than those with HMID.

For Luban's analysis to be correct, he must show that in general people cannot tell right from wrong on basic issues, but this is a dead end. Most adults across cultures understand and live according to basic moral values: tell the truth, do not harm others, keep your word, and help those in distress when it is not too costly. The problem Luban seems to be tracking is that people who develop a professional identity and have an expedient ideology too willingly compromise moral values for the subjective goods of pleasure and advantage. They may be the ones who are highly susceptible to being confused regarding what is right and wrong.

As discussed above, Luban admits that he cannot claim moral uncertainty is prevalent because people do not change their beliefs about right and wrong easily. He is concerned for those people who must spend 2,400 hours a year in a job that reinforces

moral standards that do not match those held by the human community. Eventually, the person may commit herself to a professional identity and an expedient ideology which permits her to rationalize her immoral conduct. Luban argues that from the inside this person believes that she has integrity even though she holds to contradictory moral principles. He states, “We would like our moral compass to point north, but our only instrument for detecting north is our moral compass. . . the *quest* for integrity can drive us to the high road or the low road, without any landmarks to alert us about which path we have taken.”<sup>25</sup> Luban’s problem with the high road is that people will not realize that they hold inconsistent moral principles because they will not be able to distinguish between them.

I have a similar concern regarding those who appear to hold contradictory moral principles, but I do not agree that it is a problem of not being able to distinguish right from wrong along the high road of integrity. I suggest that the problem is one of compromise, not knowledge. If we consider moral exemplars, we see people who have ID-commitments to moral values that are integrated into their self-conceptions. Their moral identities are central and constantly accessible. As moral exemplar researchers Anne Colby and William Damon have found, “individuals who define the self in terms of their moral goals will more likely interpret events as moral problems and will more likely

---

<sup>25</sup> Luban, 2007, 291.

see themselves implicated in the solutions to these problems.”<sup>26</sup> A person of substantive integrity recognizes when moral values are threatened and refuses to compromise.

When we consider the duplicitous people in Luban’s examples, the moral identity account would describe them as having a highly accessible identity that is guided by an expedient ideology towards the subjective goods of pleasure and advantage. They do know moral principles and practice them in parts of their lives, but they are willing to compromise these principles for subjective goods. As my evidence, consider the actual differences between those who have a constantly accessible and highly centralized moral identity and those without. As discussed in Chapter 4, Schlenker and other moral identity researchers have found evidence that people with ID-commitments to moral values do not easily dismiss or change their moral principles even when encouraged to do so and they tend to live out these principles across situations. People with ID-commitments to an expedient ideology, however, believe moral values as important but too rigid and impractical for all occasions.<sup>27</sup>

Luban argues that practicing a profession for 2,400 hours a year can change a person so that she is unable to perceive moral landmarks. I agree that this could happen, but we cannot forget that the employee must allow this culture to slowly compromise her moral commitments. On the moral identity account, the inability to perceive moral landmarks comes from moral identity inaccessibility and moral disengagement, not the

---

<sup>26</sup> Anne Colby and William Damon, “The Uniting of Self and Morality in the Development of Extraordinary Moral Commitment.” *The Moral Self*. Eds. Gil G. Noam and Thomas E. Wren. (Massachusetts: MIT Press, 1993): 152.

<sup>27</sup> See Barry R. Schlenker, Marisa L. Miller, and Ryan M. Johnson, “Moral Identity, Integrity, and Personal Responsibility” in *Personality, Identity, and Character: Explorations in Moral Psychology*. Eds. Darcia Narvaez and Daniel K. Lapsley. (New York: Cambridge University Press, 2009).

quest for integrity on the high road. Take as an example two lawyers. Lawyer A is in a law firm that is proud that they rarely lose a case. In her case, Lawyer A views a particular witness' personal life as fair game for discrediting her testimony. Lawyer A exaggerates the witness's personal failures to win the case. On the moral identity account, Lawyer A may not perceive moral landmarks because she works from a professional identity that is guided by an expedient ideology. In other words, she ignores moral values at work because they are impractical for the court room. If she has a moral identity, it is rarely accessible. However when she is teaching her child about being honesty, she is clear about its importance as a moral value because her moral identity is accessible at home.

In the exact same circumstances, imagine Lawyer B looking through the same witness' personal life and respectfully weighing what she will and will not bring up in trial. To be true to her moral identity, she will not overly exaggerate the witnesses mistakes to win a case because she knows to do so would be immoral and contrary to her deepest commitments. Instead, she may conduct more research and develop creative trial methods to win the case without compromising her moral identity. Even under pressure to win the case, Lawyer B can perceive moral landmarks because of her moral identity accessibility.

On the job, moral disengagement practices may also cloud Lawyer A's perception of right and wrong. She may justify her actions beforehand by arguing, "I must vigorously defend my client" or she may blame the witness for "being at the wrong place at the wrong time." Lawyer B, however, does not need to resort to these practices

because she is not going to unnecessarily harm the witness for her or her client's own ends. She sees her moral landmarks clearly because of her ID-commitment to moral values. Also, if Lawyer B daily reminds herself of the moral principles to which she has an ID-commitment, she will be less likely to lose sight of her moral landmarks. In other words, the high road of integrity is covered in shadow for those who have an inaccessible or undeveloped moral identity as well as an expedient ideology. It is not that these people are genuinely confused about right and wrong; I would suggest that they have chosen to ignore it for the sake of other commitments and subjective ends.

An excellent example of a lawyer who was clear about moral landmarks even under pressure to compromise is a young John Adams. Early in his legal career, Adams was asked to defend British soldiers who had fired on a group of citizens in Boston. Soon after the altercation it was hyped as the "Boston Massacre" and public opinion was against Adams taking the case. According to historian David McCullough, Adams believed that every person in a free country deserves the right to counsel and a fair trial. He took the case on principle.<sup>28</sup> Adams did worry about damaging his reputation by defending the unpopular British captain and his soldiers. However he steeled himself by writing in his journal often. In one entry he copied the following from an Italian opponent of the death penalty: "If, by supporting the rights of mankind, and of invincible truth, I shall contribute to save from the agonies of death one unfortunate victim of tyranny, or of

---

<sup>28</sup> David McCullough, *John Adams* (New York: Simon and Schuster, 2002), 66.

ignorance, equally fatal, his blessings and years of transport will be sufficient consolation to me for the contempt of all mankind.”<sup>29</sup>

Adams won the case for the Captain and for six of the eight soldiers. Two were convicted of manslaughter and had their thumbs branded. The victory was bitter-sweet for Adams as he claims that he lost half of his legal practice for taking the case. His relative Samuel Adams, however, did not seem to disapprove of him taking the case. McCullough speculates that Samuel Adams may have even encouraged it “out of respect for John’s fierce integrity” and that this may play well in the future for John’s political career.<sup>30</sup>

Adams’ example again brings me to a missing aspect of Luban’s account of the quest for integrity that I mentioned much earlier. If I accept that the quest for integrity can be defined as cognitive dissonance reduction coupled with a desire to appear upright, and I also accept that the quest can lead some people down a slippery slope of rationalization and self-deception, then why can’t the quest for integrity be turned up the slope as well? It seems that cognitive dissonance reduction combined with a desire to appear upright could lead people in the other direction – possibly to the peak of ethical conduct. The quest for integrity in this positive direction, especially for lawyers at a highly ethical law firm, would also clarify moral landmarks and make them bedrock supports for all cases taken and decisions made. In other words, it seems that the same drives that lead down a slippery slope could help a lawyer who does good deeds to

---

<sup>29</sup> McCullough, 2002, 67.

<sup>30</sup> McCullough, 2002, 68.

change her immoral beliefs to match her conduct. Even though Luban's main objective is to clarify the moral slippery slope that lawyers often confront, his account of integrity should also consider those with a different moral trajectory who actually take the high road by changing their conduct to match their beliefs.

Regardless of the direction of a person's moral trajectory, Luban can still make his case that both the low road and the high road of integrity produce self-deception, and that self-deception in either case is harmful. The desire to harmonize one's actions and beliefs may drive a person to deceive herself regarding the actual state of her self-conception. At this point the second form of Luban's sufficiency objection comes into play. He could argue that the quest for integrity may not be necessary but it is sufficient for self-deception. Particularly in certain professions such as law and politics, the quest for integrity is sufficient to deceive a person into thinking that her professional and personal moral standards are consistent. She may even claim that she has integrity even though her actions seem to deny that fact. In this regard, I discuss an example of self-deception that is prominent in the integrity literature: Michel de Montaigne.

### **1.3 - Montaigne and Integrity**

Sixteenth century essayist and politician Montaigne worked for thirteen years as a counselor for the Parliament of Bordeaux in the late 1500's. He found the laws he was administering unjust and he attempted to retire. Unfortunately for Montaigne, he was obliged to be a two-term Mayor of Bordeaux. He considered the government a diseased



and worm-eaten body of which he was the least diseased member.<sup>31</sup> He lamented that public offices often require vices that are in direct tension with virtue. Montaigne states, “Likewise in every government there are necessary offices which are not only abject, but also vicious. Vices find their place in it and are employed for sewing our society together, as are poisons for the preservation of our health. . . The public welfare requires that a man betray and lie and massacre . . .”<sup>32</sup>

Montaigne must live and work in this realm while at the same time maintaining his virtue. Luban assumes that he practices the vices required by his position. A few paragraphs of his writings have been used to argue that Montaigne deceives himself into thinking that he can be a vicious Mayor and a moral person at the same time. I quote parts of these paragraphs below.

I have been able to take part in public office without departing one nail’s breadth from myself, and to give myself to others without taking myself from myself.

Most of our occupations are low comedy. *The whole world plays a part* [Petronius]. We must play our part duly, but as the part of a borrowed character. Of the mask and appearance we must not make a real essence, nor of what is foreign what is our very own. . . It is enough to make up our face without making up our heart. I see some who transform and transubstantiate themselves into as many new shapes and new beings as they undertake jobs, who are prelates to their very liver and intestines, and drag their position with them into the privy. . . *They give themselves up so much to their fortune that they even unlearn their natures* [Quintus Curtius].

The mayor and Montaigne have always been two, with a very clear separation. For all of being a lawyer or a financier, we must not ignore the knavery there is in such callings. An honest man is not accountable for the vice or stupidity of his trade, and should not therefore refuse to practice it; it is the custom of his country, and there is profit in it. We must live in the world and make the most of it such as

---

<sup>31</sup> Montaigne, 1958, 760. I take some of the historical account from John Kekes, “Constance and Purity.” *Mind*, Vol. 92, No. 368. (Oct. 1983).

<sup>32</sup> Montaigne, 1958, 600.

we find it. But the judgment of an emperor should be above his imperial power, and see and consider it as an extraneous accident; and he should know how to find pleasure in himself apart, and to reveal himself like any Jack or Peter, at least to himself.<sup>33</sup>

Luban interprets Montaigne as saying he can be true to a moral self and put on the Mayor position as a totally separate role or mask that he must wear. Arthur Isak Applbaum also criticizes Montaigne along a similar line. Applbaum states, “Montaigne thinks that the moral upshot of judgment’s independence is that the person is not accountable for the vice and stupidity of one’s role. Quite the opposite: not only is the deceitful lawyer or financier also a dishonest man, but he has the critical distance from whence to judge this.”<sup>34</sup> Luban also notes Montaigne’s apparent avoidance of accountability and interprets it as dissonance reduction. To accomplish this reduction in his “quest for integrity,” Montaigne may employ two strategies.

The first is a “schizophrenic” strategy in which Montaigne claims he has two distinct identities and is therefore not accountable for the stupidity of his trade. The second strategy is to say that he is only one unified self because he has not departed “one nail’s breadth” from himself.<sup>35</sup> But contrary to these two strategies, Montaigne is still the Mayor performing the Mayor’s duties. He is accountable for his actions no matter how much he blames his role. Luban states, “If I am right that schizophrenia and restricted

---

<sup>33</sup> Montaigne, 1958, 770, 774.

<sup>34</sup> Applbaum concludes his book with a chapter titled “Montaigne’s Mistake.” The last sentence of the chapter is, “Montaigne is wrong: lawyers and financiers, politicians and public servants, *are* responsible for the vice and stupidity of their trades, and *should* refuse to practice them in vicious and stupid ways.” Arthur Isak Applbaum, *Ethics for Adversaries: The Morality of Roles in Public and Professional Life*. (New Jersey: Princeton University Press, 1999). 75 and 259.

<sup>35</sup> Luban takes this interpretation from Gerald J. Postema, “Self Image, Integrity, and Professional Responsibility.” *The Good Lawyer: Lawyers’ Roles and Lawyers’ Ethics*. Ed. David Luban (Maryland: Rowman & Littlefield Publishing Inc., 1983). Luban, 2007, 290.

identification *are* fictions, and that Montaigne's arguments for non-accountability fail, we are left with the situation of someone whose practice of a customary, profitable profession drives him to stable, self-justifying belief-changes whose only drawback is that they happen to be lies."<sup>36</sup> He concludes that Montaigne and those like him are driven to self-deception by their quest for integrity.

In responding to Luban's argument I present another interpretation of Montaigne's position in the context of his life. I also use Montaigne's example as evidence that Montaigne is on an *actual* quest for integrity that does not involve cognitive dissonance reduction and self-deception. I conclude that Luban's concerns about the quest for integrity do not account for those people who are genuinely on a quest for substantive integrity.

Luban argues that Montaigne's quest for integrity drives him to the false belief that his immoral actions as Mayor are not actually his actions. Luban interprets Montaigne's comments as defending his integrity while holding a role that requires the use of vice. But is Montaigne self-deceived or is he discussing another topic altogether? Looking into the context of his discussion and other comments about how he maintains his integrity, I come to a different conclusion.

When we consider Montaigne's comments about taking on a role like a mask, we see that he is actually *criticizing* those who change their self-conception to fit these roles. He derisively talks of those who, "transform and transubstantiate themselves into as many new shapes and new beings as they undertake jobs." Montaigne is not arguing that

---

<sup>36</sup> Luban, 2007, 290.

this is what he does, but rather that people who have no grounded self-conception too easily change their passions to fit their roles. He is concerned with those who put their passions into pursuits or positions that weigh them down. Instead, he recommends that we must detach our passions from professions and some relationships. He quotes Statius who says, “Passion handles all things ill.”<sup>37</sup> Montaigne recommends that we approach professional positions and even chess with judgment and skill in order to avoid injecting our passion and violent intensity into a position. He condemns those who are so lost in their positions that they drag it “with them even into their privy.”<sup>38</sup>

Far from being deceived about who he actually is, Montaigne is going to great lengths to maintain a true self in defiance of his profession and position. In light of Montaigne’s argument about avoiding passion in positions, we can better understand the statement that Luban interprets as indicating self-deception. “The mayor and Montaigne have always been two, with a very clear separation. For all of being a lawyer or a financier, we must not ignore the knavery there is in such callings. An honest man is not accountable for the vice or stupidity of his trade, and should not therefore refuse to practice it.” On the surface, Montaigne appears to be justifying knavery when it is part of a profession and arguing that those who practice these professions are not accountable. Luban and Applbaum interpret the passage this way and argue that Montaigne is arguing for non-accountability and is deceiving himself regarding the harmony he believes he has within himself.

---

<sup>37</sup> Montaigne, 1958, 774.

<sup>38</sup> Montaigne, 1958, 774.

My interpretation is different given the context, the passage, and other comments by Montaigne. The context is Montaigne arguing that we should be more Stoic and less passionate about our professions so as to live a more peaceful life. The passage states that the honest man must take account of the knavery in some professions and is not accountable for the trade he enters. Montaigne is not saying that the lawyer is not accountable for his actions, only for the current state of his profession. Montaigne is not arguing for non-accountability, but rather for seeing the professions as they are. The line following this passage is, “We must live in the world and make the most of it such as we find it.”<sup>39</sup> Montaigne is being a realist about the state of the professions and the world and is encouraging those who enter the professions to do their best given its current state. Montaigne is not self-deceived. He is looking at the world with open eyes.

Upon a further reading of his essay “Of Husbanding Your Will,” I find additional evidence that, Montaigne is not caught in self-deception nor is he advocating it. For example, Montaigne believes that morality must be a commitment that cannot change with different circumstances. In reference to legal settlements that are dishonest and shameful, he laments that, “we seek only to save appearances, and meanwhile betray and disavow our true intentions. . . We give ourselves the lie to save the lie we have given to someone else.”<sup>40</sup> These words are hardly those of one who is self-deceived about maintaining an upright appearance. On the same topic he states,

You must not consider whether your action or your word may have another interpretation; it is your true and sincere interpretation that you must henceforth

---

<sup>39</sup> Montaigne, 1958, 774.

<sup>40</sup> Montaigne, 1958, 780.

maintain, whatever it costs you. Your virtue and your conscience are addressed; these are not parts to put behind a mask. Let us leave these vile means and expedients to the chicanery of the Palace of Justice.<sup>41</sup>

Montaigne is adamant that virtue and conscience should not be put behind a mask in general. Even at great cost we should be frank about how we interpret events and not bend to our adversaries just for appearances sake. He does say that the practices found in the Palace of Justice (*Palais de Justice*) may require some type of mask, but I take his comment as a realistic and condemning assessment more than an endorsement of living two lives. In actual life he wants us to be true to ourselves.

In another place Montaigne reaffirms his desire to be true to himself and to avoid false appearances. He describes himself as a mild person who puts on the mask of anger only when it is needed. He says, “I was born of a family that from way back has flowed along without glamour and without tumult, a family ambitious above all for integrity.” Contrary to Luban’s interpretation of Montaigne as deceiving himself, it appears that Montaigne is actually quite clear about who he is and who holds the office of Mayor. Montaigne is not deceiving himself about having two identities. He is claiming that he is who he is and he takes on the office of Mayor as himself. In line with his true self he does not feign passion for the duties of his office, but performs them in accordance with his self-conception and moral identity. He states, “I did not leave undone, as far as I know, any action that duty genuinely required of me. I easily forgot those that ambition mixes up with duty and covers with its name. Those are the ones that most often fill the

---

<sup>41</sup> Montaigne, 1958, 780.

eyes and ears, and satisfy men. Not the thing but the appearance pays them.”<sup>42</sup>

Montaigne is doing the exact opposite of what Luban claims. Instead of trying to appear upright, Montaigne is living according to his deepest ID-commitments and avoiding false ambition and appearances that have attached to the office of Mayor.

Montaigne’s clearest articulation of his genuine quest for integrity comes from his comments about how he handles princes who want to confide in him or use him.

Montaigne resists tying himself too closely to these people, but he must still deal with them. He states, “But these are princes who do not accept men halfway and scorn limited and conditional services. There is no remedy. I frankly tell them my limits.”<sup>43</sup> Montaigne is on a quest for integrity, but it is not producing self-deception and rationalization.

Instead, his quest for substantive integrity (i.e. self-consistency and an ID-commitment to moral values) is producing a dialogue about what he will and will not do. John Kekes calls these limits Montaigne’s *unconditional commitments* that “are the core of a person’s pattern, the fundamental components of his identity, his most basic allegiances.”<sup>44</sup>

I spend time on Montaigne’s example because it cuts to the heart of Luban’s claims about the quest for integrity along both the low road and the high road. In his position as Mayor, Montaigne exhibited integrity along the lines of the identity account not the cognitive dissonance account. A quest for integrity as mere coherence did not drive him to rationalization and self-deception as Luban wants to show. Instead, Montaigne identifies with his ID-commitments to moral values, avoids pouring false

---

<sup>42</sup> Montaigne, 1958, 781.

<sup>43</sup> Montaigne, 1958, 603.

<sup>44</sup> Kekes, 1983, 514.

passion into the duties of his office, and tells those who would pressure him to compromise that he has certain limits that he will not pass. In this context, Montaigne can say without self-deception that, “I have been able to take part in public office without departing one nail’s breadth from myself, and to give myself to others without taking myself from myself.” As for Montaigne’s realistic and critical comments about lawyers, financiers and the Palace of Justice, these do make Luban’s point about how some people act only for appearance and personal gain. But Montaigne is not recommending that these people deceive themselves about the world and their professions. Instead, they should make the best of it, tell others their limits, be true to themselves as much as possible, and pursue substantive integrity with one’s eyes wide open to the professional pressures to compromise.

My conclusion is that Luban’s description of the quest for integrity on both the low road and the high road is incorrectly named. Montaigne was on a quest for *substantive* integrity, but it moved him to clearly define his ID-commitments and avoid rationalization and self-deception. But then how do we describe those Luban and Montaigne criticize for transforming themselves for each position and losing themselves in false appearances? Luban has already provided us with the description. “From the inside, the quest for integrity and the process of rationalizing our actions are nearly impossible to distinguish.”<sup>45</sup> The reason they are impossible to distinguish is because what Luban calls the quest for integrity is probably a combination of cognitive dissonance reduction, moral disengagement, and an expedient ideology. Integrity, at

---

<sup>45</sup> Luban, 2007, 291.



least the kind Montaigne pursued, has nothing to do with changing one's moral principles, rationalizing immoral conduct, or trying to appear upright. These are all mechanisms for pursuing the subjective goods of pleasure and advantage at the expense of living an authentic life.

#### **1.4 - Section Summary**

I have argued that the quest for integrity, as defined by Luban, is not necessary to produce the dangers that he outlines. I first pointed out that Luban's concerns are based on a person who is already heading down the slippery slope towards immoral behavior and that the quest for integrity as coherence could also lead a person to greater moral certainty. I then argued that the quest for integrity is not necessary for changing one's beliefs and commitments, rationalizing one's immoral conduct, or deceiving oneself. The actual targets of Luban's concern seem to be people who have an inaccessible or undeveloped moral identity and possibly an expedient ideology. Most of his concerns do not apply to people of substantive integrity.

As for his concern that a person on the high road of integrity cannot find north on her moral compass, I refer Luban to those moral exemplars who have consistently maintained access to their moral identity and acted accordingly in their quest for substantive integrity. Contrary to Luban's interpretation, Montaigne's example confirms that a quest for substantive integrity appears to produce less self-deception, the clarification of one's moral values, and the setting of clear moral limits within a profession.

I did allow that the quest for integrity, as narrowly defined, may be sufficient for belief change, rationalization and self-deception. But to take this path, Luban needs to explain how the quest for integrity so defined can also produce moral exemplars who do not change their beliefs, who do not rationalize their behaviors, and who take responsibility for their actions.

I conclude this chapter and the dissertation by considering different recommendations for maintaining one's integrity across situations.

## **Section 2: Integrity and Defending One's Self**

In my criticism of Luban's arguments against the quest for integrity, I do not mean to take away from a mutual concern we have about the seeming fragility of moral values when confronted with cultures that support an expedient ideology. The situationist critique of character and personality traits highlights human susceptibility to compromise basic moral values when confronted with unique situational factors. Cognitive dissonance reduction and moral disengagement are two other factors that may encourage a person to compromise her moral values. In this section I present four recommendations for maintaining one's moral values and then evaluate them using the moral identity account. By considering these recommendations in light of the moral identity account, I underscore the benefits of the account in accurately describing what it means to have substantive integrity. I first summarize the necessary conditions for being a person of substantive integrity and then present and evaluate the four recommendations for maintaining one's moral values across situations.

In earlier chapters I outlined three necessary conditions for having substantive integrity. First, to be a person of substantive integrity one must have coherent ID-commitments to moral values and the determination to consistently act on them. This condition assumes that a person of substantive integrity does not violate moral values when actualizing ID-commitments to other values to which they are committed. Second, a person of substantive integrity must have a disposition to be true to her self-conception. I described this particular disposition as having a serious concern to maintain a coherent self-concept and the motivation for this concern can be an ID-commitment to self-consistency and/or acting from a particular ID-commitment. Third, a person of substantive integrity must have a self-conception that is constituted such that her moral identity is highly central to this self-conception and also constantly accessible to her working self-concept.

These necessary conditions do not require a person to consciously select, recognize, or articulate her ID-commitments nor do they require a person to “understand why” she has such a strong disposition to be true to herself. They do require that a person of substantive integrity be capable of understanding reasons for action and consider new information and evidence regarding her values. As I explained in Chapter 3, even though the moral identity account is built around core ID-commitments that constitute a self-conception, a person of substantive integrity is always maturing and learning how to wisely live out her moral values. Merely having an ID-commitment does not excuse a person from considering different perspectives on situations and role obligations.

Regardless of the strength of one's ID-commitment to moral values, situational forces and professional roles can create great pressure to compromise. To solve this problem, I consider four general recommendations to protect us from failing to maintain moral values across situations. The four are to 1) reflect on forces that work against substantive integrity and consistent moral behavior; 2) avoid compromising situations; 3) set limits on behavior; and 4) daily reinforce one's ID-commitments to moral values.

The first is to reflect on the psychological heuristics that can trip us up in our quest for substantive integrity. John Doris has argued that we are often overly confident about how character can withstand situational factors. As an example, he notes that Stanley Milgram asked subjects to predict the maximum shock they would use to punish a victim if commanded to do so. No subject said they would go beyond 300 volts.<sup>46</sup> But when 40 subjects were put in the Voice Feedback condition (i.e. subject could hear but not see the victim being shocked), 65% of the subjects shocked the learner past the 450 volt mark.<sup>47</sup> Doris claims that we hinder our deliberation when we assume that character can stop us from compromising moral values across situations. The same concern arises about rationalizing behavior in order to appear upright. Luban considers the recommendation that we learn the truth about the dynamics of cognitive dissonance reduction and rationalization so that we can avoid their influence or at least recognize them when we see them.<sup>48</sup>

---

<sup>46</sup> John Doris, *Lack of Character* (New York: Cambridge University Press, 2002). 146-147.

<sup>47</sup> Milgram, 1974, 141.

<sup>48</sup> Luban, 2007, 291.

The problem with these strategies is that even if a person knows that her character may not withstand situational factors and that cognitive dissonance reduction may lead her to rationalize her behavior, she may not recognize when to apply these insights in her daily life. Simply knowing a fact about one's psychology and tendencies may not translate to the motivation or skill to overcome these habits. Luban even admits that once while walking with three other people, he did not immediately help an unconscious person lying on the ground (i.e. diffusion of responsibility). Knowledge about the weakness of character, situational influences, and the power of cognitive dissonance can help us interpret our actions and possibly avoid others, but knowledge alone is unlikely to produce the consistent behavior expected from a person of substantive integrity. The moral identity account suggests that the problem is not the knowing of something but having that knowledge constantly accessible. I discuss this under the fourth point.

Doris also recommends that we must go beyond learning about psychological factors and must invest "more of our energies in attending to the features of our environment that influence behavioral outcomes."<sup>49</sup> The second recommendation is to avoid situations that may lead to moral compromise. Doris provides an excellent example of a person who is invited to dinner by a colleague with whom she has been flirting for a long time. The person is in a committed relationship and doesn't see how dinner and some wine could result in compromising her fidelity. She may think to herself, "I am in a committed relationship and am morally upright. Nothing is going to happen." Doris insists that unless she wants to be like an obedient Milgram subject, she

---

<sup>49</sup> Doris, 2002, 146.

should avoid the dinner. If she discounts the strength of her character and focuses on the unpredictability of behavior across situations, then she should “recognize the situational pressures may all too easily overwhelm character and avoid the dangerous situation.”<sup>50</sup>

The strategy is to avoid situations that may overwhelm our commitments to moral values and to seek out those situations that promote ethical conduct.

The moral identity account supports this strategy because it recognizes that a person with an ID-commitment to moral values should not necessarily seek out situations where she will be sorely tempted to compromise. Different people have different weaknesses and a person of substantive integrity would understand through experience what makes her moral identity less accessible. These are the situations, professions, or people who trigger those desires and even identities that she has rejected as part of her self-conception. In the dinner case, the person of substantive integrity knows that being in a committed relationship means that flirting with a colleague is compromising that ID-commitment. I would suggest that the process of avoiding infidelity begins at the flirting stage which blocks the dinner invitation from ever arising.

A problem with the situation avoidance recommendation is that situations can present us with a variety of factors and activate motivations that cannot be understood and managed in a high-pressure situation. For example, consider a scientist who must decide if she is going to falsify data “just this once” and just for “this article” to keep her lab’s research funding, thereby saving the jobs of four researchers. Moral disengagement is knocking at the door before she can assess the situation and determine what to do. The

---

<sup>50</sup> Doris, 2002, 147.

problem is not so much that the situational factors tempt us but that we are too willing to listen to the temptation. I agree that the person of substantive integrity must avoid situations that are particularly difficult for her to handle given her self-conception. But some situations are inescapable and unexpected and the person of substantive integrity must be able to maintain her moral identity accessibility with little or no notice.

The third strategy is to set limits on one's conduct and this may explain how a person can maintain her moral identity accessibility even in unexpected situations. Luban was inspired to suggest this strategy after reading about David Heilbroner, a former New York City prosecutor. Heilbroner had promised himself before he joined the district attorney's office that he would never take a case if he doubted the defendant's guilt. After being assigned a case in which his star witness was a known drug dealer and notorious liar, Heilbroner decided that the case and one's like it pushed his personal ethics too far. The defendant was acquitted and Heilbroner quit his job not long afterwards. He said, "To stay on much longer meant maintaining a blindered belief in the rectitude of our work."<sup>51</sup> Luban compares conduct limits to a canary in a mineshaft. He says the formula to follow is something like, "Whatever else I do, and however else my views change, I will never, ever . . ."<sup>52</sup> It is interesting that this is the strategy used by Montaigne when dealing with those who wanted to use him for their own ends: "I frankly tell them my limits."<sup>53</sup>

---

<sup>51</sup> Luban, 2007, 294. See David Heilbroner, *Rough Justice: Days and Nights of a Young D.A.* (New York: Dell, 1990).

<sup>52</sup> Luban, 2007, 295.

<sup>53</sup> Montaigne, 1958, 603.

A person of substantive integrity would seem to live by this strategy. Recall the story of Socrates who risked his life because he would do nothing unjust or impious. Or Thomas More who knew that his limits would not allow him to affirm Henry VIII's marriage because it was contrary to his religious convictions. Consider also Anthony Flew who turned from being a life-long atheist to a theist because he had an ID-commitment to "follow the argument wherever it leads." The person of substantive integrity sets limits, either consciously or non-consciously, and reliably obeys those limits because they are integrated into her self-conception. The moral identity account explains how these limits are integrated into a person's self-conception by the creation of ID-commitments to moral values. The motivation to be true to oneself then motivates the person to act on these limits even under extreme pressure to compromise.

The final strategy for maintaining one's substantive integrity is to reinforce one's ID-commitments to moral values daily. I described this process in detail in Chapter 4 in reference to the experiment in which people with HMID cooperated more than other groups if they had been primed with a moral stimulus. I also explained how people who performed daily religious devotions were less prejudiced than those who did not. The point of daily reinforcement is not to fundamentally change an immoral person into a moral person. In fact, the research covered in Chapter 4 demonstrates that moral primes had little effect on subjects with low moral identity centrality. Daily moral reinforcement does seem to help a person with HMID to maintain constant moral identity access.

Both Doris and Luban also hint at some type of self-priming ritual. Doris notes that not much empirical work has been done on how reflecting on psychological



processes and behavior, what could be construed as a moral prime, can facilitate deliberation but he does cite one interesting experimental study. In this study, students who attended either a lecture or short film on group effects and helping behavior were more likely than those in a control group to help someone in a staged emergency when they were a bystander. The researcher and his colleagues found that subjects were more likely to help in these situations even two weeks after attending the lecture or film.<sup>54</sup> In this context, I am also reminded of the defiant professor in the Milgram experiments. Being a divinity professor who taught the Old Testament (i.e. Hebrew) scriptures, he probably constantly read and discussed moral standards of justice, helping those in need, and answering to a higher power. When asked, “What in your opinion is the most effective way of strengthening resistance to inhumane authority?” he responded, “If one had as one’s ultimate authority God, then it trivializes human authority.”<sup>55</sup>

Luban offers several practical examples of ways to reinforce one’s ID-commitment to behavioral limits. He states, “If necessary, write down the ‘I will never, ever’ formula. Put it in an envelope, keep it in a drawer, and pull it out sometimes to remind yourself what it says. And the moment the canary dies, get out of the mineshaft.”<sup>56</sup> Leaders have often lived by this advice and kept mottos, quotes, or proverbs sitting on their desks or hanging on their walls to keep their moral identity accessible when they must make important decisions. A textbook case of how

---

<sup>54</sup> Doris, 2002, 148. See A. L. Beaman, P. J. Barnes, B. Klentz, and B. McQuirk. “Increasing Helping Rates through Information Dissemination: Teaching Pays.” *Personality and Social Psychology Bulletin* 4, (1978).

<sup>55</sup> Milgram, 1974, 49.

<sup>56</sup> Luban, 2007, 295.

reinforcing moral values makes one's moral identity accessibility is Johnson & Johnson corporation's (J&J) handling of tainted Tylenol in 1982. After seven people died in the Chicago area, it was revealed that a non-employee was tampering with bottles and lacing the medicine with cyanide. At an estimated cost of \$100 million, CEO Jim Burke decided to pull all Tylenol from the US market. His reasoning was that the J&J credo, which outlines the moral values of the company, requires him to put the lives of customers over profits. Far from being a dusty plaque on the wall, the credo was a constant and accessible reminder of what decisions should and should not be made and J&J employees knew that their actions would be held up to that standard. Burke estimated that as CEO he spent 40% of his time making sure the credo was part of the J&J culture.<sup>57</sup>

While all four recommendations have potential for helping a person maintain her moral values, those who rely on maintaining access to their moral identity seem the most promising. The first two recommendations are important before an event occurs or if one is trying to avoid big mistakes while still enjoying temptation from a far. In contrast, the last two prepare a person for whatever situation may come because they help a person regularly clarify and publish her ID-commitments to moral values. The key to conducting oneself morally across situations appears to be developing and maintaining one's moral identity centrality and accessibility by setting moral limits and reinforcing moral standards each day. By following these recommendations and developing deep

---

<sup>57</sup> James C. Collins and Jerry I. Porras, *Built to Last: Successful Habits of Visionary Companies*. (New York: Harper Collins, 1997). 58-61. The first part of the J&J Credo states, "We believe our first responsibility is to the doctors, nurses and patients, to mothers and fathers and all others who use our products and services. In meeting their needs everything we do must be of high quality. We must constantly strive to reduce our costs in order to maintain reasonable prices. Customers' orders must be serviced promptly and accurately. Our suppliers and distributors must have an opportunity to make a fair profit."

ID-commitments to moral values, a person can pursue substantive integrity with little fear of the belief change, rationalization and self-deception that Luban describes.

## Bibliography

- Abelson, Robert P., Kurt P. Frey, and Aiden P. Gregg. *Experiments with People: Revelations from Social Psychology*. New Jersey: Lawrence Erlbaum Associates, 2003.
- Adams, Robert Merrihew. *A Theory of Virtue*. New York: Oxford University Press, 2006.
- Annas, Julia. "My Station and Its Duties: Ideals and the Social Embeddedness of Virtue." *Proceedings of the Aristotelian Society* (January 2002): 109-123.
- Appelbaum, Arthur Isak. *Ethics for Adversaries: The Morality of Roles in Public and Professional Life*. New Jersey: Princeton University Press, 1999.
- Aquino, Karl and Americus Reed, "The Self-Importance of Moral Identity." *Journal of Personality and Social Psychology*, 83, No. 6. (2002): 1423-1440.
- Aquino, Karl and Dan Freeman, "Moral Identity in Business Situations: A Social Cognitive Framework for Understanding Moral Functioning." *Personality, Identity, and Character: Explorations in Moral Psychology*. Edited by Darcia Narvaez and Daniel K. Lapsley, 375-395. New York: Cambridge University Press, 2009.
- Aquino, Karl, Americus Reed II, Stefan Thau, and Dan Freeman. "A Grotesque and Dark Beauty: How Moral Identity and Mechanisms of Moral Disengagement Influence Cognitive and Emotional Reactions to War." *Journal of Experimental Social Psychology*, 43, (2007): 385-392.

Aquino, Karl, Dan Freeman, Americus Reed II, Vivien K. G. Lim, and Will Felps.

“Testing a Social-Cognitive Model of Moral Behavior: The Interactive Influence of Situations and Moral Identity Centrality.” *Journal of Personality and Social Psychology*, 97, No. 1. (2009): 123-141.

Aristotle, *Nicomachean Ethics*. Translated by Terence Irwin. Indianapolis: Hackett Publishing Company, 1999.

Asch, Solomon. *Social Psychology*. New Jersey: Prentice Hall, 1952.

Audi, Robert and Patrick E. Murphy. “The Many Faces of Integrity.” *Business Ethics Quarterly*, 16, No. 1. (2006): 3-21.

Bandura, Albert. “Social Cognitive Theory of Personality.” *The Coherence of Personality: Social Cognitive Bases of Consistency, Variability, and Organization*. Edited by Daniel Cervone and Yuichi Shoda, 185-241. New York: The Guilford Press, 1999.

Batson, C. Daniel, P.J. Cochran, M.F. Biederman, J. L. Blosser, M. J. Ryan, and B. Vogt. “Failure to Help When in a Hurry: Callousness or Conflict”, *Personality and Social Psychology Bulletin* 4 No. 1 (1978): 97-101.

Beaman, A. L., P. J. Barnes, B. Klentz, and B. McQuirk. “Increasing Helping Rates through Information Dissemination: Teaching Pays.” *Personality and Social Psychology Bulletin* 4, (1978): 406-411.

Beebe, John. *Integrity in Depth*. College Station, Texas: Texas A&M University Press, 1992.

- Benjamin, Martin. *Splitting the Difference: Compromise and Integrity*. Kansas: University Press of Kansas, 1990.
- Bigelow, John and Robert Pargetter. "Integrity and Autonomy." *American Philosophical Quarterly* Vol. 44 No. 1 (2007): 39-49.
- Blasi, Augusto. "Moral Character: A Psychological Approach". *Character Psychology and Character Education*. Edited by Daniel K. Lapsley and F. Clark Power, 67-100. Indiana: University of Notre Dame Press, 2005.
- Blasi, Augusto. "Moral Cognition and Moral Action: A Theoretical Perspective." *Developmental Review*, 3. (1983): 178-210.
- Blass, Thomas. "Understanding Behavior in the Milgram Obedience Experiment: The Role of Personality, Situations, and Their Interaction." *Journal of Personality and Social Psychology*, Vol. 60, No. 3, (1991): 398-413.
- Blount Jr., Roy. *Robert E. Lee*. New York: Viking, 2003.
- Blum, Lawrence. *Moral Perception and Particularity*. New York: Cambridge University Press, 1994.
- Blustein, Jeffrey. *Care and Commitment: Taking the Personal Point of View*. New York: Oxford University Press, 1991.
- Boisjoly, Russell, Ellen Foster Curtis, and Eugene Mellican. "Roger Boisjoly and the Challenger Disaster: Ethical Dimensions." *Journal of Business Ethics* 8, (1989): 217-230.

Calhoun, Cheshire. "Standing for Something." *The Journal of Philosophy*, Vol. 92, No. 5.  
(May 1995): 235-260.

Cicero. *The Verrine Orations*. Translated by L. H. G. Greenwood. Cambridge: Harvard  
University Press, 1928.

Cohen, T. R., R. M. Montoya, and C. A. Inkso. "Group Morality and Intergroup  
Relations: Cross Cultural and Experimental Evidence." *Personality and Social  
Psychology Bulletin*, 32. (2006): 1559-1572.

Colby, Anne and William Damon, "The Uniting of Self and Morality in the Development  
of Extraordinary Moral Commitment." *The Moral Self*. Edited by Gil G. Noam  
and Thomas E. Wren, 149-174. Massachusetts: MIT Press, 1993.

Colby, Anne and William Damon, *Some Do Care: Contemporary Lives of Moral  
Commitment*. New York: Free Press, 1992.

Collins, James C. and Jerry I. Porras. *Built to Last: Successful Habits of Visionary  
Companies*. New York: Harper Collins, 1997.

Cox, Damian, Marguerite La Caze, and Michael P. Levine, "Integrity", *Stanford  
Encyclopedia of Philosophy*. Updated August 10, 2008.  
<http://plato.stanford.edu/entries/integrity/>.

Cox, Damian, Marguerite La Caze, and Michael P. Levine, *Integrity and the Fragile Self*.  
Vermont: Ashgate, 2003.

- Darely, John M. and C. Daniel Batson. “‘From Jerusalem to Jericho’: A Study of Situational and Dispositional Variables in Helping Behavior”, *Journal of Personality and Social Psychology* 27, No. 1, (1973): 100-108.
- Darley, John M. and Bibb Latané, “Bystander Intervention in Emergencies: Diffusion of Responsibility.” *Journal of Personality & Social Psychology* 8. (1968):377-383.
- Darwall, Stephen. *The Second Person Standpoint: Morality, Respect, and Accountability*. Massachusetts: Harvard University Press, 2006.
- Descartes, Rene. *Meditations on First Philosophy*. Translated by Ronald Rubin. California: Areté Press, 2001.
- Doris, John. *Lack of Character*. New York: Cambridge University Press, 2002.
- Elms, A. C. and Stanley Milgram, “Personality Characteristics Associated with Obedience and Defiance Toward Authoritative Command.” *Journal of Experimental Research in Personality*, 1 (1966): 282-289.
- Festinger, Leon and J. M. Carlsmith, “Cognitive Consequences of Forced Compliance,” *Journal of Abnormal and Social Psychology* 58, (1959): 203-211.
- Finnis, John. *Fundamentals of Ethics*. Washington D.C.: Georgetown University Press, 1983.
- Flew, Anthony. *There Is a God: How the world’s most notorious atheist changed his mind*. New York: Harper Collins, 2007.



Foot, Philippa. "Virtues and Vices". *Virtue Ethics*. Edited by Stephen Darwall, 105-120. Massachusetts: Blackwell Publishing, 2007.

Fox, John. *Foxes Book of Martyrs*. Edited by William Byron Forbush. Accessed August 12, 2010. Online book: <http://www.ccel.org/f/foxe/martyrs/fox112.htm>.

Frankfurt, Harry. "Identification and Wholeheartedness." *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. Edited by Ferdinand Schoeman, 27-45. New York: Cambridge University Press, 1987.

Frimer, Jeremy A. and Lawrence J. Walker. "Towards a New Paradigm of Moral Personhood." *Journal of Moral Education* Vol. 37. No. 3. (2008): 333-356.

Gershon, Elliot S. "Ernst Rüdin, a Nazi Psychiatrist and Geneticist." *American Journal of Medical Genetics* 74 (1997): 457-458.

Glock, Charles Y. and Rodney Stark, *Christian Beliefs and Anti-Semitism*. New York: Harper and Row Publishers, 1966.

Goodstein, Jerry D. "Moral Compromise and Personal Integrity: Exploring the Ethical Issues of Deciding Together in Organizations." *Business Ethics Quarterly* 10, No. 4. (2000): 805-819.

Goruch, Richard and Danile Aleshire, "Christian Faith and Ethnic Prejudice: A Review and Interpretation of Research." *Journal for the Scientific Study of Religion* 13, (1974): 281-307.

- Graham, Jody L. "Does Integrity Require Moral Goodness?" *Ratio*. XIV, No. 3 (September 2001): 234-251.
- Halfon, Mark S. *Integrity: A Philosophical Inquiry*. Philadelphia: Temple University Press, 1989.
- Harman, Gilbert. "Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error," *Proceedings of the Aristotelian Society*. (1999): 315-331.
- Heilbroner, David. *Rough Justice: Days and Nights of a Young D.A.* New York: Dell, 1990.
- Howe, Leslie A. "Self and Pretence: Playing with Identity." *Journal of Social Philosophy* Vol. 39, No.4. (2008):564-582.
- Irwin, T. H. "Aristotle's Conception of Morality." *Proceedings of the Boston Area Colloquium in Ancient Philosophy* Vol. 1 (1985): 115-143.
- Joan, McCarthy. *Dennett and Ricoeur on the Narrative Self*. New York: Humanity Books, 2007.
- Johnson, Paul. *Churchill*. New York: Penguin Group, 2009.
- Kant, Immanuel. *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy*. Translated by Mary J. Gregor. New York: Cambridge University Press. 1999.

- Kateb, George. "Socratic Integrity." *Integrity and Conscience*. Edited by Ian Shapiro and Robert Adams, 77-112. New York: New York University Press, 1998.
- Kekes, John. *Enjoyment*. New York: Oxford Press, 2008.
- Kekes, John. "Constancy and Purity" *Mind* Vol. 92, No. 368. (Oct. 1983): 499-518.
- Kripke, Saul. "On Rules and Private Language." *The Philosophy of Language – Fifth Edition*. Edited by A.P. Martinich. New York: Oxford University Press. 2008.
- Lacayo, Richard and Amanda Ripley, "The Whistleblowers." *Time*. (December 30, 2002): 45-50.
- Leadershipnow.com. "Integrity." Accessed September 9, 2010.  
<http://www.leadershipnow.com/integrityquotes.html>"Integrity."
- Linder, Douglas. "Trial of Thomas More" (2006). Accessed September 9, 2010.  
<http://www.law.umkc.edu/faculty/projects/ftrials/more/moreaccount.html>.
- Luban, David. *Legal Ethics and Human Dignity*. New York: Cambridge University Press, 2007.
- MacIntyre, Alasdair. *After Virtue*. Indiana: University of Notre Dame, 1984
- May, Larry. "Integrity, Self, and Value Plurality." *Journal of Social Philosophy* Vol. 27 No. 1 (Spring 1996): 123-139.
- McCullough, David. *John Adams*. New York: Simon and Schuster, 2002.
- McFall, Lynne. "Integrity." *Ethics* Vol. 98, No. 1 (1987): 5-20.

- McFerran, Brent, Karl Aquino, and Michelle Duffy. "How Personality and Moral Identity Relate to Individuals' Ethical Ideology." *Business Ethics Quarterly*, 20, 1. (January 2010): 35-56.
- Milgram, Stanley. "Some Conditions of Obedience and Disobedience to Authority." *The Individual in a Social World*. Edited by John Sabini and Maury Silver. New York: McGraw-Hill, Inc., 1992. 136-161.
- Milgram, Stanley. *Obedience to Authority: An Experimental View*. New York: Harper and Row Publishers, Inc, 1974.
- Mischel, Walter. *Personality and Assessment*. New York: Wiley, 1968.
- Monroe, Kristen Renwick. *The Hand of Compassion: Portraits of Moral Choice During the Holocaust*. New Jersey: Princeton University Press, 2004.
- Montaigne, Michel de. *The Complete Essays of Montaigne*, Translated by Donald M. Frame. California: Stanford University Press, 1958.
- New Oxford American Dictionary*. New York: Oxford University Press, 2005.
- Oakley, Justin and Dean Cocking. *Virtue Ethics and Professional Roles*. New York: Cambridge University Press, 2001.
- Oliner, Pearl M. *Saving the Forsaken: Religious Culture and the Rescue of Jews in Nazi Europe*. New Haven: Yale University Press, 2004.
- Oliner, Sam. P. and Pearl. M. Oliner, *The Altruistic Personality: Rescuers of Jews in Nazi Germany*. New York: Free Press, 1988.

- Palanski, Michael E. and Francis J. Yammarino. "Integrity and Leadership: Clearing the Conceptual Confusion." *European Management Journal* (June 2007): 171-184.
- Pence, George E. "The Tuskegee Study." *Contemporary Issues in Bioethics*, 7e. Edited by Tom L. Beauchamp, LeRoy Walters, Jeffrey P. Kahn, and Anna C. Mastroianni, 46-53. California: Thomson-Wadsworth, 2008.
- Plato. *The Apology in The Trial and Death of Socrates*. Translated by G. M. A. Grube. Indianapolis: Hackett Publishing Company, 2000.
- Postema, Gerald J. "Self Image, Integrity, and Professional Responsibility." *The Good Lawyer: Lawyers' Roles and Lawyers' Ethics*. Edited by David Luban. Maryland: Rowman & Littlefield Publishing Inc., 1983.
- Puka, Bill. "Altruism and Character." *Moral Development, Self, and Identity*. Edited by Daniel K. Lapsley and Darcia Narvaez. 161-187, New Jersey: Lawrence Erlbaum Associates, 2004.
- Rand, Ayn. *The Journals of Ayn Rand*. Edited by David Harriman. New York: Penguin Putnam Inc., 1997.
- Rawls, John. *A Theory of Justice – Revised Edition*. Massachusetts: Harvard University Press, 1999.
- Reed II, Americus and Karl Aquino, "Moral Identity and the Expanding Circles of Moral Regard Toward Out-groups." *Journal of Personality and Social Psychology*, 84 No. 6. (2003): 1270-1286.

- Ricoeur, Paul. *One Self as Another*. Translated by Kathleen Blamey. Chicago: University of Chicago Press, 1992.
- Ross, Lee and Richard E. Nisbett, *The Person and the Situation*. Philadelphia: Temple University Press, 1991.
- Schauber, Nancy. "Integrity, Commitment, and the Concept of a Person." *American Philosophical Quarterly*, Vol. 33 No. 1. (January 1996): 119-129.
- Schlenker, Barry R. "Integrity and Character: Implications of Principled and Expedient Ethical Ideologies." *Journal of Social and Clinical Psychology* Vol. 27 No. 10) (December 2008): 1078-1125.
- Schlenker, Barry R., Marisa L. Miller, and Ryan M. Johnson, "Moral Identity, Integrity, and Personal Responsibility." *Personality, Identity, and Character: Explorations in Moral Psychology*. Edited by Darcia Narvaez and Daniel K. Lapsley, 316-340. New York: Cambridge University Press, 2009.
- Schlenker, Barry R., T. W. Britt, J. W. Pennington, R. Murphy, and K. J. Doherty. "The Triangle Model of Responsibility." *Psychological Review*, 101. (1994): 632-652.
- Seneca, Lucius Anneaus. *Seneca Moral Essays*, Translated by John W. Basore. Cambridge: Harvard University Press, 1989.
- Shinert, Gregory and Charles Ford. "The Relation of Ethnocentric Attitudes to Intensity of Religious Practice. *Journal of Educational Sociology* 32 (1958): 157-162.

Sidgwick, Henry. "Essays on Ethics and Method." Edited by Marcus G. Singer. New York: Oxford University Press, 2000.

"Sir Thomas More." *Encyclopædia Britannica*. 2010. Encyclopædia Britannica Online. Accessed August 12, 2010.

<http://www.britannica.com/EBchecked/topic/392018/Sir-Thomas-More>.

Solomon, Robert. *True to Our Feelings*. New York: Oxford University Press, 2007.

Sommers, Tamler. *A Very Bad Wizard: Morality Behind the Curtain*. Canada: Believer Books, 2009.

Strawson, P.F. "Freedom and Resentment." *Studies in the Philosophy of Thought and Action*. New York: Oxford University Press, 1968.

Tavris, Carol and Elliot Aronson, *Mistakes Were Made (but not by Me)*. New York: Harcourt, 2007.

Taylor, Gabriele. "Integrity." *Proceedings of the Aristotelian Society*, Supplementary Vol. 55 (1981): 143-159.

Taylor, Gabriele. *Pride, Shame and Guilt: Emotions of Self-Assessment*. Oxford: Oxford University Press, 1985.

Torisky, Eugene V. "Integrity and Supererogation in Ethical Communities," Paper given at the Twentieth World Congress of Philosophy, Boston, Massachusetts. August 10-15, 1998. Accessed September 9, 2010.

<http://www.bu.edu/wcp/Papers/Soci/SociTori.htm>.

- Velleman, David, "The Self as Narrator." *Self to Self: Selected Essays*. New York: Cambridge University Press, 2006.
- Velleman, David. "From Self Psychology to Moral Philosophy." *Philosophical Perspectives* 14. (2000): 349-377.
- Walker, Lawrence J. and Jeremy A. Frimer, "Moral Personality Exemplified." *Personality, Identity, and Character: Explorations in Moral Psychology*. Edited by Darcia Narvaez and Daniel K. Lapsley, 232-255. New York: Cambridge University Press, 2009.
- Walker, Lawrence J. and Karl H. Hennig. "Different Conceptions of Moral Exemplarity: Just, Brave, and Caring," *Journal of Personality and Social Psychology* Vol. 86 No. 4 (2004): 629-647.
- Wallace, James. *Virtues and Vices*. Ithaca: Cornell University Press, 1978.
- Waller, James. *Becoming Evil: How Ordinary People Commit Genocide and Mass Killing*. New York: Oxford University Press, 2007.
- Webster, Noah. "Integrity." *Webster's 1828 Dictionary*. Accessed October 30, 2010. <http://1828.mshaffer.com/d/search/word,integrity>.
- "William Tyndale." *Encyclopedia Britannica*. 2010. Encyclopædia Britannica Online. Accessed August 12, 2010. <http://www.britannica.com/EBchecked/topic/611569/William-Tyndale>.



- Williams, Bernard and J.J.C. Smart. *Utilitarianism: For and Against*. New York: Cambridge University Press, 1973.
- Williams, Bernard. "Replies." *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Edited by J.E.J. Altham and Ross Harrison, 185-224. New York: Cambridge University Press, 1995.
- Williams, Bernard. *Ethics and the Limits of Philosophy*. Massachusetts: Harvard University Press, 1985.
- Williams, Bernard. *Moral Luck*. New York: Cambridge University Press, 1981.
- Wilson, Glenn and Francis Lillie. "Social Attitudes of Salvationists and Humanists." *British Journal of Social and Clinical Psychology* 11. (1972): 220-224.
- Wittgenstein, Ludwig. *Philosophical Investigations*. Translated by G.E.M. Anscombe. Massachusetts: Blackwell, 2001.
- Wordnik. "Integrity." Accessed July 25, 2010. <http://www.wordnik.com/words/integrity>.
- Wulff, David M. *Psychology of Religion: Classic and Contemporary (2<sup>nd</sup> Edition)*. New York: John Wiley and Sons, 1997.