ABSTRACT

Taking Responsibility for Ourselves: A Kierkegaardian Account of the Freedom-Relevant Conditions Necessary for the Cultivation of Character

Paul E. Carron, Ph.D.

Mentor: Robert C. Roberts, Ph.D.

What are the freedom-relevant conditions necessary for someone to be a morally responsible person? I examine several key authors beginning with Harry Frankfurt that have contributed to this debate in recent years, and then look back to the writings or Søren Kierkegaard to provide a solution to the debate. In this project I investigate the claims of semi-compatibilism and argue that while its proponents have identified a fundamental question concerning free will and moral responsibility—namely, that the agential properties necessary for moral responsibility ascriptions are found in scenarios where the agent acts on her own as opposed to her action resulting from freedom-undermining external causes such as manipulation, phobias, etc.—they have failed to show that the freedom-relevant agential properties identified in those actual-sequence scenarios are compatible with causal determinism. My argument is that only a voluntarist-libertarian theory can adequately account for the kinds of cases that the semi-compatibilist identify. I argue that there are three freedom-relevant conditions necessary for someone to be a morally responsible person: a hierarchical understanding of human

desires [specifically and mental states generally], an incompatibilist (non-deterministic) understanding of human action, and a historical understanding of character development. The ability to reflect critically about one's own desires and emotions, and thus to have a kind of self-knowledge and understanding with regard to the springs of one's own actions, is required to make it possible for the agent to be the "source" of her own actions and character. The non-deterministic understanding of human action is needed for a similar reason: if determinism is true, then every action a person performs can be ultimately traced to and exhaustively explained in terms of factors outside the agent's control, thus making the agent's responsibility for his actions an illusion. And finally, human nature must be such that, over time, one's choices leave a dispositional residue of self-understanding and motivation in the person's self, out of which, in mature understanding and motivation, the person acts as a fully responsible agent.

Taking Responsibility for Ourselves: A Kierkegaardian Account of the Freedom-Relevant Conditions Necessary for the Cultivation of Character

by

Paul E. Carron, B.A., M.A., M.Div.

A Dissertation

Approved by the Department of Philosophy

_____
Michael D. Beaty, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of
Doctor of Philosophy

Approved by the Dissertation Committee

_____
Robert C. Roberts, Ph.D., Chairperson

_____
C. Stephen Evans, Ph.D.

_____
J. Todd Buras, Ph.D.

_____
Robert B. Kruschwitz. Ph.D.

_____
Paul Martens, Ph.D.

Accepted by the Graduate School
May 2011

_____
J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

Copyright © 2011 by Paul E. Carron

All rights reserved

TABLE OF CONTENTS

Acknowledgements vi

Chapter One: Setting the Stage: An Overview of the Project 1
Introduction 1
Chapter Two 2
Chapter Three 3
Chapter Four 5
Chapter Five 8

Chapter Two: Harry Frankfurt on Free Will and Moral Responsibility 10
Introduction 10
Frankfurt's understanding of Freedom and Responsibility 15
Frankfurt's Counterexamples against PAP 21
Frankfurt's Hierarchical Actual-Sequence Theory of Moral Responsibility 29
Identification 38
Volitional Necessity 52
Wholeheartedness, Moral Self-Cultivation and Morally Responsible Agency 74

Chapter Three: Fischer's "Semi-compatibilism" 76
Introduction 76
Fischer's Semicompatibilism 78
Fischer's Response to Frankfurt 81
Direct and Indirect Arguments for Incompatibilism 85
Guidance Control 86
Direct Arguments for Incompatilibilism 95
Critiques of Semicompatibilism 102
Conclusion 133

Chapter Four: Stump's "Source Incompatibilism" 135
Introduction 135
Stump's Thomistic Theory of Free Will and Moral Responsibility 138
Stump's Frankfurt-Style Libertarianism 149
Stump's Arguments against PAP 161
Sourcehood and Origination 172
Critique of Stump's revised FSC 172
Timpe's Critique of Stump` 183
Source Incompatibilism and the Direct Argument 184
Leeway Incompatibilism 188
Conclusion 191

Chapter Five: A Kierkegaardian Account of the Freedom-Relevant Conditions

Necessary for Morally Responsible Agency.....193
Chapter Overview.....198
Kierkegaard's Modest Libertarianism.....200
Kierkegaard's 'Leap' Reconsidered.....202
Kierkegaard's Hierarchical-Libertarian Solution to the problem of Identification and Moral Self-Cultivation.....236
Conclusion.....258

Bibliography.....261

ACKNOWLEDGMENTS

Thanks to my committee—Bob Roberts, Steve Evans, Todd Buras, Bob Kruschwitz, and Paul Martens—and to the philosophy department for training and support through the years.

CHAPTER ONE

Setting the Stage: An Overview of the Project

Introduction

This project began with an intuition: Søren Kierkegaard's decidedly libertarian understanding of free will does not neatly fit into any of the contemporary categories of free will and moral responsibility. Kierkegaard's complex moral psychology and his insistence that the goal of a fully formed self is submission to God do not square well with many contemporary libertarian theories. I turned to Harry Frankfurt's work on free will and moral responsibility and discovered that despite his compatibilist orientation he has certain elements in common with Kierkegaard, and that there are mutually illuminating elements in both thinkers. While Frankfurt's theory on its own does not adequately provide the necessary freedom-relevant grounding conditions, Kierkegaard's understanding of freedom and moral responsibility is unsystematic and benefits from Frankfurt's analytic framework. Together they combine to provide a rich and satisfying theory of the freedom-relevant conditions necessary for grounding moral responsibility.

Chapter Two

This project has an essentially Hegelian dialectic structure. In chapter two, I summarize Frankfurt's solution to his articulation of an ancient Aristotelian moral responsibility problem. Frankfurt argues that the main differentiating factor between responsible and non-responsible action is whether or not the action results from responsibility-undermining external sources such as manipulation or coercion, or stems

from the agent's own will. Frankfurt's solution to this problem is inadequate based on his own requirements for what would make the agent responsible (namely that it must stem from the agent in a way that an action that results from manipulation does not). Furthermore, there are internal contradictions within Frankfurt's own understanding of morally responsible action. I ask whether some (or any) of Frankfurt's freedom-relevant conditions necessary for grounding moral responsibility can be utilized to solve the problem he powerfully articulates. In each of the subsequent chapters I proceed to reject or substantially modify one or more of his freedom-relevant conditions. I utilize some of his insights to solve his problem. What emerges is a position that both annuls some and preserves other key aspects of this theory. My solution is hierarchical, historical, and libertarian in the classical sense and includes a knowledge component (including a cognitive understanding of desires and emotions) that Frankfurt rejects. The resulting modified version of his hierarchical view of the will makes sense of his understanding of care. Most importantly, it provides the freedom-relevant conditions for grounding morally responsible agency.

In a Hegelian spirit, I will turn to Kierkegaard to articulate the freedom-relevant conditions necessary for grounding morally responsible agency. Specifically, I address areas where the Frankfurt-style compatibilist and libertarian models fail without losing sight of the fundamental problem that Frankfurt articulates.[1] The Kierkegaardian inspired theory of the freedom-relevant conditions necessary for grounding moral responsibility

[1] The Kierkegaardian account I am articulating is transcendental in nature: Kierkegaard gives an account of the freedom-relevant conditions necessary for the possibility of morally responsible action and character formation, not an account of whether or not these conditions can be met in the actual physical world. This is both an advantage and disadvantage of my Kierkegaardian account. The accounts of Fischer, Stump, Pereboom, and Robert Kane all attempt in various ways to accommodate their accounts of free will and moral responsibility with contemporary scientific theory. Obviously Kierkegaard did not have the advantage of contemporary neuroscience, for example, so it is not surprising that he does not incorporate much scientific theory into his understanding of free will.

incorporates the previous three attempts examined in this dissertation, with two crucial additions/modifications to the previous attempts. Frankfurt illuminates the key problem, namely, differentiating cases of free and responsible volitions from volitions that result from responsibility-undermining external forces such as manipulation, phobias, coercion, etc. Furthermore, he argues that free and responsible volitions can sometimes appear similar to volitions that result from sources such as manipulation, and therefore the freedom-relevant conditions necessary for grounding moral responsibility must be able to account for this similarity. Therefore he argues that a hierarchical understanding of the will coupled with a notion of caring or volitional necessity simultaneously differentiates free and responsible volitions from non-responsible ones and accounts for the fact that these two instances can appear quite similar.

Chapter Three

In chapter three I argue that Fischer rightly criticized Frankfurt's hierarchical notion of the will because his time-slice mesh theory of moral responsibility could not differentiate free and responsible volitions from non-responsible volitions as Frankfurt intends. Fischer believes that an agent who forms free and responsible volitions exhibits guidance control, whereas the agent whose actions result from responsibility-undermining external sources does not exhibit the same control. Fischer argues that Frankfurt's theory of moral responsibility is inadequate due to the lack of a primary role for the intellect and the ahistorical nature of Frankfurt's mesh-theory account of moral responsibility. Fischer argues that the agential control necessary for moral responsibility requires moderate reasons-responsiveness as well as the proper causal history of the action. Fischer believes that Frankfurt rightly argues that agents are the source of

morally responsible actions in ways that they are not the source of non-responsible actions; but, against Frankfurt, he argues that one of the main differences between these two instances is the causal history of the action. In instances of responsibility-undermining external causes such as manipulation, the history of the action traces back to the manipulator and not to the agent's own moderately reasons-responsive mechanism. Furthermore, even if the agent were aware of sufficient reasons to form a different volition, she could not form a different volition because the action of the manipulator is trumping her ability to respond to reasons. Thus Fischer rightly argues that in order to differentiate free and responsible action from action that results from responsibility-undermining external causes the intellect must play a primary role and there must be a certain causal history of the action.

Due to Fischer's similarities to Frankfurt, he also does not properly distinguish free and responsible actions form nonresponsible actions. Fischer is committed to the doctrine of semi-compatibilism, the belief that while moral responsibility is compatible with the truth of causal determinism, freedom traditionally conceived of as access to alternative possibilities is incompatible with the truth of causal determinism. This claim– that one can simultaneously argue that morally responsible action is distinct from action that results from responsibility-undermining external causes such as manipulation and that morally responsible action is compatible with determinism – is incoherent and ultimately fails. Stump and Pereboom's manipulation arguments sufficiently demonstrate that Fischer's notion of guidance control does not distinguish free and responsible actions from nonresponsible actions *in a deterministic context*. *All* behavior in a deterministic world is a result of causes that are out of the agent's control, and control is what Fischer

argues is necessary for moral responsibility. Furthermore, the requirements that Fischer deems sufficient for conferring moral responsibility are attainable by both globally and locally manipulated agents. Therefore Fischer succeeds in demonstrating that the only real options that can distinguish free and responsible action from action resulting from manipulation are hard compatibilism (which admits that the necessary responsibility-grounding agential structures can obtain in cases of global manipulation or *covert nonconstraining control*) and some form of incompatibilism.

Chapter Four

The most obvious place to turn next is to a form of source incompatibilism, since source incompatibilists retain one of Frankfurt's major contributions–his rejection of the principle of alternative possibilities (*PAP*)–while emphasizing that the necessary responsibility-grounding freedom-relevant agential structures are incompatible with the truth of causal determinism (a major element lacking in Fischer's account). Stump's source incompatibilism relates well to this discussion since she incorporates a revised version of Frankfurt's hierarchical mesh theory into her own theory of moral responsibility. Similar in certain respects to Fischer, Stump criticizes Frankfurt's mesh theory as inadequate for differentiating responsible action from action that results from manipulation. Furthermore, she agrees with Fischer that an adequate responsibility-grounding source component must include a primary role for the intellect. She disagrees with Fischer and agrees with Frankfurt that a mesh theory of moral responsibility is essential for distinguishing responsible action from action that results from manipulation, but only when supplemented by a *constraining* role of the intellect. Finally, she argues that the sourcehood component for responsibility ascriptions requires that the action must

stem from the agent's own intellect and will and not be traceable to external causes. She concludes, therefore, that causal determinism must be false.

Ultimately Stump's intellectualist source incompatibilism also fails to distinguish responsible action from action that results from manipulation. The main problem for Stump's account is: how does the indeterminism to which Stump is committed bring about moral responsibility in contexts where moral responsibility otherwise would be lacking?[2] Another way of putting the problem is: how does the truth of indeterminism *alone* allow for responsibility-grounding components that could not exist if determinism were true? Stump thinks that if determinism were true, then an agent's own intellect and will would not cause her action; instead, her action could be traced to some ultimate external causal factor. She agrees with Pereboom's assessment that

> In the deterministic view, the first and second-order desires and the reasons-responsive process that result in…[an agent's action] are inevitable given their causes, and those causes are inevitable given their causes. In assessing moral responsibility…[for this action] we wind our way back along the deterministic chain of causes that results in his reasoning and desires, and we eventually reach causal factors that are beyond his control—causal factors that he could not have produced, altered, or prevented.[3]

Stump agrees that if we can trace an agent's action back to sufficient explanatory causal factors beyond the agent's control, then the agent cannot be responsible for the action. Stump argues that one necessary component for grounding responsibility ascriptions is that the action stem from the agent's own intellect and will, and that the agent's own intellect and will must be the ultimate cause of the action. In the case of responsibility-

---

[2] It should be noted that on this point I partly agree with Fischer's response to Stump. Although I do not think that she simply is assuming that determinism rules out moral responsibility, there needs to be more than the truth of indeterminism (coupled with reasons-responsiveness or Stump's version of strong evaluation) to generate moral responsibility in cases where responsibility would otherwise be lacking.

[3] Pereboom, "Determinism Al Dente," p. 23.

undermining external factors, the agent's intellect is not the ultimate cause of the action. The problem with Stump's solution is that it looks much like Pereboom's compatibilistic rejoinder: an agent is morally responsible despite the deterministic causal process that led to the action because the first and second-order desires are *the agent's* desires, and those desires are generated through *the agent's* rational evaluation of the reasons. In other words, "[s]ince the causal history of his action has the right pattern…[the agent] is free and morally responsible."[4]

Stump likewise argues that the agent is responsible for her action as long as the desires are generated through the proper interaction of the agent's own intellect and will. But why can't the agent's own intellect and will produce the action in a deterministic context? In my revised case in chapter 3, $Irene_2$, I propose a scenario in which Irene acts (in a deterministic context) against her better judgment but for reasons she perceives as good. Furthermore, she feels no effectual regret for her actions (that is, regret that leads her to disassociate herself from that desire), and because her action stemmed from the proper interaction of her intellect and will (on Stump's account) she is responsible despite the fact that Irene's world is *nomically* determined. Stump must admit that Irene is not responsible since her action was nomically determined, yet according to Stump's own account it is nothing other than the truth of determinism that makes this the case. Irene's evaluative and volitional capacities were not compromised in any responsibility-undermining way, and she was not manipulated or subject to severe compulsion or phobia. The question that arises is what makes the agent's intellect and will her own in the sense necessary for moral responsibility? Stump thinks that one necessary condition

4 Ibid.

is the falsity of causal determinism. But this alone does not reveal the difference between the causal sources of Irene's action in a deterministic context versus an indeterministic one. Therefore the problem with Stump's intellectualist hierarchical theory of moral responsibility is that while she claims that it is a form of incompatibilism, her theory alone cannot properly distinguish between instances where mental actions are causally determined and instances where they are not.

Chapter Five

Like Frankfurt's voluntarist theory, Stump's intellectualist hierarchical theory of moral responsibility cannot properly distinguish causally determined action from indeterministic action because it is a time-slice theory of moral responsibility. Stump aims to solve the problem of authority that is present in Frankfurt's theory by including a robust role for the intellect, but an action that results from the proper interaction of the agent's intellect and will can still be causally determined unless there is a strong alternative possibilities condition, coupled with an historical notion of moral responsibility, and a robust role for the emotions in the moral life.

Kierkegaard's picture of human agency and the will provides the necessary freedom-relevant responsibility-grounding conditions. He has a hierarchical view of the will that is in some respects similar to Frankfurt's, but with a (modest) knowledge component (what I call his motivational cognitivism) that both Fischer and Stump agree is essential in order for the agent to be considered the source of her actions. Kierkegaard agrees with Fischer and others that the causal history of the action is essential for differentiating free and responsible action from action resulting from manipulation. Kierkegaard believes

that any account of the ethical life – the life of a person who takes responsibility for the formation of her character – must be *historical* in nature. As I will show in examples from Kierkegaard's pseudonym Judge William, the aesthete A is building a history through his choices whether he is aware of it or not. Judge William argues that A's choices form his character and solidify his self, which in turn contributes to his future actions. Though A believes that he is subject to necessity and therefore refrains from choosing, Judge William will not let A off the hook. Instead the Judge deems A responsible for making character-forming choices even as he refuses, in a sense, to choose. Therefore Kierkegaard articulates a view of freedom and responsibility that is hierarchical *and* historical. Frankfurt's early account of free will and moral responsibility includes an ambiguity as to whether and in what way an historical element is necessary. Later comments about a manipulator providing the agent with an entirely new character and that agent being responsible for the actions that stem from that new character clearly indicate that Frankfurt believes that an agent can be morally responsible *without* any historical element.[5] However, his notion of care and love requires that the agent maintain her cares over time, and for Frankfurt love is one of the highest expressions of human freedom. I argue that in this one fundamental area where Frankfurt needs the correction that Kierkegaard can offer, and that Kierkegaard's combination of a hierarchical account of the will and an historical account of agency and responsibility is uniquely suited to this task.

[5] See Frankfurt's "Response to John Martin Fischer" in *Contours of Agency*, Edited by Sarah Buss and Lee Overton, 2002, The MIT Press, Cambridge MA, pp. 27-31.

CHAPTER TWO

Harry Frankfurt on Free Will and Moral Responsibility

Introduction

In this chapter I will discuss Harry Frankfurt's contributions to the contemporary debates over what constitutes free will and what kind of freedom in necessary for moral responsibility. Specifically, I will examine his semi-compatibilism—the claim that the freedom necessary for moral responsibility is compatible with determinism while libertarian freedom is not compatible—and evaluate whether his understanding of the freedom relevant conditions necessary for grounding moral responsibility are able to distinguish cases where the action of the agent results from responsibility-undermining external forces such as manipulation, compulsion, etc., from cases whether the agent acts on her own and therefore meets at least one requirement for free and responsible action. I conclude that Frankfurt's voluntarist hierarchical mesh-theory of moral responsibility alone is unable to account for this distinction, and thus must be either modified or discarded altogether. I will argue in subsequent chapters that his hierarchical theory of moral responsibility can make this distinction with the proper modifications, namely libertarian and virtue elements found in Kierkegaard's thought. After discussing some introductory issues, the chapter will divided neatly into four main movements. First, I will summarize Frankfurt's three main contributions to this debate, which are his arguments against what he calls the principle of alternative possibilities, his notion of first and second order-desires, and his understanding of what he calls care or volitional

necessity. I will then discuss each of these three main contributions in detail. Finally, I will briefly suggest modifications to Frankfurt's understanding of free will and moral responsibility that are necessary so that his theory can usefully contribute to an adequate understanding of the freedom-relevant conditions necessary for grounding morally responsible agency.

One of the earliest and most influential discussions over what constitutes a free will and what kind of freedom is necessary for moral responsibility occurs in book III of Aristotle's *Nicomachean Ethics.* Though rudimentary in many ways, his conclusions set the stage for the next 2500 years of debate. In book III, Aristotle discusses what constitutes voluntary action and choice and opposes voluntary action to action that results from compulsion or ignorance. Aristotle cites three main conditions for voluntary (and consequently responsible) action. First, for an agent to act voluntarily, her actions must *not* be such that "the cause (or what he later calls the "moving principle") is in the external circumstances and the agent contributes nothing."[1] Although Aristotle argues that not all coercion or constraint undermines voluntariness equally, voluntary—and thus morally responsible—action requires that the agent have within herself "the 'origin' (*arche*) of the action."[2] If constraint or coercion makes the agent no longer the source of the action but rather that coercion causes and is thus the source of the action, then the agent is not acting voluntarily and hence is not morally responsible. This first condition is in line with what contemporary philosophers sometimes refer to as *autonomy* or self-

[1] Aristotle, *Nicomachean Ethics*, 1113b6. Quoted in Watson, Gary, "Free Action and Free Will". *Mind*. 96 (382) (1987), 145 (fn1).

[2] Sorabji, Richard., *Necessity, Cause, and Blame: Perspectives on Aristotle's Theory*, (Ithaca, N.Y.: Cornell University Press, 1980), 234. Quoted in Kane, Robert, *The Significance of Free Will*, New York: Oxford University Press, 1998), 33.

determination. To act freely and responsibly, the agent must determine or be the source of her actions. Someone who is forced to perform the action is not acting freely. In what follows, I call this freedom-relevant[3] condition necessary for grounding moral responsibility the *source condition*.

The second condition that Aristotle hints at is that in order for an agent to act voluntarily and thus responsibly she must have a kind of dual power: "…when acting is up to us, so is not acting."[4] Although Aristotle did not fully articulate either causal or theological determinism, or human freedom, many scholars give this "dual power" statement a libertarian interpretation. For instance, Sorabji argues that Aristotle usually cites two conditions for an agent's action being "up to us" in the sense required for moral responsibility: the source condition cited in the previous paragraph, and what Robert Kane calls the "alternative possibilities" condition. Sorabji and Kane both interpret Aristotle's dual power statements as indicating that free and morally responsible action requires that the agent have it within her power to perform the action or not perform the action.[5] In other words, in order to ascribe moral responsibility to an agent she must have the requisite control or meet what I call the *control condition*.

Finally, Aristotle claims that an agent cannot perform a voluntary action (much less make a choice) if she is ignorant of some fundamental aspect of her action. "Since that which is done under compulsion or by reason of ignorance is involuntary, the voluntary would seem to be that of which the moving principle is in the agent himself, he

---

[3] "Freedom-relevant" is a technical term commonly used in the literature. It is used to distinguish conditions necessary for grounding moral responsibility that have to do with freedom of the will from conditions having to do with other factors such as knowledge or normative competence. See footnote 7.

[4] Aristotle, *Nicomachean Ethics*, 1113b6. Quoted in Watson, "Free Action and Free Agency," 1 (fn1).

[5] Kane, Robert, *The Significance of Free Will*, (New York: Oxford University Press, 1998), p. 33.

being aware of the particular circumstances of the action."[6] Aristotle's account suggests that an action cannot be free (and thus contribute to the agent's virtue and character) if the agent lacks the requisite knowledge. As Kevin Timpe puts it, in order for the agent to meet what he calls the "epistemic requirement" necessary for grounding moral responsibility, the agent cannot be "ignorant of the relevant facts or else that there be something that she should have done at some earlier time such that, had she done it, she would not now be ignorant of the relevant moral facts."[7] Thus the third requirement is that the agent must have some awareness of the relevant moral facts. I refer to this as the *epistemic condition*.

On the surface the epistemic condition may not seem like a *freedom-relevant* condition, but I show that it is connected to the other two freedom-relevant conditions in important ways. For instance, an agent cannot have the *control* over her actions necessary for moral responsibility unless she is aware of certain relevant moral facts and has a certain kind of connection to the external world that requires a level of epistemic awareness. However, there may well be epistemic concerns related to moral responsibility that are not freedom-relevant; I will not discuss those concerns but only concerns related to the freedom-relevant conditions of sourcehood and control. In the context of my thesis it is most accurate to say that the epistemic condition is an *indirect* freedom-relevant condition while the other two conditions—source and control—are *direct* freedom-relevant conditions.[8]

---

[6] Aristotle, *Nicomachean Ethics*, book III chapter 1 (find number). The Internet Classics Archive, http://classics.mit.edu/Aristotle/nicomachaen.3.iii.html. Accessed May 13, 2010.

[7] Timpe, Kevin, *Free Will: Sourcehood and its Alternatives*, (London: Continuum ,2008), 10.

[8] Gary Watson suggests a fourth condition that is not included in the freedom-relevant conditions, namely the issue of "normative competence" or what Susan Wolf calls "sanity." I actually think that sanity

Many of the most prominent contemporary accounts of the freedom-relevant conditions necessary for grounding moral responsibility incorporate, or at least account for, these three components in some way. However, many of these accounts differ from Aristotle's account because he is not interested only in the freedom-relevant conditions of particular actions,[9] but also in the conditions necessary for the formation of *character*. Virtue and vice are states of character for which the agent is responsible. Voluntary action and choice play a vital role in Aristotle's account of the cultivation of virtue and vice as characteristics, because "by choosing what is good or bad we are men of a certain character."[10] Aristotle is not only interested in ascribing praise and blame to agents for particular actions, but in ascribing praise and blame for the kind of character that results from habitual actions that shape character. In other words, Aristotle is laying out the conditions necessary for taking responsibility for the self, or for what I call morally responsible agency. Following Aristotle, I argue that any adequate theory of the freedom-relevant conditions necessary for grounding moral responsibility must deal not only with the three conditions of 1) source (the agent must determine her actions in the relevant sense), 2) control (the agent must be able to guide herself along the path she wishes to traverse) and 3) knowledge (the agent must have some kind of fundamental epistemic connection to both her own mental states and the external world), but also must

---

is also an indirect freedom-relevant condition since the agent must have a certain level of psychological integration or sanity in order to have the freedom that is required in order that praise or blame be ascribed to her actions. The inclusion of this condition would make this project too broad, however, and therefore I will only discuss it incidentally.

[9] Throughout this dissertation, unless otherwise specified, "action" will be used in the broad sense to include mental actions such as intentions, choices, etc.

[10] Aristotle, *Nicomachean Ethics*, book III chapter 2.

take into account how these three components contribute to the cultivation of the agent's character.

I find that Kierkegaard offers a compelling account of free will and moral responsibility that includes all these elements. Furthermore, I argue that he can help solve a number of problems in the contemporary literature on free will. However, Kierkegaard's account is unsystematic and difficult to grasp. Harry Frankfurt's notions of higher-order desires and volitional necessity lend clarity to Kierkegaard's account. The interaction of contemporary scholars with Kierkegaard yields a rich account of human freedom that includes the freedom-relevant conditions necessary for the cultivation of character that makes his voice relevant to the contemporary discussion.

Frankfurt's Understanding of Freedom and Responsibility

Harry Frankfurt's work on free will and moral responsibility has altered the way philosophers think about these issues. He introduced three crucial ideas in three separate essays that shifted the focus of the debates. In his seminal article, "Alternative Possibilities and Moral Responsibility," published in 1969,[11] he introduced now infamous counterexamples[12] against what he termed the principle of alternate[13] possibilities or *PAP*. *PAP* maintains that in order for a person to be held morally responsible for her actions she must have the ability to act in a different way than she did. Whatever one

---

[11] Frankfurt, "Alternative Possibilities and Moral Responsibility," *The Journal of Philosophy*, Vol. 66, no. 23. (Dec. 4, 1969), 829-839. Reprinted in *The Importance of What we Care About*, pp. 1-10.

[12] Frankfurt's counterexamples are discussed in section III below.

[13] Although Frankfurt used the term "alternate" and even later defended his use of that term, almost all other philosophers refer to "alternative" possibilities since it makes more sense to say that the debate is about whether or not agents have alternatives when willing freely, and not about alternates to a preferred action. After all, we often are not sure which action we prefer, and that is why we deliberate. I follow general usage and use "alternative possibilities."

thinks of the efficacy of the counterexamples, by introducing them Frankfurt shifted the focus of discussions of moral responsibility to what makes an action free in the actual sequence of events as opposed to the alternative or counterfactual sequence.[14]

In his "Freedom of the Will and the Concept of a Person" published in 1971, Frankfurt introduced his hierarchical understanding of freedom which, although not new,[15] has given more precise language to centuries of debate about what it means to say that a person's actions proceed from a will that is free. Frankfurt argued that being a person means having the ability to form higher-order desires or volitional states about one's basic or first-order desires. Furthermore, his notion is broadly compatibilistic in that it can be reconciled with a deterministic view of the world.[16] Both libertarian and compatibilist theories have benefitted and evolved from the conversation that Frankfurt initiated.

Finally, in "The Importance of What we Care About" published in 1982, Frankfurt introduced his notions of caring and volitional necessity. Here he argues that a person can come to care so much about something that "it is impossible for him to forbear from a certain course of action."[17] A person who cares in this way and to this extent is driven by an irresistible or constraining passion and finds that he has no choice but to accede to that force, yet he accedes because he is committed to the object of his

---

[14] By actual sequence I mean that sequence of events in which the action occurred. This is opposed to the counterfactual sequence in which a different sequence of events issues in a different action.

[15] For instance, Stump claims that the hierarchical view is found in Aquinas and Augustine, albeit with crucial differences. See chapter four for a detailed discussion.

[16] I argue that his view does not have to be combined with any sort of determinism. However, it seems to have been Frankfurt's intent to come up with an articulation of freedom of the will that is both true to our experience and intuitions and also compatible with determinism.

[17] Frankfurt, "The Importance of What We Care About," 86.

care and cannot imagine letting it go. Finally, because this caring is something that the person actually wants, this "[v]olitional necessity may have a liberating effect…"[18] In other words, Frankfurt intriguingly suggests that a person who is unable to resist the force of her own care(s) may in fact be experiencing the summit of human freedom.

Taken together, these three elements have altered the trajectory of the free will debate. Any systematic account of what it means for a person to will freely and responsibly must deal with Frankfurt's provocative, if underdeveloped, notions of actual sequence freedom, the hierarchical or structuralist account of the will, and his notion of caring or volitional necessity. However, compatibilists and incompatibilists alike have raised objections to all the main elements of Frankfurt's account of freedom and moral responsibility. Perhaps the most ink has been spilled in response to Frankfurt's famous counterexamples to *PAP*. The counterexamples are designed to show that an agent can be morally responsible despite the lack of alternative possibilities. Incompatibilists argue that Frankfurt's construction of the counterexamples leaves room for the agent to form, for instance, intentions to act in one way or another before the manipulator intervenes (these examples will be explored at length below). Despite Frankfurt's argument to the contrary, incompatibilists have found what John Martin Fischer calls "flickers of freedom" in the counterexamples. In response, compatibilists have sought to refine the counterexamples so as to eliminate these flickers of freedom. Two main issues have arisen from this debate. First, whether an *FSC* (Frankfurt-style case) can be constructed that eliminates these flickers of freedom is questionable. Even if such a case cannot be constructed, compatibilists such as Fischer argue that whatever flicker of freedom is

[18] Ibid., 88.

available to the agent in an *FSC* is not sufficiently robust to ground ascriptions of moral responsibility.

The second issue that has arisen out of the debate over *FSCs* is a distinction between two different kinds of arguments for the incompatibility of determinism and moral responsibility. Classically, incompatibilists have argued that determinism is incompatible with moral responsibility because moral responsibility requires alternative possibilities and determinism is assumed to eliminate alternative possibilities. Causal determinism is taken to eliminate alternative possibilities because if causal determinism is true then the unalterable events of the past combined with the fixed natural laws result in the absolute determination of every future event. Arguments for incompatibilism that claim that causal determinism rules out moral responsibility by eliminating alternative possibilities have come to be known as indirect arguments for incompatibilism.

Some incompatibilists think that FSCs successfully show that moral responsibility is at least possible without the agent having access to alternative possibilities, but these incompatibilists are still convinced that moral responsibility is incompatible with determinism; therefore, the reasons for their incompatibilism rest on the argument that causal determinism *directly* rules out moral responsibility. The direct argument can take a variety of forms, but usually goes something like this: In order for an agent to be morally responsible for her actions, she must be the ultimate source of those actions. However, if causal determinism is true, then whenever we trace the causal source of an agent's action, we will find that the causal history stretches beyond the agent to sources that are ultimately out of the agent's control. Since the ultimate cause of the actions is outside the agent, she is not the source of her actions and therefore cannot be held

morally responsible. The agent does not have the freedom that the incompatibilist thinks is necessary for grounding ascriptions of moral responsibility because she is not free to be the source of her actions. Thus proponents of this view conclude that determinism *directly* rules out moral responsibility without reference to alternative possibilities. The direct argument poses a particular challenge to Frankfurt and other Frankfurt-style compatibilists because it is compatible with the conclusion of FSCs yet argues against Frankfurt's implied conclusion that an agent can be morally responsible in a deterministic world.

Perhaps the most famous objection responds to Frankfurt's hierarchical account of the will. Frankfurt claims that in order to be a person who acts with free will one must have higher order volitions that endorse first-order volitions, and that this higher-order endorsement makes a person's volitions authoritative and freedom-granting, thus distinguishing a person from a creature lacking higher volitions and thus lacking freedom. A number of philosophers, perhaps the most famous being another compatibilist Gary Watson, have objected to this hierarchical account of freedom claiming that nothing inherent in this model prevents an *infinite regress* of desires. In other words, if an agent can have second-order desires, then why does she not have third or fourth-order desires or beyond? If a person can have the desire to desire A, then what is preventing her from having the desire to desire to desire A, and so on? This objection leads to the more crucial related objection sometimes referred to as the problem of *authority*. Nothing in Frankfurt's (especially early) account articulates how exactly a second-order desire (or any higher-order desire) becomes more authoritative than a first-

order desire.[19] Furthermore, why are desires about desires constitutive of a person? Frankfurt argues that higher-order desires are authoritative because they are desires with which a person decisively identifies and about which a person cares. These desires distinguish a person from a wanton who "has no preference concerning which of his first-order desires is to be his will…"[20] However, why does identifying with a second-order desire make a person less wanton and more certain of what he wants? If he is wanton with respect to his first-order desires, he may be wanton with respect to his second-order desires, and so forth. Taken together, these two objections point to the need to clarify exactly what Frankfurt means when he says that by identifying with certain desires and rejecting others, an agent reveals her freedom and takes responsibility for her actions.[21]

Furthermore, in clarifying what he means by identification, Frankfurt claims that "[t]he higher-order attitudes that are formed in processes leading to identification involve 'evaluations' only in a sense that is strictly value-neutral."[22] Here as in many other places, Frankfurt makes clear that no knowledge component or epistemic condition forms the moral responsibility grounding higher-order volitions, other than the person's awareness of what she cares about and that she "gets behind" that desire. John Davenport calls this Frankfurt's "existential subjectivism." This idea has lead to a number of

---

[19] Watson (1975 p. 215) answered his own now famous objection by saying that authoritative desires are desires that stem from the agent's "valuational system" or the agent's all-things-considered judgment about what is best in the particular situation. Watson (1987) later criticized his own schema for being too rationalistic and unable to account for addictions, phobias, and so forth. Inherent in his critique of Frankfurt is the notion that Frankfurt is actually Humean in his understanding of the nature of desire and its relationship to reason, and on this point Watson is certainly correct.

[20] Frankfurt, "Freedom of the will," 21.

[21] Several other objections to be dealt with later include the ahistorical nature of Frankfurt's hierarchical account, his motivational noncognitivism, and his lack of any account of moral self-cultivation.

[22] Frankfurt, "Reply to Michael E. Bratman," in *Contours of Agency*, edited by Sarah Buss and Lee Overton (The MIT Press: Cambridge MA, 2002), 87.

criticisms and modifications to Frankfurt's view of caring, including Davenport's "existential objectivism" and John Martin Fischer's notion of "reasons-responsiveness." This objection applies to Frankfurt's notion of higher-order volitions and identification, as well as to his concept of care or volitional necessity.[23]

Frankfurt's Counterexamples Against PAP

In his groundbreaking article "Alternate Possibilities and Moral Responsibility," Frankfurt attempts to defend the claim that moral responsibility is compatible with determinism. Frankfurt agrees with Aristotle that the cause or *arche* of morally responsible action has to lie within the agent; but he denies Aristotle's second condition, namely, that the agent must be able to perform or refrain from performing the action.[24] These two conditions respectively can be called the source and leeway[25] conditions. Frankfurt calls the leeway condition "the principle of alternate possibilities" and subsequently defines *PAP* as the claim that "a person is morally responsible for what he does only if he could have done otherwise."[26] In other words, *PAP* is the claim that in order for an agent to be considered responsible for action X, she must have had some leeway in regards to X. She must have been able to perform that action or another action, or at least have been able to avoid performing that action (therefore this is sometimes referred to as the avoidability condition). If the action was causally determined, then there can be no leeway. Frankfurt claims that the majority of philosophers—libertarians

[23] I deal with these objections in the subsequent chapters.

[24] Kane, *The Significance of Free Will*, 34.

[25] Leeway is a stronger version of the control condition. For instance, John Martin Fischer argues that an agent can meet a weaker control condition without meeting the stronger leeway condition.

[26] Frankfurt, "Alternate Possibilities and Moral Responsibility," 1.

and compatibilists alike—affirm some form of this principle. Furthermore, he claims that many if not most philosophers believe that a coerced action cannot be a free action because coercion eliminates the possibility of leeway and therefore coercion is simply a more particularized version of *PAP*.

Frankfurt proceeds to deny *PAP* by first denying that coercion is a version of *PAP*. He imagines three scenarios is which the agent, Jones, is coerced by means of punishment into performing an action *X* that he already has decided to perform for reasons of his own.[27] The coercion is such that any reasonable person would succumb to the pressure and perform the action. In the first instance, Jones is not a reasonable person and does what he has decided to do no matter what; therefore the coercion exerted no influence on him. Against *prima facia* intuitions, Frankfurt argues that if Jones did what he wanted despite the threat of torture, then this is not a case of coercion, because the external force had no influence on the Jones's actions. Since the external force did not prevent him from not performing action *X*, the agent seemingly had alternative possibilities and therefore was morally responsible. In other words, since Jones is unreasonable, he as easily could have refused to perform *X* as perform X. Coercion in this first, albeit unusual, instance does not preclude the existence of alternative possibilities. In the second instance Jones is affected so profoundly by the threat that he forgets what he previously decided to do. He is so upset by the threat that he performs *X* only because he is threatened, not because of his previous decision. Frankfurt concludes that Jones is not morally responsible because his action is the result of coercion. In other words, he does not meet Aristotle's condition for responsible action because he is not the source of his action; rather, the causal history of the action can be traced to the torturer.

[27] Ibid., 3.

In the third scenario the agent is "neither stampeded by the threat nor indifferent to it."[28] The threat affects him to the point that he would have performed *X* even if he had not previously made the decision on his own. However, this agent performs *X* "on the basis of the decision he had made before the threat was issued."[29] In other words, he is not motivated by the threat but by his own decision, so Frankfurt concludes that in this third scenario Jones meets Aristotle's source condition and is therefore responsible for his action despite his lack of alternative possibilities. Furthermore, Frankfurt argues that in the third scenario Jones was *not* coerced because the coercion did not play a causal role in his action. He acted on his own, for reasons of his own. However, Frankfurt does admit that this is not a decisive counterexample to *PAP* because one could argue that morally relevant alternative possibilities were available to Jones despite the fact that he would have performed X no matter his choice.

However, Frankfurt thinks that a decisive counterexample can be formulated as follows:

> Suppose someone—Black, let us say—wants Jones to perform a certain action. Black is prepared to go to considerable lengths to get his way, but he prefers to avoid showing his hand unnecessarily. So he waits until Jones is about to make up his mind what to do, and he does nothing unless it is clear to him (Black is an excellent Judge of such things) that Jones is going to decide to do something *other* than what he wants him to do. If it does become clear that Jones is going to decide to do something else, Black takes effective steps to ensure that Jones decides to do, and that he does do, what he wants him to do. Whatever Jones's initial preferences and inclinations, then, Black will have his way… Now suppose that Black never has to show his hand because Jones, for reasons of his own, decides to perform and does perform the very action Black wants him to perform. In that case, it seems very clear, Jones will bear precisely the same moral responsibility for what he does as he would have borne if Black had not been ready to take steps to ensure that he do it. It would be quite unreasonable to excuse Jones for his action, or to withhold the

[28] Ibid., 4.

[29] Ibid.

> praise to which it would normally entitle him, on the basis of the fact that he could not have done otherwise."[30]

Jones's action *X* is something that he does on his own, but is also something that Black would have caused him to do if he had shown somehow that he was not going to do *X*. A number of salient points come from this short passage, and for each point, numerous potential problems and objections. Frankfurt's goal is to envision a scenario in which the agent is responsible for her actions but cannot refrain from the act or perform a different action (and of course refraining would be a different action). In other words, Frankfurt argues that since Jones is the cause of his action, he is responsible. And if another agent (or even a natural cause) makes it such that the agent could not perform a different action, as long as the agent still performs the action on her own, for reasons of her own, she is responsible. On the one hand this is not a revelation. Many philosophers and theologians through the centuries have thought that the freedom necessary for ascriptions of moral responsibility is compatible with various kinds of determinism, particularly the determinism of a God that is omniscient and omnipresent. However, Frankfurt's scenario shows this potential compatibility without reference to such mysterious notions as the character of God. Given the advances of modern science it seems possible that a person could monitor another agent's activity and be able to predict that agent's actions. If prediction is possible, then an agent might well act freely and responsibly without access to alternative possibilities. Furthermore, if the agent can act freely and responsibly without access to alternative possibilities, then moral responsibility may be compatible with determinism.

---

[30] Ibid., 6, 7.

If Frankfurt's case of Jones and Black is successful, then the primary point is to show that freedom-relevant conditions necessary for grounding ascriptions of moral responsibility do not include alternative possibilities. Furthermore, if morally responsible action does not require alternative possibilities, then Frankfurt's case has removed a major barrier to the compatibility between moral responsibility and causal determinism. Causal determinism is generally taken to be "the claim that a complete statement of the laws of nature and a complete description of the (temporally nonrelational or 'genuine') facts about the world at some time *T* entail every truth about the world after *T*."[31] If this definition of causal determinism is right, then the clear consequence of causal determinism appears to be the elimination of alternative possibilities, a claim that is referred to as the "consequence argument" due to van Inwagen's formulation:

> If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what event went on before we were born, and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (including our present acts) are not up to us.[32]

In other words, any choice *C* made by an agent at time T_2 is logically entailed by the state of the world at time T_1 together with the laws of nature. In order for the agent to make a different choice C_2 at T_2 either something about the past would need to change or some natural law would have to change. Since neither of these is the case (or at least it is not up to the agent to change the events of the past or the relevant laws) the agent could not have chosen otherwise at time T_2.[33] Fischer points out that some philosophers do not find

---

[31] Fischer, John Martin, "Recent Work on Moral Responsibility," Ethics. 110 (1999), 99.

[32] Van Inwagen, Peter, *An Essay on Free Will*, (Oxford: Clarendon Press, 1983), 16. Quoted in Mele, *Autonomous Agents*, 195.

[33] Fischer, John Martin, "Recent Work on Moral Responsibility," 99-100.

the consequence argument sound; but, many (including Fischer) do find some version or another to be sound, and if one is convinced by the consequence argument then only a few conclusions can be drawn about the relationship between freedom and causal determinism.

The most obvious implication of the consequence argument for the relationship between freedom and determinism is that if causal determinism obtains, then human beings never have access to alternative possibilities, no matter what phenomenological human experience of the world indicates. Some philosophers (like hard incompatibilists such as Derk Pereboom) conclude that if causal determinism obtains then human beings are not morally responsible for their actions since moral responsibility requires features such as alternative possibilities that are incompatible with causal determinism.[34] One can draw the opposite conclusion from the consequence argument—collective phenomenological human experience is impossible to reconcile with the truth of causal determinism. Human agents experience the world as open to some extent; and, human agents experience the world as containing real alternative possibilities, as a result of those choices we make a distinct mark on or difference to the world. If causal determinism were true and some form of the consequence argument is valid, then we are radically mistaken in the way that we subjectively encounter the world and thus one conclusion to draw is that causal determinism has to be false.

These two extreme responses to the consequence argument are not the only plausible responses. Philosophers such as Fischer who are persuaded by both the

[34] I will discuss Pereboom's "hard incompatibilism" in chapters two and three, but it is worth noting here that Pereboom is convinced by FSCs that alternative possibilities are not a freedom-relevant condition necessary for grounding moral responsibility. Therefore he thinks that what is eliminated by causal determinism (or indeterminism that entails randomness at the microlevel) is the control that is required for moral responsibility.

consequence argument and FSCs go in a different direction and attempt to flesh out just what kind of freedom is required for moral responsibility and conclude that perhaps this freedom is distinct from leeway (AP) freedom. In other words, Fischer believes that if causal determinism is true, the consequence is that agents do not have access to genuinely open alternative possibilities. However, he also is convinced that Frankfurt has succeeded in constructing a case that generates the intuition of moral responsibility without alternative possibilities. In one sense, Jones acts freely when he performs *X* because he was not coerced, forced, etc., and because he did what he wanted. Furthermore, Jones acts as if he has genuinely accessible alternative possibilities, even though in reality he does not (if the argument is successful). Therefore, Fischer argues that the freedom-relevant features necessary for grounding moral responsibility are found not in the agent's access to some alternative or counterfactual scenario, but in the actual sequence in which the action occurs. Fischer refers to himself as a semi-compatibilist because he believes that while what we normally mean by human freedom—access to alternative possibilities—is incompatible with causal determinism, a very important kind of freedom is compatible with determinism, namely, the freedom to guide one's self along a certain path in a certain way. Fischer argues that this second kind of freedom—which he labels "guidance control" —is necessary for moral responsibility and is compatible with causal determinism.[35]

However, there are also incompatibilists who are persuaded by the consequence argument and believe that FSCs are successful in showing that alternative possibilities are not necessary for morally responsible action and that the locus of the freedom-relevant

[35] I will discuss and critique Fischer's semi-compatibilism and his understanding of guidance control in chapter two.

conditions necessary for grounding moral responsibility are found in the actual sequence. These incompatibilists disagree with Frankfurt that the freedom necessary for moral responsibility is compatible with determinism.[36] This disagreement has led to the development of two distinct arguments in favor of the incompatibilism of determinism and moral responsibility. The traditional argument that Frankfurt seeks to reject is now referred to as the indirect argument because it states that moral responsibility requires alternative possibilities and therefore causal determinism eliminates moral responsibility via its elimination of alternative possibilities (i.e. the consequence argument). Incompatibilists such as Eleonore Stump and Derk Pereboom, who agree with Frankfurt that alternative possibilities are not a freedom-relevant condition for ascriptions of moral responsibility, still think that moral responsibility is impossible if causal determinism is true. Incompatibilists such as Stump and Pereboom think that the freedom-relevant requirement for moral responsibility is that the agent be the source of her actions (Aristotle's first requirement). They think that if the ultimate cause of the agent's actions is extrinsic to the agent—such as another agent, a natural cause, or God—then that extrinsic cause is the source of the action and not the agent and therefore the agent is not morally responsible. This is a direct argument for the incompatibilism of causal determinism and moral responsibility—direct because it makes no reference to alternative possibilities. The direct argument has opened a new area of debate between Frankfurtians and his critics.

[36] Stump, Pereboom, and Michael McKenna are all examples of source incompatibilists that I will discuss at length in chapter four.

Frankfurt's Hierarchical Actual-Sequence Theory of Moral Responsibility[37]

Frankfurt's second major contribution to the issues surrounding the freedom-relevant conditions necessary for grounding moral responsibility is his hierarchical or structuralist theory. Generally speaking, Frankfurt argues that what makes a person free is that the person wants what she wants to want; i.e. the person endorses or identifies with her desires. Several problems emerge from this hierarchical account, namely, the problems of infinite regress and authority. In other words, what prevents a person from having an infinite number of higher-order desires; and, what makes a second or higher-order desire authoritative or freedom-granting in a way that first-order desires are not? Note that the problem of authority is the fundamental problem here, but it is intimately tied to the infinite regress problem. The question is what makes the agent the source of her free and responsible actions. Frankfurt responds that it is the right mesh of desires, namely a higher-order desire that reflectively endorses a first-order desire. But this hierarchical account opens up the possibility of an infinite number of reflective desires. Without something to cut off this infinite regress the hierarchical theory has failed because the agent may never act at all, and if so the action may be purely arbitrary. There must some element present in the mesh theory that accounts for the agent's ability to make responsible choices. Here we see the connection to the authority question. If the agent's higher-order desires are fundamentally connected to her evaluative faculties, then both problems are solved at once because her evaluative faculties give the her the resources to cut off the regress and makes her (at least potentially) the source of her action in the sense required for moral responsibility. Frankfurt's answer to these

[37] Here I am following Fischer in calling Frankfurt's account an "actual-sequence" theory. See Fischer, "Recent Work on Moral Responsibility," 125.

problems is that a person with free will wholeheartedly *identifies* with her desires. Frankfurt often uses the word wholehearted as a virtual synonym for volitional necessity, so the concepts of higher-order desires and identification are connected closely to the concept of volitional necessity.

Frankfurt first elucidated his hierarchical account in his 1971 article, "Freedom of the Will and the Concept of a Person."[38] Unlike his article on alternative possibilities, where his main concern is moral responsibility, in this article Frankfurt is concerned primarily with the volitional structure or features of free agency. Frankfurt thinks that some amount of freedom is necessary for moral responsibility, but that a person need not act with free will in the most complete sense to be morally responsible. He suggests that what separates human persons from higher functioning animals is the unique ability to form higher-order desires and volitions about other motivational states.[39] As Alfred Mele puts it, "a higher-order desire is a desire whose representational content encompasses a representation of another actual or possible desire of the person whose desire it is."[40] The representational content of a higher-order volition is generally a first-order or basic motivational state (in this specific case a desire), while the condition of satisfaction of a higher-order desire is the actualization of the first-order desire that is represented in the higher-order desire.

---

[38] Frankfurt, Harry G., "Freedom of the Will and the Concept of a Person," *The Journal of Philosophy*. 68 no. 1 (1971): 5-20.

[39] In general I will use "higher-order desires" to refer to a desire about another desire; however, the term can refer generically to any mental state about another mental state. Although Frankfurt does not spend much time on the subject, it is clear from some of his comments that he thinks that agents form higher-order mental states about a variety of mental states; and, in later articles he will sometimes refer to higher-order attitudes.

[40] Mele, *Autonomous Agents*, 65.

Frankfurt argues that human freedom requires that agents not only have higher-order desires but also volitions. He specifies two varieties of higher-order desires, since a second order desire can either be the desire "simply to have a certain desire" or the desire for a certain desire to be one's will.[41] In the first instance, a person has desires (second or higher-order) about other desires (first-order) but the individual either is not concerned with whether or not the first-order desire is effectual or perhaps wishes to have the first-order desire but does not want it to be effectual.[42] In the second instance, the person has a desire about another desire and wants that first-order desire to be her will but this second-order desire is thwarted due to a conflicting second-order desire. In this instance, the person might have mutually exclusive desires or loyalties that cannot both be fulfilled despite the fact that she wants to have both desires.[43] In a third instance, the individual has a desire about a desire and wants that desire to be effectual and the first-order desire is effectual (in other words she carries out that desire whether that be the formation of an intention or an attempt at physical action, etc.).[44]

---

[41] Frankfurt, "Freedom of the Will," 16 (book).

[42] Frankfurt gives the fascinating example of a doctor who wants to understand his patient's drug addiction and thus wishes to have his patient's addiction desire (first-order) but does not actually want to take drugs. He just wants to understand what addiction feels like. Therefore the doctor has second-order desires but not volitions (at least in relation to this particular desire). Needless to say this is probably not very common. See Frankfurt, "Freedom of the Will and the Concept of a Person," 16 (book).

[43] This is an instance of what Frankfurt calls "ambivalence" and is discussed below. Part of what's at issue here is the fact that Frankfurt's specific construal of the hierarchical account cannot help the person who is truly ambivalent. I will argue that with the right modifications a hierarchical account is particularly useful in instances like these.

[44] Eleonore Stump usefully divides Frankfurt's first and second-order desires into four categories: first-order desires and volitions on the one hand, and second-order desires and volitions on the other. Thus an agent can have conflicting first-order desires (D_1), for instance, to indulge in a piece of cake for dessert while at the same time desiring to hold to the Lenten commitment to refrain from sweets. If the agent wants to want to hold to her Lenten commitment, then she has a second-order desire. If her second-order desire is effectual and she follows her desire to hold to her Lenten commitment and refrains from eating the cake, then she has both a first-order volition (V_1) not to eat the cake, as well as a second-order volition (V_2) not to eat the cake. But she also has a D_1 desire to eat the cake. Thus whenever a desire D_1 or D_2 is

Thus Frankfurt specifies three different varieties of second order desires, two of which are actually second-order volitions. The first variety—a person who simply has desires about her desires but is not interested in whether or not those desires are effectual—does not have second-order volitions because she is ultimately *wanton* in regards to the constitution of her will. A wanton "does not care about his will. His desires move him to do certain things, without its being true of him either that he wants to be moved by those desires or that he prefers to be moved by other desires."[45] Furthermore, a wanton can display the reflective capacity necessary for higher-order desires, but does not take an active role in self-constitution and therefore is not displaying the distinctly human characteristic of higher-order volitions.

> What distinguishes the rational wanton from other rational agents is that he is not concerned with the desirability of the desires themselves. He ignores the question of what his will is to be. Not only does he pursue whatever course of action he is most strongly inclined to pursue, but he does not care which of his inclinations is the strongest.[46]

Thus, a wanton can be highly reflective and rational displaying complex deliberation over how to fulfill her desires. For instance, she might reflect on her desires and form second-order desires. She may decide, through reflection on her desires, to pursue the strongest of those desires and figure out how to attain the goal of fulfilling those desires. However, she does not reflect on whether or not she wants those desires to constitute her will. She does not reflect on the kind of person that she wants to be and whether or not the desires that she is pursuing will help her become that kind of person. The wanton may reflect,

---

effectual it becomes a volition. This of course is not quite in line with Frankfurt's own delineation, so I will stick with Frankfurt's account at this juncture. See Stump's "Sanctification, Hardening of the Heart, and Frankfurt's Concept of Free Will," *The Journal of Philosophy* 85:8 (1988): 401.

[45] Frankfurt, "Freedom of the Will and the Concept of a Person," 16 (book).

[46] Ibid., 17 (book).

but "ignores the question of what his will is to be;" therefore, the wanton lacks what Frankfurt calls higher-order volitions.

In the other two cases, where the individual has higher-order desires, those desires are volitions because whether or not the desires are effectual, the person takes an active role in the constitution of her will. She cares which inclinations are strongest even if she cannot determine which inclination is strongest. In other words, she is not wanton in regards to her desires. Frankfurt usefully illustrates these three different forms of higher-order desires with examples of three addicts. All three addicts have a "physiological addiction accounting for their condition" and they all "succumb inevitably to their periodic desires for the drug to which they are addicted."[47] The first addict is addicted to the drug and may have conflicting desires to take the drug and refrain from taking the drug. Furthermore, he may even have higher-order desires—desires about his desire for the drug—but he does not take a stance with regard to those desires. "His actions reflect the economy of his first-order desires, without his being concerned whether the desires that move him to act are desires by which he wants to be moved to act."[48] He acts because of his addiction and his desires cannot be otherwise. Nonetheless, he does not decide whether or not he even *wants* those desires to be different than they are. This first addict is a *wanton*.

The second addict suffers from the same compulsion as the first; and, like the first, because of his addiction he cannot act otherwise than he does. However, because

[47] Ibid.

[48] Ibid., 18.

the unwilling addict chooses to endorse his desire to refrain from taking the drug, he has second-order volitions as well.

> The unwilling addict has conflicting first-order desires: he wants to take the drug, and he also wants to refrain from taking it. In addition to these first-order desires, however, he has a volition of the second order. He is not neutral with regard to the conflict between his desire to take the drug and his desire to refrain from taking it. It is the latter desire, and not the former, that he wants to constitute his will; it is the latter desire, rather than the former, that he wants to be effective and to provide the purpose that he will seek to realize in what he actually does.[49]

The unwilling addict has conflicting desires that the wanton addict may or may not have. More importantly, the unwilling addict takes a stance in regard to which desire he identifies with; that is, he wants to want to refrain from taking the drug. Needless to say, his second-order volition is ineffectual, but he is a person in Frankfurt's sense because he takes a stance in regards to his desires. Finally, the unwilling addict is distinguishable from the wanton addict because he is responsible for his actions due to his second-order volition whereas the wanton is not responsible.[50]

The third example is that of a willing addict. The willing addict is like the first two in that he has the compulsive desire to take the drug and that desire would be effectual whether or not the addict wants it to be. However, in this case the addict fully endorses his addiction. He is so committed to his addiction that if it were to fade he would take steps to rekindle it. Like the unwilling addict, the willing addict is a person and acts responsibly because he takes an active interest in which desires constitute his will.

---

[49] Ibid., 17-18.

[50] I discuss this point below.

Both the unwilling and willing addicts have higher-order volitions. Whereas the unwilling addict clearly does not possess freedom of will because he is not free to make the desire that he wants his will, the question is more complicated in the case of the willing addict. On the one hand, the willing addict cannot help but desire to take the drug and therefore he is not free because he cannot constitute his will the way that he wants. He is not free to refrain from taking the drug. On the other hand this addict *wants* to take the drug and when he takes it "he takes it freely and of his own free will."[51] Frankfurt calls this a case of "overdetermination" of a first-order desire and concludes that the willing addict does not act with free will since he is not able to will other than how he actually does and therefore "[h]is will is outside of his control…"[52]

Because the willing addict's will is overdetermined and he is both free and unfree when he wills to take the drug he is not the best example of a person with a successful higher-order volition. Instead, the best example of a person with successful higher-order volitions would be someone not in the grips of any addiction. The person with a truly free will meets three requirements. The first two requirements—which are co-extensive

---

[51] Ibid., 25.

[52] Ibid. I find this conclusion inconsistent with Frankfurt's notion of volitional necessity discussed below as well as with comments he makes in later papers and lectures. Frankfurt is consistent throughout his writings that the willing addict is responsible because his will conforms to his higher-order volitions and thus he wants what he wants to want, namely the drug to which he is addicted. Frankfurt is less consistent on whether or not the willing addict has free will when he takes the drug. As noted, he says in "Freedom of the Will" that the willing addict is *not* free because his desire is effectual due to a physiological compulsion and therefore his will is outside of his control. Yet as we will see below, the person who wills something with volitional necessity also has a will that is out of her direct voluntary control; but, because she has the will she wants to have, is satisfied with, and does not seek to change it, she wills freely. The only difference I can see between Frankfurt's account of a person who wills with a free (but volitionally necessitated) will and one who wills both because he is addicted and he wants to will that way is that addiction is like a kind of coercion or manipulation. But that cannot be right because then we would have to conclude that volitional necessity is also like manipulation, which Frankfurt clearly does not believe. Thus this example appears to obscure Frankfurt's definition of freedom of the will. Therefore I argue that the only way to distinguish between free will and a will that is the result of addiction, coercion, or manipulation is to look at the history of how the agent acquired the volitional necessity. If the addict intentionally formed his addiction through his own free choices then his will is no more overdetermined than the agent who is subject to volitional necessity.

with the requirements for moral responsibility are the agent's ability to want what she wants to want (form higher-order desire), and to make the desire that she wants to want her will (higher-order volitions). Frankfurt states that "…the notion of the will… is not coextensive with…the notion of something that merely inclines an agent in some degree to act in a certain way. Rather, it is the notion of an *effective* desire…(not merely) what an agent intends to do."[53] The willing addict meets this requirement because he is able to make the desire that he wants to want his own. Yet, he is not able to make a desire other than the one that he actually has his own. Frankfurt states the third requirement clearly in the following passage:

> A person's will is free only if he is free to have the will he wants. This means that, with regard to any of his first-order desires, he is free either to make that desire his will or to make some other first-order desire his will instead. Whatever his will, then, *the will of the person whose will is free could have been otherwise*; he could have done otherwise than to constitute his will as he did.[54]

Here Frankfurt claims that in order to will with a will that is truly free, the agent must be able to will otherwise than she actually did. The agent must have what I call *motivational alternative possibilities*.[55] She must be able to constitute her will in a different way than she did by choosing to make another desire her will. Despite the fact that the willing addict possesses effectual higher-order volitions because he is able to make the desire that he wants to want his will, he does not possess freedom of the will because he does

---

[53] Ibid., 14 (book). Italics in original.

[54] Ibid., 24 (book). This passage is controversial and subject to interpretation that need not be delineated here. It is important to note, however, that this passage is consistent with a modest libertarian approach where the agent can select from various internal motivational states. I will deal with this passage and its potential implications for Kierkegaard's modest-libertarianism in chapter four.

[55] My thanks to Robert Roberts for this suggestion.

not meet the third requirement. The willing addict is not able to constitute his will differently than he actually did.

In light of this difference between the willing and unwilling addict, I find it more accurate to say that three different types of higher-order volitions can be identified. In the first instance (the unwilling addict) the agent has a higher-order volition but lacks freedom because he is unable to make the desire that he wants his will. The unwilling addict has an ineffectual higher-order volition. In the second instance (the willing addict) the agent lacks true freedom of the will because although he is able to make the desire that he wants his will, he is not able to make a *different* desire his will. The willing addict has an unalterable effectual higher-order volition. In the third instance a person has both higher-order volitions and freedom of will because he is both able to make the desire he wants his will and able to constitute his will differently than he in fact does.[56] The truly free agent has an alterable effectual higher-order volition.

Importantly, even in his early work Frankfurt is adamant that only instances one and three demonstrate moral responsibility. Whether or not some sort of alternative motivational possibilities are necessary for an agent to possess freedom of the will, Frankfurt's case of Jones and Black shows that all that is necessary for moral responsibility on Frankfurt's account is that the agent does what she wants for reasons of her own and that the agent does not act because of some freedom-undermining manipulation or coercion. Apparently, the early Frankfurt thinks that there are conditions

[56] To complicate matters further, Frankfurt states that the satisfaction of second-order desires is freedom of the will, whereas the absence of freedom equals the frustration of these desires. In this case the willing addict is free (Frankfurt of course says that he both is and is not free). He wills with a kind of momentary freedom of the will when he wills to take the drug (in other words that particular willing is free), but he does not actually possess freedom of the will because he is not able to will otherwise; he does not have access to alternative motivational possibilities.

for the most desirable freedom of the will that is possible for a person to have but that these conditions are not necessary in order to ground ascriptions of moral responsibility.[57] Frankfurt's later work does not clarify this situation much. He focuses more on clarifying his notions of identification and wholeheartedness, to which I now turn.

Identification

Frankfurt's notion of higher-order desires as the basis of free will is not uncontroversial. Some scholars wonder what keeps an agent from forming higher and higher-order desires *ad infinitum*, presenting a problem of infinite regress.[58] Others wonder whether second-order desires are authoritative and freedom-granting in a way that first-order desires are not.[59] Yet others wonder whether second-order desires must be present in *akratic* behavior, as Frankfurt seems to indicate in some places.[60] Some scholars even wonder whether the notion of a second or higher-order desire is even intelligible.[61] Frankfurt responds to these criticisms by developing his notion of

---

[57] This distinction between conditions of full metaphysical freedom and freedom-relevant grounding conditions of moral responsibility has led some scholars to argue that humans may potentially possess free will attributes that are not necessary for responsibility ascriptions. I do not think that this is what Frankfurt is doing; rather I think that he is simply confused. He thinks that determinism is probably right, and he thinks that the more simplistic conditional analyses of freedom fall short, so he comes up with a compatibilist account of freedom that has a much richer psychology than previous accounts. The problem, however, is that it is difficult to see how this account can work without some sort of alternative possibilities (strong control) condition.

[58] Stump, "Persons: Identification and Freedom", Philosophical Topics 24 (1996): 184ff.

[59] Most notably Gary Watson.

[60] Mele, *Autonomous Agents*, especially pp. 65-80. Although Frankfurt never tackles the problem of akratic behavior head-on, in a number of places he does deal with ambivalence, by which he means something like a divided will. Akratic behavior is related in important ways to ambivalence, and Frankfurt argues that ambivalence "must arise out of a person's higher-order, reflective attitudes." Frankfurt, *Necessity, Volition, and Love*, (Cambridge, U.K.: Cambridge University Press, 1999), 99.

[61] See for example Loughrey, Dennis, "Second-order Desire Accounts of Autonomy," *International Journal of Philosophical Studies* 6 no. 2 (1998): 211-229.

identification. Roughly, identification is the theory of how an agent chooses one desire while rejecting others. By identifying with a desire, the agent makes it her own and secures the freedom of will absent in the instance of conflicting desires. I argue that although Frankfurt's notion of identification makes some headway against these criticisms, his account needs augmentation in important ways to respond to these objections adequately.

Recall that Frankfurt's basic claim is that for a person to will with a will that is free she must have second-order desires that correspond with her effectual first-order desires. But what keeps that person from also having a desire about her second-order desire, and a desire about her third-order desires, etc. ad infinitum? Furthermore, what makes that higher-order desire authoritative and freedom-granting in a way that the first-order desire is not? As Gary Watson notes:

> One job that Frankfurt wishes to do with the distinction between lower and higher orders of desire is to give an account of the sense in which some wants may be said to be more truly the agent's own than others…(and) the sense in which the agent 'identifies' with one desire rather and another and the sense in which an agent may be 'unfree' with respect to his own 'will.'[62]

In other words, Frankfurt introduces the concept of second-order desires to provide a framework for understanding the difference between a person who has the will she wants to have and a person who does not have the will she wants to have. But the mere framework of first and second-order desires alone does not articulate this distinction, since a person could have endlessly higher orders of desires and therefore fail to have the will that she wishes to have.

Frankfurt's first article did attempt to address these dual concerns of infinite regress and authority:

[62] Watson, Gary, "Free Agency". *The Journal of Philosophy.* 72 (8) (1975): 217.

> When a person identifies himself *decisively* with one of his first-order desires, this commitment "resounds" throughout the potentially endless array of higher orders…The fact that his second-order volition to be moved by this desire is a decisive one means that there is not room for questions concerning the pertinence of volitions of higher orders…the decisiveness of the commitment he has made means that he has decided that no further questions about his second-order volition, at any higher order, remain to be asked.[63]

However many philosophers were dissatisfied with what Frankfurt later admits were "terribly obscure" terms of "'identification,' 'decisive commitment,' and 'resounding'."[64] Reflecting on this passage from "Freedom of the Will," Watson famously replied that "[s]ince second-order volitions are themselves simply desires, to add them to the context of conflict is just to increase the number of contenders; it is not to give a special place to any of those in contention."[65] This objection has come to be known as the "problem of authority." An agent's desire to desire *X* does not *itself* grant the agent freedom or autonomy that the agent lacked without the higher-order desire. Watson claims that this process of identification is an arbitrary way of not permitting the "interminable ascent to higher-orders" and thus does nothing to solve the problem of authority.[66]

Frankfurt attempts to clarify his notion of identification and the authoritative role of higher-order desires in a series of articles. Several main points emerge about the nature of identification. First, identification involves the reflective evaluation and endorsement of a desire, though this evaluation is not a matter of weighing the worth of a particular desire. Second, identification is a method by which the agent cuts off a

---

[63] Frankfurt, "Freedom of the Will," 21-22 (book). Quoted in Frankfurt, "Identification and Wholeheartedness," 167.

[64] Frankfurt, "Identification and Wholeheartedness," 167.

[65] Watson, Gary, "Free Agency," 218. Quoted in Frankfurt, "Identification and Wholeheartedness," 166.

[66] Watson, Gary, "Free Agency," 218-19.

potentially endless array of higher-order desires by making a resolute decision. Third, through the process of identification, the agent internalizes the desires that she identifies with and externalizes or alienates the desires that she repudiates or chooses not to endorse. Finally, through the process of identification the agent "constitutes" herself, by which Frankfurt means that the agent decides which desires will have a place in her volitional structure, but more fundamentally the agent takes responsibility for her character and her actions.

The first point that Frankfurt makes about identification is simply an elaboration on his description in "Freedom of the Will and the Concept of a Person." He states that persons are unique in the ability to form second-order desires which require "reflective self-evaluation."[67] He later unpacks what he means by reflective self-evaluation: "To be a person entails evaluative attitudes (not necessarily based on moral considerations) towards oneself…Instead of responding unreflectively to whatever he happens to feel most strongly, he undertakes to guide his conduct in accordance with what he really cares about." Identification is a matter of reflecting on one's desires and deciding what desires one really cares about and choosing to endorse or identify with those desires. Despite the fact that Frankfurt notes that this reflective evaluation need not *necessarily* be based on moral considerations, many commentators have understood Frankfurt's notion of identification as involving estimations of value, particularly moral value. Frankfurt recognizes that his own terminology has led to this misunderstanding.

> The higher-order attitudes that are formed in processes leading to identification involve 'evaluations' only in the sense that is strictly value-neutral. In speaking of these matters, I have regrettably made use of terms—such as "endorse"—that naturally suggest a positive evaluation. However,

[67] Frankfurt, "Freedom of the will and the Concept of a Person," 12 (book).

> what I have actually intended to convey by referring to "endorsement" is not that the agent *approves* of what he is said to endorse, or that he considers it to *merit* his support, but nothing more than that the agent *accepts* it as his own.[68]

Identification is not a matter of ranking desires on some value-scale, or determining which desires have more merit. Frankfurt stresses in his most recent lectures that identification is simply a matter of figuring out which desire the agent wishes to accept as her own. This process contributes to the person's free will because that desire is something that the agent cares about, not because the desire has worth independent of the agent's cares and commitments.[69]

Furthermore, in Frankfurt's response to Bratman, he gives a nice summary of the process of identification. He states that four elements are distinguishable within the process of identification: First, the agent provisionally suspends or brackets her relationship to desire D_1, then the agent reflects on that desire. Again, he stresses that this reflection need not be based in any estimation of value whatsoever, but the agent simply is deciding whether D_1 is a desire that she wishes to accept as her own. Then the agent forms a higher-order desire or attitude toward desire D_1. Finally, in forming this higher-order desire D_2 the agent either identifies himself with desire D_1 or alienates himself from it.[70] These elements are the minimum requirements for identification. Thus, the mere approval or identification with a desire does not mean that the agent endorses the desire as an effective motive or as a justifying reason for action. It means that the agent takes the desire seriously, and identifies with it. "For someone to identify

---

[68] Frankfurt, "Response to Michael E. Bratman," in *Contours of Agency*, 87.

[69] Frankfurt, *Taking Ourselves Seriously and Getting it Right*," 22.

[70] Frankfurt, "Response to Michael E. Bratman," 87.

with a desire means merely that—for whatever reason, or for no reason whatever—he joins himself to the desire and accepts it as his own."[71]

The second salient point is that through the process of identification the agent cuts off the potentially endless array of higher and higher orders of desires. In "Identification and Wholeheartedness," Frankfurt compares identification to solving a math equation, where the recalculation of the problem could extend indefinitely; but, at some point the mathematician decides to "cut off" this potentially endless recalculation. This decision could be a result of boredom, or resignation that the problem can never be solved confidently. But it could also result from the mathematician's confidence that she has solved the equation and no amount of recalculation is going to change her answer.[72] Frankfurt says that "[i]n this respect, the future is transparent to him (the mathematician), and his decision that a certain answer is correct resounds endlessly in just this sense: It enables him to anticipate the outcomes of an infinite number of possible further calculations."[73] He goes on to use this mathematics analogy to clarify his notion of resonance:

> The fact that commitment resounds *is* simply the fact that the commitment is *decisive*. For a commitment is decisive if and only if it is made without reservation, and making a commitment without reservation means that the person who makes it does so in the belief that no further accurate inquiry would require him to change his mind. It is therefore pointless to pursue the inquiry any further. This is, precisely, the resonance effect.[74]

---

[71] Ibid., 89.

[72] Frankfurt, "Identification and Wholeheartedness," p. 168.

[73] Ibid.

[74] Ibid.,168-69.

Here Frankfurt homes in on a clearer understanding of identification. When an agent identifies with a desire, she takes a stance in relation to that desire. She decides that that desire is what she wants to be her will and not any other. Furthermore, she decides that since this desire is what she wants to be her will, she will not reflect endlessly on her desire for that desire. Frankfurt states that to decide literally means "to cut off" and that "…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."[75] The picture emerges of an agent with various desires who decisively chooses to identify with one of those desires while rejecting the others, and through this decision terminates the potentially endless "sequence of desires or preferences of increasingly higher orders."

The last two points about identification are closely related. Frankfurt states that when the agent decisively identifies with a desire, she *constitutes* herself.[76] This provocative statement needs unpacking. Frankfurt has a notion of self-identity or personhood that is closely linked to his understanding of identification. He points out in a number of places that agents have all kinds of desires, but that these desires are not necessarily constitutive of the agent. In other words, not all of an agent's given desires belong to that agent or contribute to that agent's own self-understanding. Consider the example of a father who has the sudden desire to kill his son. The father likely will reject this desire immediately and find it repulsive, refusing to identify with it. Frankfurt thinks that there is a sense in which the father alienates that desire from his psyche, even though

[75] Ibid., 170.

[76] Ibid.

the desire did originate within the agent. Frankfurt says that psychic elements such as these are "exogeneous—that is, if the person is not identified with them…they are…external to his will."[77] In a later essay he elaborates on this notion of externality by noting that human agents often find themselves with psychic elements that the person deems "dangerously antithetical to his intentions and his conception of himself."[78] Agents can disassociate themselves from these rogue desires and externalize them by pushing them away through repression or other psychological ways. Most importantly, by externalizing unwanted psychic elements "we deny them any entitlement to supply us with motives or with reasons."[79] Identification is the way that an agent decides which desires will be motivating and which will not.

Furthermore, this process of externalizing unwanted desires allows the agent to resolve fundamental tensions within her psyche, tensions that Frankfurt refers to as ambivalence. Frankfurt argues that ambivalence is constituted by conflicting volitional elements and must meet two conditions. First, the volitional elements must be "inherently and hence unavoidably opposed…" Second, the volitional elements must be "wholly internal to a person's will rather than alien to him…" In other words, the agent has not rejected or externalized either element and thus is not passive to either element. If these two conditions are met, then true ambivalence cannot occur between "first-order psychic elements alone…but must arise out of a person's higher-order, reflective

---

[77] Frankfurt, "The Faintest Passion," in *Necessity, Volition, and Love,* 99.

[78] Frankfurt, Harry G., and Debra Satz, *Taking Ourselves Seriously & Getting it Right*, (Stanford, Calif: Stanford University Press, 2006), 10.

[79] Ibid.

attitudes."[80] An ambivalent person is inclined in two different directions at the same time; furthermore her attitude *towards* these inclinations is unsettled. Frankfurt does not think that conflicting first-order desires alone can cause ambivalence; if the agent is ambivalent, she must take a reflective stance towards the desires in question.

Identification makes certain desires internal to the agent's volitional structure and makes others external. Through the process of identification, the agent acquires motives and reasons for action. Thus, the agent constitutes herself by incorporating certain desires into her will by choosing whether or not to endorse them. Self-constitution is a matter of reflecting on a desire and choosing either to endorse that desire and accept it as the agent's own or reject it and alienate it, making it external to the self in a fundamental sense. This indicates that Frankfurt's notion of identification is not only a theory about the necessary conditions for free and responsible agency, but also a theory about how agents can participate freely in character formation. Frankfurt does not believe that agents can create themselves *ex nihilo*. He does believe agents can decide, through the process of identification, which desires will be a part of their volitional structure and what kind of person they will be. "The willing acceptance of attitudes, thoughts, and feelings transforms their status. They are no longer merely items that happen to appear in a certain psychic history. We have taken responsibility for them as authentic expressions of ourselves."[81] Frankfurt believes that all agents find themselves with "psychic raw elements that nature and circumstances have provided us…"[82] However, this idea that

---

[80] Frankfurt, "The Faintest Passion," 98.

[81] Ibid., 8.

[82] Frankfurt, "Taking Ourselves Seriously," 7.

agents actively participate in character formation is at odds with Frankfurt's disagreement with the Aristotelian idea that "we are responsible for what we are to the extent that we have caused ourselves—by our voluntary behavior—to become that way."[83] Instead he argues that becoming responsible is not a matter of developing character traits out of which desires arise but taking responsibility for certain desires through the process of identification. I return to this important ambiguity in chapter four. The salient point here is that free and responsible agency is accomplished through internalizing certain desires through the process of identification while repudiating and externalizing other desires.

Frankfurt's continued refinement of his notion of identification is helpful and makes some headway towards answering his critics' concerns. However, the fundamental problems still linger. Claiming that the potentially endless array of higher-order desires can be terminated through a resolute decision does not mean that there cannot still be a potentially infinite array of higher-order desires. The critics want to know if something in the structure of a second-order desire will prevent the infinite regress. Frankfurt's answer indicates nothing to that end. Furthermore, the problem of authority is not only unsatisfactorily addressed by Frankfurt's revisions, it is amplified. The issue is how second-order desires solve certain problems of free will and moral responsibility by demonstrating the difference between an agent that acts freely and responsibly from an agent that does not. In other words, how does a hierarchical view address the autonomy, control, and knowledge components that an adequate theory of free will and moral responsibility must address? An agent's possession of a higher order-desire or volition does not indicate that she is the source of her actions in the way a manipulated person is not the source of her actions. Furthermore, the mere existence of

[83] Ibid., 6.

higher-order desires and volitions does not *eo ipso* demonstrate that an agent has the control over her desires necessary for grounding responsibility ascriptions. Frankfurt admits the claim about manipulation in a response to Fischer's revision of his view:

> A manipulator may succeed, through his interventions, in providing a person not merely with particular feelings and thoughts but with a new character. That person is then *morally responsible* for the choices and the conduct to which having this character leads…We are the sorts of persons we are; and it is what we are, rather than the history of our development, that counts.[84]

The far-reaching implications of this passage require unpacking. But for the purposes here, I want to note that Frankfurt admits that someone whose character is the direct result of manipulation still can meet the requirements for morally responsible action, including higher-order desires. This goes against Frankfurt's original intent of formulating a hierarchical account. The hierarchical account could explain why some agents are responsible and/or free while others are not. Yet, Frankfurt states that even higher-order desires can be the result of manipulation and therefore admits that his hierarchical account cannot do what he had previously claimed it could do. Namely, his theory cannot distinguish between agents whose actions are the result of manipulation or addiction and agents who act responsibly and/or freely.

Recall Frankfurt's distinction between the willing and unwilling addicts. Although Frankfurt claims that neither addict is free in the most robust sense, he claims that the willing addict is the only addict *responsible* for his actions. Frankfurt explains why the willing addict is responsible in an important passage:

> Suppose that a person has done what he wanted to do, that he did it because he wanted to do it, and that the will by which he was moved when he did it was his will because it was what he wanted. Then he did it freely and of his own free will. Even supposing that he could have done otherwise, he would not have done otherwise; and even supposing that he could have had a different

[84] Frankfurt, "Reply to John Martin Fischer" 28 (my emphasis).

> will, he would not have wanted his will to differ from what it was. Moreover, since the will that moved him when he acted was his will because he wanted it to be, he cannot claim that his will was forced upon him or that he was a *passive bystander to its constitution*. Under these conditions, it is quite irrelevant to the evaluation of his moral responsibility to inquire whether the alternatives that he opted against were actually available to him.[85]

The willing addict is morally responsible for his actions because he did what he wanted to do, and as long as he wants his will to be constituted in the way that it is, then he is responsible despite the fact that due to his addiction he cannot want to have a will to refrain from the drug. This claim about the moral responsibility of the willing addict appears consistent with the claim that a person whose character is the result of manipulation can be morally responsible for her actions. The same requirements could be true of the willing addict. The willing addict wanted to have the will that he did have; he willed the way he did because he wanted to; and finally, even if he could have constituted his will differently, he would not have wanted to. Frankfurt concludes that for these reasons the willing addict *cannot claim* that "his will was forced upon him or that he was a *passive bystander to its constitution*."[86] He was an active contributor to the constitution of his will, thus he is morally responsible for his will. Even if an agent (or perhaps a natural cause) had manipulated him to have the will that he does, that is not relevant to ascriptions of moral responsibility.

The case of the willing addict initially harmonizes with Frankfurt's requirements for moral responsibility, yet what is puzzling is why the unwilling addict cannot be morally responsible on Frankfurt's own account. Consider this passage about the unwilling addict:

---

[85] Frankfurt, "Freedom of the Will," 24 (book). My emphasis.

[86] Ibid, 17.

> The unwilling addict identifies himself…through the formation of a second-order volition, with one rather than the other of these conflicting first-order desires. He makes one of them more truly his own and, in so doing, he withdraws himself from the other. It is in virtue of this identification and withdrawal, accomplished through the formation of a second-order volition, that the unwilling addict may meaningfully make the analytically puzzling statements that *the force moving him to act is a force other than his own*, and that it is not of this own free will but rather against his will that this force moves him to act.[87]

The unwilling addict is person for whom freedom is a problem because he wants certain desires to be motivating and to constitute his will, but he lacks the freedom to makes those desires his will. Through the process of identification, the unwilling addict alienates the unwanted desires and is not responsible for them. Those desires move him to act like "a force other than his own." In other words, because he has withdrawn from his own motivating desires he is, in Frankfurt's words, a "passive bystander" to the constitution of his own will and therefore is not morally responsible for his will. But if the unwilling addict is not morally responsible due to the fact that he is a passive bystander to his own will, then how can Frankfurt say that a person (such as the willing addict) is responsible even if a manipulator causes him to have the will that he has? To press the point, consider what have come to be known as Walden Two examples.[88] In these examples, the citizens are happy only because they have been fully conditioned to want what they want. They have had no say in the development of their character. Frankfurt's statement about a manipulator causing an agent to have the character that she does could be taken to posit that this not only means that the agent's first-order desires, but her higher-order volitions as well, are determined by the manipulator. Would

[87] Frankfurt, "Freedom of the Will," 18 (book). My emphasis.

[88] This term is taken from B.F. Skinner's *Waldon Two.*

Frankfurt still declare that she is responsible for her actions? He indicates that the willing addict is responsible despite the fact that she is not fully free because she wants to want what she wants, even though the initial wants (the desire to take the drug) is something is she cannot change and has no say in. The willing addict's first-order desires are out of her control. Since they could be forged by a manipulator, she is able to make the first-order desire that she wants her will. Apparently this higher-order volition is not the result of addiction or compulsion like the first-order desire, and therefore is in her control. However, if both her first and higher-order desires and volitions were out of her control in the same way that her first-order desires are, Frankfurt would have to say that she is responsible since he argues that an agent is morally responsible for her actions even if a manipulator provides her with a new character. With this claim, Frankfurt's only remaining basis must rest on the claim that the willing addict is morally responsible while the unwilling addict is not since the willing addict has the right mesh of desires.[89]

The notion of the right mesh of desires has important implications for the authority objection to Frankfurt's hierarchical view. If all that matters for moral responsibility is the right mesh of desires, then an agent is responsible for her actions even if her higher-order desires and volitions are not her own but are the result of manipulation, addiction, etc. Frankfurt's initial counterexample against PAP argued that an agent can be considered morally responsible for his actions despite his lack alternative possibilities, but only because Jones acted on his own without intervention. If the Black would have intervened to *cause* Jones to perform a certain action, then Jones would not be morally responsible for that action. Based on his counterexample against PAP, one

[89] Eleonore Stump tends to refer to Frankfurt's hierarchical theory of the will as a "mesh theory" since an agent is free if she has the right "mesh" of higher and lower-order desires.

can plausibly draw the conclusion that Frankfurt believes that agents are responsible for their actions only if they are the source of those actions, and that manipulation (or addiction in the case of the unwilling addict) makes something other than the agent the source of his actions and in this way undermines responsibility ascriptions. Frankfurt's hierarchical theory alone does not tell us why someone who has the right mesh of desires is responsible in a way that a manipulated agent is not.

All is not lost for hierarchical views of moral responsibility; however, Frankfurt's view requires supplementation. He must demonstrate how higher-order volitional states *belong* to the agent in a way that first-order desires do not. In chapter four, I argue that Kierkegaard shows how an agent forms higher-order volitional states deeply imbedded in that agent's character in a way that makes them the agent's own through a striving will. Furthermore, I argue that an agent must be able to make open decisions[90] to develop her character and be the source of her own actions. At this point I discuss Frankfurt understanding of care or what he sometimes refers to as "volitional necessity."

Volitional Necessity

The third way that Frankfurt helped to shape the discussion of free will and moral responsibility is with his notion of care or volitional necessity. Volitional necessity is a state of the will where the agent is wholeheartedly committed to certain ends and cannot bring herself to will in ways opposed to those ends. However, the commitment that constrains her is not is not an external force. Instead, the agent's own will does the constraining. The agent cares about the object of her will so much that willing otherwise than she does is "unthinkable." She cannot bring herself to will something else through a

[90] I define "open" decisions as decisions where more than one action is compossible with the events of the past and the laws of nature.

singular mental act because of her volitional state. However, the agent may intentionally end this state of volitional necessity as the result of some longer process and not by one mental act.[91] Frankfurt calls this state *volitional* necessity because he thinks that the agent is not constrained primarily by the dictates of reason or of morality, but by what the agent cares about, which is primarily a function of the agent's will.[92] He calls this volitional *necessity* because he believes that there is a real sense in which the agent is constrained by her commitments. One objective of this section is to unpack these two connected ideas.

Caring and volitional necessity/wholeheartedness[93] are important parts of the discussion of free will and moral responsibility. Having a will that is free closely connects with having an integrated self with which the agent is satisfied. The will cannot be truly free if elements or characteristics within the self are present that the self does not wish to have. Sourcehood, or *autonomy*, provides a grounding condition for moral responsibility. At first, sourcehood and autonomy may seem like different concepts, but the close relationship is seen by considering the opposite of autonomy, *heteronomy*. Literally meaning laws given by another, heteronomy in moral philosophy means an action performed by an agent that is caused by something external to the agent. Thus autonomous action is action that is caused by or comes from within the agent, making the

[91] As noted below, Frankfurt does not specify how this process of "undoing" a state of volitional necessity might unfold.

[92] Frankfurt refers to himself as a "voluntarist" because he thinks that care is primarily a function of the will, not of the intellect or of the emotions. *Taking Ourselves Seriously and Getting It* Right, 105, n. 1.

[93] Frankfurt tends to use the terms volitional necessity and wholeheartedness as synonyms. I however prefer to use wholeheartedness to refer to a state of the agent's volitional structure as a whole whereas volitional necessity can refer to one particular aspect of the agent's will. In other words, an agent can have certain desires that are subject to volitional necessity while still being heteronomous in other ways, but the agent who is wholehearted does not have any desires, inclinations, or characteristics that she does not wish to have.

agent the source of the action. An agent that is truly free and morally responsible is an agent who is the source of her own actions and does not have desires, inclinations, characteristics, etc. that she does not want and that, in turn, cause her to have volitions that she does not wish to have. Following Frankfurt's lead, I call this state of absolute autonomy *wholeheartedness*. Like Frankfurt's concept of the hierarchical will, volitional necessity is not new, despite the new terminology. Something similar to volitional necessity can be traced back at least as far as Augustine in the Christian tradition. One could argue that Aristotle's notion of states of character has similarities as well. What is unique about Frankfurt's concept is the language he employs and the way that he ties the concept of volitional necessity to his hierarchical conception of the will. A number of problems with Frankfurt's concept of volitional necessity appear when taken on its own. I find that when coupled with a libertarian understanding of freedom and supplemented by Kierkegaard's concept of becoming a self, Frankfurt's notions of a hierarchical will and volitional necessity yield a compelling picture of human freedom that successfully grounds not only moral responsibility for an agent's particular actions but moral responsibility for the cultivation of an agent's character.

Frankfurt introduced the main contours of volitional necessity in his 1982 article, aptly titled "The Importance of What We Care About."[94] This notion of care builds on his earlier distinction between first and second-order desires and the claim that what is most distinctive about personhood is the capacity to form higher-order volitions. When the concepts of care and volitional necessity are added to these distinctions, Frankfurt's notion of personhood expands into several subdivisions. The main difference between a

[94] Hereafter all references are to the article as reprinted in *The Importance of What We Care About*, 80-94.

higher-order desire and a higher-order volition is that the when an agent forms a volition she takes an interest in the constitution of her will; she actually wants to will certain things as opposed to others. A wanton, on the other hand, "cannot or does not care which of his conflicting first-order desires wins out…When a *person* acts, the desire by which he is moved is either the will he wants or the will he wants to be without. When a *wanton* acts, it is neither."[95]

A person differs from a wanton in that a *person* forms higher-order volitions and a wanton does not. A *person* takes an interest in and decides which desires will constitute her will. However, not every *person* necessarily displays the distinctive qualities of caring. An agent forms a higher-order volition merely by preferring one desire over another. But, as Frankfurt notes in his essay "On Caring," the fact that an agent has desires or preferences—even fulfilled desires—does not mean that that person is happy or satisfied. After all, many people have desires that they do not care about.[96] This relates to the distinction between autonomous and heteronomous volitions. A person can attempt to identify with all of her desires. She may find that this does not make her happy because she does not care about all of those desires. In fact, she wishes that she did not have some of those desires. When a *person* chooses to identify with certain desires, this choice could be made on a whim, or it could be made due to some deeper commitment or belief the agent holds. In the former instance, the agent is a *person* since she has higher-order volitions, even though she does not exhibit care in her decision. In

[95] Frankfurt, "Freedom of the Will," 18-19 (book).

[96] Frankfurt, "On Caring," *Necessity, Volition, and Love*, 157.

the latter instance, the person's identifications stem from a larger framework[97] of attachments, commitments, beliefs, etc; this *person* cares.

Holding higher-order volitions is a necessary but not sufficient condition of caring, since one can desire things *without* caring (in the instance of the *person* with higher-order volitions but lacking in care). But without higher-order volitions, one finds it impossible to care. Imagine a woman, Janet, who claims to care for the homeless. Surely Janet cannot care about the homeless unless she has some first-order mental states about the homeless: anger at seeing a homeless person mistreated, desire to volunteer at a food kitchen, etc. One could easily imagine a situation where Janet has the requisite first-order mental states, yet does not really care about the homeless. First, Janet may feel anger at seeing a homeless person mistreated, but she may not actually want to feel this way because she knows that a commitment to that mental state entails other commitments such as acting on the behalf of the homeless. Second, she may feel the desire to help the homeless person and not care about the desire either way. She has the passing desire to help a homeless person but can quickly dismiss this feeling without deciding whether or not she wants to have the desire to help the homeless. In other words, Janet could have certain emotions and desires about the homeless and still be a *wanton* with regard to those desires, having no higher-order mental states about her first-order mental states. The essential characteristic of a wanton is that she does not care about her will. Desires move her, or happen to her (as they do to Janet in this first scenario), but she does not care what those desires are. In other words, a wanton is not concerned with the

[97] I am intentionally referring to Charles Taylor and his notion of 'inescapable frameworks" because I think that there is a certain correlation between Taylor's understanding of an inescapable framework and Frankfurt's notion of volitional necessity.

desirability of her desires.[98] To be a person who cares, Janet must not only have first-order desires and emotions, but she also must want to have these desires and want them to motivate her, that is, want them to be her will.[99] Alternatively, Janet can have *weak* higher-order desires—desires that are not rooted in some deeper guiding framework and that are merely preferences—and still not *care* in Frankfurt's sense.

In Frankfurt's sense, for Janet to care about the homeless, she must not only have certain first and second-order desires and volitions, but those higher-order volitions must be *strong*. Strong higher-order desires cannot be mere preferences; they must stem from a guiding framework, and give the self a kind of endurance through time.[100] As Frankfurt puts it, "[w]hen we do care about something, we go beyond wanting it…the caring entails, in other words, a commitment to the desire."[101] For a desire to represent or be the result of a care, the desire must not only persist if frustrated, it must also "endure through an exercise of his (the agent's) own volitional activity rather than by its own inherent momentum."[102] Thus, if Janet really wants to have those desires about the homeless, she must take action to ensure that they do not whimsically come and go. Care entails not merely the approval or endorsement of a desire (lip-service), but the active commitment to seeing that the desire is not abandoned or neglected.

---

[98] Frankfurt, "Freedom of the Will and the Concept of a Person," 11 (journal article).

[99] She can also wish *not* to have those desires. The point is that she must have higher-order attitudes one way or another about her first-order desires.

[100] Frankfurt is not entirely clear what exactly is enduring here, as he rejects an essentialist understanding of human beings. He seems to view the self as a cluster of commitments, desires, and other mental states, and insofar as those mental states remain constant the self "endures through time."

[101] Frankfurt, *Taking Ourselves Seriously*, 18.

[102] Frankfurt, "On Caring," 160.

Active commitment or intentionality[103] is essential to caring, because we could imagine a case where an individual cares about nothing (in the sense of care that Frankfurt means) and yet still wants to have certain desires (perhaps even motivating desires), and does not want to have or be motivated by other desires. Kierkegaard's case of the pseudonymous A provides an interesting example.[104] He has moved beyond vulgar hedonism—the commitment to physical pleasure—and decides that he is better off not satisfying certain desires, namely desires that would require long-term commitments and would rule out satisfying whatever other desires happen to come along. A's only commitment is his refusal to make any long-term commitments. A has *weak* second-order volitions. A's higher-order desires are *volitions* because he does want certain desires to constitute his will (desires that will minimize pain and boredom, but will not develop a character that will constrain his desires), but his higher-order volitions are *weak* because they do not stem a larger framework of commitments and change from moment to moment based on what desires will yield the most pleasure. Because of A's lack of intentionality about his desires, Judge William argues that A may well find these weak second-order desires and volitions solidify and that he develops a stable character, albeit unintentionally.[105] For instance, A may wake up one day and find that he *cannot* make any long-term commitments even if he wants to. He might become the kind of person who cannot commit, although he has become this kind of person unintentionally. In this way, a person's weak higher-order desires and volitions might give volitional consistency

[103] Here I am using intentionality according to its everyday usage, not in any technical philosophical sense.

[104] See Kierkegaard's *Either/Or* volumes I and II. This example is fleshed out in chapter five.

[105] See for instance Kierkegaard's *Either/Or II*, 164.

or stability to his life. Judge William and Frankfurt agree that this kind of stability is "…merely fortuitous and inadvertent. It would not be the result of any deliberate or guiding intent on our part."[106] Thus, a person's capacity for higher-order desires and volitions might remain intact despite the fact that she cares about nothing, rendering it necessary that a person have more than the mere existence of those states. A person who cares is different from someone like A in that the person who cares plays an active, willing role in the continued existence of those higher-order dispositions and attitudes by identifying with her desires.

Frankfurt states that caring "coincides in part with the notion of something with reference to which the person guides himself in what he does with his life and in his conduct."[107] This guiding framework of the *person* who cares means that she will guide herself "along a distinctive course…"[108] On the other hand, a person who shows weak higher-order volitions (such as A) can drift through life since she takes minimal interest in the constitution of her will. Furthermore, the *person* who has weak higher-order volitions can have a will that is free in particular instances insofar as she has the right mesh of lower and higher-order desires at those moments. In a recent essay, Frankfurt puts the point this way:

> Willing freely means that the self is *at that time* harmoniously integrated. There is, within it, a *synchronic* coherence. Caring about something implies a *diachronic* coherence, which integrates the self across time…By our caring…we engage ourselves in guiding the course of our desires. If we cared

[106] Frankfurt, "On Caring," 162.

[107] Frankfurt, "The Importance of What We Care About," 82.

[108] Ibid, 83.

> about nothing, we would play no active role in designing the successful configurations of our will.[109]

Caring gives direction and shape to a person's life by integrating the self and building a history that is essential to morally responsible agency.

Caring can give direction and shape to a person's life only if it not only serves as a kind of guide to the person, but also endures over time. Frankfurt distinguishes manifestations of care from mere desires and beliefs because the latter do not have any necessary or inherent persistence. In other words, "nothing in the nature of wanting or of believing *requires* that a desire or belief must endure."[110] One can easily imagine a teenager who sees on television a puppy suffering from maltreatment and immediately decides to become an activist for the humane treatment of animals, only to change her mind the next week (or hour). She acts as if she cares about animals enough to orient her entire life around this goal, but soon changes her mind. She may form a weak higher volition freely, but that weak volition is not an instance of care, but more akin to being moved by impulse, or a kind of volitional spasm. Something more is needed to move from mere impulse to where our projects and goals have a consistency that orients the agent's life. Part of what being a fully mature person means for Frankfurt includes this ability to guide oneself along a steady path, to do something purposely with oneself.[111]

---

[109] Frankfurt, *Taking Ourselves Seriously and Getting it Right*, 19. Quoted by Korsgaard in the same volume, 60.

[110] Ibid. 84. My emphasis.

[111] Ibid. 83.

In this way, a person who cares about something (as opposed to merely liking or desiring it) builds a history and a narrative continuity that otherwise would be lacking.[112]

This narrative continuity also requires that a person take what I call the "long-view" of her life. Or as Frankfurt puts it, "[t]he outlook of a person who cares about something is inherently *prospective*."[113] A person who cares knows that she has a future, and thinks about her life in those terms. When she takes stock of her desires and goals, she thinks about the long-term viability of those desires and goals, and considers whether or not they should be adjusted. I do not think that Frankfurt means that a care can never be about something that the agent knows cannot be long lasting, but that the agent maturely recognizes the duration of that for which she cares and takes stock of the consequence of orienting her life around something temporary. A person who cares knows that the satisfaction of temporary desires is not what brings about happiness. The prospective person takes a broader view of life and understands that happiness requires a level of continuity that mere desire-fulfillment does not.

The combination of these of these last two characteristics—that cares endure over time and provide narrative continuity, and that a person who cares is prospective—yields another key component of Frankfurt's notion of caring. Cares are the source of a person's personal identity or what Kierkegaard's pseudonym Judge William calls personality. One reason for this is that apart from building a history there is no subject.

> Desires and beliefs can occur in a life which consists merely of a succession of separate moments, none of which the subject recognizes—either when it

[112] As I note below, it is possible to achieve a level of narrative unity without actually caring about anything in the way that Frankfurt means, but it is not a narrative that truly *belongs* to the agent, thus other qualifications must be added.

[113] Ibid, 83 (italics mine).

> occurs or in anticipation or in memory—as an element integrated with others in his own continuing history. When this recognition is entirely absent, there is no continuing subject.[114]

Frankfurt claims that caring takes a number of separate moments and disconnected desires and gives them coherence and meaning. This coherence allows a person to know who she is—where she has come from and where she is going. It also allows her to recognize that these cares serve as guide, belong to the agent, and give the agent a sense of personal identity.

One of the most interesting characteristics of Frankfurt's notion of care is that care concerns a cluster of questions that are "not properly within the scope either of epistemology or of ethics…"[115] He clarifies what he means by saying that care is not properly within the scope of ethics, in the recent lecture "Getting It Right:"

> Suppose you are trying to figure out how to live. You want to know what goals to pursue and what limits to respect…your most fundamental problem is not to understand how to identify what is valuable…neither judgments of value in general nor moral judgments in particular can settle this for you…What we actually care about—what we are to regard as really important to us—cannot be based simply upon judgments concerning what has the most value.[116]

Frankfurt believes that caring, and not moral considerations of value judgments, is the source—or should be the source—of an agent's practical normativity. Frankfurt rejects "normative realism", the belief that "there are objective reasons for us to act in various ways, whether we know about them, or care about them, or not." Furthermore, he does "not believe that anything is inherently important...[because] the standards of volitional

---

[114] Frankfurt, "The Importance of What We Care About," 83.

[115] Ibid., 80.

[116] Frankfurt, *Taking Ourselves Seriously and Getting It Right*, 27-28.

rationality and of practical reason are grounded…only in ourselves."[117] Two related ideas are at work here. First, Frankfurt is affirming internalism about reasons. Second, he is not allowing any place for objectivity in caring.

Internalism states that "there are no reasons for acting that apply to everyone, regardless of their prior desires, commitments, and projects."[118] For example, I cannot give another person a reason not to rob banks unless she already cares about not robbing banks; not robbing banks must connect up somehow with her "subjective motivational set."[119] If she *does* have a reason not to rob the bank, then that reason must have its origin in his preexisting care(s).[120] Bernard Williams thinks that this reason can originate in one of two ways. Either it can be a preexisting, conscious care, or it can be reached through the agent's reasonable deliberation based on his already existing cares. Thus it may be possible to convince a bank robber to stop robbing banks by appealing to cares that she already has, assuming that she can make a reasonable inference from her actual cares to conclude that she also should care about not robbing banks.

Frankfurt is an internalist in terms of reasons and motivation.[121] He states that "[u]nless a person knows what he *already* cares about…he cannot determine what he has

---

[117] Ibid., 33.

[118] Blustein, Jeffrey. *Care and Commitment: Taking the Personal Point of View*, (New York: Oxford University Press, 1991), 21.

[119] Williams, Bernard. "Internal and External Reasons," in *Moral Luck: Philosophical Papers, 1973-1980*, (Cambridge [Cambridgeshire]: Cambridge University Press, 1981), 102.

[120] Blustein, 22.

[121] Fischer also argues that Frankfurt is a moral internalist, but Fischer defines moral internalism as evaluating moral responsibility solely in terms of the "internal features of an agent's configuration of mental states and dispositions" as opposed to the agent's history and connection to the external world. (Find source) The definition I am utilizing has more in common with the second feature of Fischer's moral externalism.

reason to care about."[122] Frankfurt thinks that "things are only important if they make a difference."[123] In other words, Frankfurt agrees with Williams that an agent will not be motivated to act in a certain way unless some motive in her subjective motivational set will be served by that action. If the agent's motives are not satisfied by the action in question, then that action will not be important to the agent. Frankfurt assumes that human agents are unable to form new motivations that are not based in already existing motivations. Therefore all Frankfurt can recommend is that a person who wants more direction in her life reflect on what she cares about. If she can figure this out (and it is quite possible that she cannot, due to the fact that often people do not know what they really care about) then she will have all that she needs to order her life and priorities.

Frankfurt's emphasis on the importance of internal reasons for caring is coupled with a rejection of any sort of objectivity. "There can be no rationally warranted criteria for establishing anything as inherently important…any answer to the normative question must be derived from considerations that are manifestly *subjective*."[124] He affirms this view saying, "The fact that something is important to us does not primarily consist in our estimate of its own value. The question of what we are to care about is not settled by arriving at judgments as to the inherent or comparative merits of various possible objects of devotion."[125] These statements represent the other half of what Davenport calls

---

[122] Frankfurt, *Taking Ourselves Seriously*, 23.

[123] Ibid.

[124] Ibid., 22,24 (my emphasis).

[125] Frankfurt, "The Faintest Passion," in *Necessity, Volition, and Love*, 106. Quotes affirming the same basic principle as the previous two abound in Frankfurt's writings. This is clearly something he does not want his reader to miss.

Frankfurt's subjective existentialism[126] that can be inferred from the internal reasons thesis. The agent cannot be motivated to act unless the reason in question connects up with her subjective motivational set. Additionally, she cannot even claim that anything is inherently important *independent* of that motivational set. Frankfurt rejects Nagel's claim that "we can hope to discover the truth by transcending the appearances and subjecting them to critical assessment."[127] Objects out there in the world cannot be ranked on some sort of universal value scale, or according to any objective criteria. "To care about something differs not only from wanting it and from preferring it but also from judging it to be valuable."[128] Susan Wolf argues for objectivity in caring, such as in the individual's affinity or fit between her needs and desires and the characteristics of the object of care.[129] Frankfurt's response is that "[a]n enthusiastically meaningful life need not be connected to anything that is objectively valuable, nor need it include any thought that the things to which it is devoted are good. Meaning in life is created by loving."[130] Thus figuring out what *should* motivate to us has nothing to do with the value of the objects of our love of care, but is a *factual* matter about what we actually do love; nothing ought to motivate everyone, because motivation is purely a subjective matter.

---

[126] Davenport, John, *Will as Commitment and Resolve: An Existential Account of Creativity, Love, Virtue, and Happiness*, (New York: Fordham University Press, 2007), 490.

[127] Nagel, Thomas, *The View from Nowhere*, (New York: Oxford University Press, 1986), 139.

[128] Frankfurt, "On Caring," in *Necessity, Volition, and Love*, 158.

[129] Wolf, "The True, the Good, and the Lovable: Frankfurt's Avoidance of Objectivity," in *Contours of Agency: Essays on Themes from Harry Frankfurt*. Edited by Buss, Sarah, and Lee Overton. Cambridge, Mass: MIT Press, 2002, p. 233. This is a very chastened objectivity, but objectivity nonetheless. Davenport elaborates on this point in *Will as Commitment and Resolve*, 508.

[130] *Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Lee Overton (Cambridge, Mass: MIT Press, 2002), 250.

The last main aspects of care are related. Frankfurt believes that we cannot bring ourselves to care about something (or to stop caring) through a single act of will. What we care about is not under out direct volitional control. He states that "…what we care about is not always up to us…[but] consists of desires and dispositions that are not under our immediate voluntary control. We are committed in ways that we cannot directly affect."[131] Care is constituted by certain desires, emotions, and other mental states, therefore not always within the agent's control. Instances can occur in which we cannot stop caring immediately about what we do because we cannot get rid of those desires, emotions, etc.[132] Emotions and desires are complex mental states. Although I argue that Kierkegaard believes that we maintain a level of control over emotions, I think that he would agree with Frankfurt that these mental states are not something over which we have *direct* or *immediate* control. I cannot wake and simply decide that I am going to feel joyful all day. And since these states are what constitute caring, it follows that caring is not something that can be willed directly either. Furthermore, cares are dispositional, and as such are entrenched deeply in a person as shown by the fact that cares endure over time.[133] These cares give the agent a personal identity and provide a key component of that person's character. Thus cares are related closely to character traits, and character traits are not something that a person chooses at will.

---

[131] *Taking Ourselves Seriously*, 24.

[132] This is one of four qualifications that Frankfurt makes on caring in this essay "On Caring," *Volition, Necessity, and Love*, 161-62.

[133] Davenport, *Will As Commitment and Resolve*, 465.

Moreover, our cares are not under our direct volitional control because "decisions," in the normal sense of the word, do not entail caring in Frankfurt's sense.[134] "A decision to care no more entails caring than a decision to give up smoking entails actually giving it up. In neither case does making the decision amount even to initializing the state of affairs decided upon unless that state of affairs actually endures."[135] To emphasize this point, Frankfurt refers to the "young man" in Sartre's famous example. Even if the young man makes the choice to stay home with his mother instead of go off to war, he may not follow through with his decision.

> …[H]e might be unable to carry out his intention. He might discover, when the chips are down, that he simply cannot bring himself to pursue the course of action upon which he decided…he might discover that he does not have and that he does not subsequently develop the feelings, attitudes, and interests constitutive of the sort of person which his decision has committed him to being.[136]

In other words, he might find out that he does not have the sort of character necessary for carrying out such an intention. Just as we simply do not get to decide what kind of person we are going to be, we do not get to decide about that which we are going to care.

Frankfurt subdivides his category of a *person* into the *person* who cares and the *person* who does not. A similar distinction can be made within his category of the *person* who cares. The *person* who cares may come to care about something so much that "it is impossible for him to forbear from a certain course of action."[137] At this point caring gives way to volitional necessity or wholeheartedness.

---

[134] Ibid.

[135] Frankfurt, "The Importance of What We Care About," 84. Even then, the complicated question remains whether the agent's decision played a causal role in the ensuing state of affairs.

[136] Ibid., 84-85.

[137] Ibid, 86.

The rest of this section discusses the distinctive characteristics of volitional necessity, how Frankfurt's idea of care and volitional necessity fits into his larger philosophy, and why these concepts are important to the contemporary discussions of free will and moral responsibility. Finally, I will offer a brief critique of volitional necessity. Frankfurt's notions of care and volitional necessity are related closely. Frankfurt uses love as the prime example of a volitional necessity, and "…love is a particular mode of caring."[138] Frankfurt uses the term "volitional" to highlight the fact that this sort of necessity is not logical in nature. It does not imply a contradiction, nor does it have any metaphysical basis.[139] This kind of necessity has to do with the will. In an instance of volitional necessity, a person is driven by an irresistible or constraining passion. She finds that she has no choice but to accede to the force of her own passion. The idea is that the agent's own will control what she does, even though the agent's "will is not itself within the scope of his voluntary control."[140] In this state, the agent's free choices necessitate that she follow a certain course. Thus, what the agent can act on volitionally is significantly limited. Certain actions are ruled out, not by anything external to the agent, but by the agent's own freely chosen will. [141]

Frankfurt notes that a "person who is subject to volitional necessity finds that he must act as he does."[142] This passivity is desirable when it plays a role in volitional

---

[138] Frankfurt, *Taking Ourselves Seriously*, 40.

[139] "The Importance of What We Care About," 86.

[140] Ibid, 88.

[141] This does not mean, however, that the agent does not experience first order desires that conflict with his highest order will. It is a matter of the desire being effectual that the agent wants to be effectual.

[142] "The Importance of What We Care About," 86-87.

necessity. However, a similar passivity can be undesirable when it takes other forms similar to yet distinct from volitional necessity. (OBSESSIONS?) To distinguish volitional necessity from its counterfeits, Frankfurt's distinction between the willing and unwilling addicts becomes useful. Neither of these individuals is fully free when they desire the drug. All three of these agents *look* as though they have volitional necessities, but only the agent who wills freely can also will wholeheartedly. Recall that the willing addict is *not* free despite the support of his higher-order desires and the lack of ambivalence about taking the drug because "his desire to take the drug will be effective regardless of whether or not he wants this desire to constitute his will."[143] So, a willing addict has a volitional "disability" or compulsion and cannot act against her will for the drug even if she wanted to. She must suffer her own volitional tendencies against her own will (although the willing addict actively supports her will). The unwilling addict is also similar in certain respects to the addict and the agent subject to genuine volitional necessity. In all three cases the agent's "will is not itself within the scope of his voluntary control." The unwilling addict is also like the addict in that she has an unrelenting craving that she cannot deny even if she wants to. However, the unwilling addict and the willing addict are separated in that the unwilling addict does not endorse or identify with her desire for the drug. The ambivalence in the unwilling addict's will makes her addiction one that she wants to change but cannot.

The agent who wills wholeheartedly and is subject to volitional necessity is different from both the willing and unwilling addicts. "Unlike the (unwilling) addict, he does not accede to the constraining force because he lacks sufficient strength of will to defeat it. He accedes to it because he is *unwilling* to oppose it and because, furthermore,

[143] Frankfurt, "Freedom of the Will and the Concept of a Person," 24-25.

his unwillingness is itself something which he is unwilling to alter."[144] In other words, a person subject to volitional necessity *wants* to will as she does, despite her lack of power to will otherwise. She is *active* in support of her volitional necessities. While an aspect of passivity to volitional necessity presents itself as similar to that found in both addicts, volitional necessity is also an active state in that the agent wants her will to be what it is.

Another major difference emerges between the agent subject to volitional necessity and both a willing and unwilling addict. While the most important characteristic that the willing and unwilling addicts have in common is that they both lack freedom of the will, Frankfurt asserts that the agent subject to volitional necessity enjoys true freedom. As noted in section five, Frankfurt claims that "[a] person's will is free only if he is free to have the will he wants."[145] By calling the willing addict unfree Frankfurt is moving beyond a simple compatibilistic definition. He adds the stipulation that volitional necessity must allow at least some room for a person to alter her commitments, since the willing addict could not will otherwise even if he wanted.[146] However, the fact that a person's will is subject to a kind of necessity means that he "cannot deliberately…stop himself from loving. His wholeheartedness means, by definition, that he has no reservations or conflicts that would move him to initiate or to

---

[144] Frankfurt, "The Importance of What We Care About," 87.

[145] Frankfurt, "Freedom of the Will and the Concept of a Person," 24.

[146] Or he must add an historical element to his notion of free and responsible willing. In other words, the willing agent is not free because he did not play an active role in cultivating the characteristics that now necessitate his will and makes it that he cannot will otherwise; whereas, the willing addict did play an active role. However, I cannot see what would make the wholehearted non-addicted agent active other than the fact that the agent was the source of her actions and had access to alterative possibilities at some point in the past.

support such an attempt."[147] The agent subject to volitional necessity experiences her own will as a kind of constraining force that cannot be changed through a singular act of will. In this respect she is similar to the addict.

However, the agent subject to volitional necessity also is distinguished from the addict because the presence of this constraining force does not mean that once a person is subject to volitional necessity she never can alter that state intentionally:

> We can sometimes take steps that inhibit us from loving, or steps that stimulate us to love; more or less effective precautions and therapies may be available, by means of which a person can influence whether love develops or whether it lasts. Love is nonetheless *involuntary*, in that it is not under the *immediate* control of the will. We cannot love—or stop loving—merely by deciding to do so.[148]

The last sentence of this quotation is crucial. Volitional necessity does not mean that the master projects and loves of the agent are irreversible, but that a single act of will cannot change them.[149] This goes against Michael Bratman's comments on Frankfurt's recent lectures. Bratman argues that what Frankfurt calls volitional necessities are not necessities at all but rather revisable, though wholehearted and psychologically entrenched, commitments.[150] Bratman is convinced that volitional necessity is identical to volitional incapacity, and thus wants to make a distinction between volitional necessity

---

[147] Frankfurt, *Taking Ourselves Seriously*, 45. I understand volitional necessity and wholeheartedness to be identical.

[148] Ibid., 41. Emphasis mine.

[149] Charles Taylor's examples are useful here. Whereas I can change my mind about whether I want cake or ice cream (or even which is my favorite), I cannot in a moment go from being a person who has spent my life fighting for justice to a person who no longer cares about justice at all. It may happen slowly (and even semi-intentionally), but it cannot be the result of a single act of will. "What is Human Agency?" In *Human Agency and Language: Philosophical Papers 1* (Cambridge: Cambridge University Press, 1985).

[150] Bratman, Michael, "A Thoughtful and Reasonable Stability: A Comment on Harry Frankfurt's 2004 Tanner Lectures" in *Taking Ourselves Seriously & Getting It Right*, 81.

and wholeheartedness. However, in his discussions of addiction, Frankfurt makes clear that he does *not* equate volitional necessity with incapacity. Furthermore, in the same lectures on which Bratman is commenting, Frankfurt says that we can take steps to inhibit or stimulate our love; and, since love is a prime instance of volitional necessity, Frankfurt must hold that the agent subject to volitional necessity retains a level of freedom unavailable to the addict.[151] Frankfurt says that volitional necessities are not under our *immediate* control. I argue that comments like these mean that Frankfurt's understanding of volitional necessity makes room for the intentional cultivation of volitional necessities.[152] However, Frankfurt does not make clear exactly how this intentional cultivation might take place. Here is where Kierkegaard is a useful corrective and supplement to Frankfurt. I return to this point in a moment.

The assertion that volitional necessity consists in a certain structure of the will where the agent now *must* act as she does appears at odds with the claim that an agent subject to volitional necessity retains some freedom to change his commitments intentionally. I think that this is why Frankfurt moves away from the term volitional necessity in his later writings and often substitutes the term wholeheartedness. Though the opposite of volitional necessity might seem to be libertarian freedom, Frankfurt argues that the opposite of wholeheartedness is not freedom but *ambivalence*. He

---

[151] One might object here that there is no difference at all, from Frankfurt's perspective, between the person with volitional necessity or love, and the willing addict. Just as Luther loves his God, so the addict loves his cocaine. Given that there are no objective standards applicable, there seems to be no difference between the two except a matter of taste in objects of love. I grant that while it may be possible that Frankfurt does not make a firm distinction between the willing addict and the wholehearted agent, he does say that the wholehearted agent has penultimate freedom while the addict does not. I am simply trying to discern—based on Frankfurt's own comments—the basis of this distinction.

[152] Cf. Kierkegaard on patience: "'to gain one's soul'…immediately turns the mind to a *quiet but unflagging activity*…not of making a conquest, of hunting and seizing something, but of becoming more and more quiet…" "To Preserve One's Soul in Patience," *Eighteen Upbuilding Discourses,* 170-71.

understands ambivalence as a kind of psychic instability or conflict within a person's will.[153] He seems to agree with Kierkegaard that the opposite of wholeheartedness (or single-mindedness as Kierkegaard calls it) is double-mindedness, or a divided will.[154] In this instance the agent "is moved by incompatible preferences or attitudes regarding his affects or his desires or regarding other elements in his psychic life."[155] In other words, the agent has higher-order endorsements about incompatible first-order psychic elements, and as a result lacks a robust motivational identity because she does not know what she wants.[156]

Wholeheartedness, on the other hand, is not the lack of all internal opposition. Humans have all kinds of desires and emotions, many of which are in direct conflict with each other. What wholeheartedness requires is that an agent must be fully resolved with respect to those conflicting elements; she must be satisfied with having certain psychic elements instead of others.[157] Satisfaction does not require the "adoption of any cognitive, attitudinal, affective, or intentional stance." "Satisfaction is a state of the entire psychic system—a state constituted just by the absence of any tendency or inclination to alter its condition."[158] A person who is wholehearted is satisfied with the

---

[153] Frankfurt, "The Faintest Passion," in *Necessity, Volition, and Love*, 98.

[154] See Kierkegaard, *Purity of Heart is the Will One Thing*.

[155] Frankfurt, "The Faintest Passion," 99.

[156] Or perhaps more accurately, she wants goods that are incommensurable and thus only one of his higher preferences can be fulfilled.

[157] "The Faintest Passion," 103.

[158] Ibid., 104.

constitution of her will and has no interest in changing that will despite the possibility of conflicting psychic elements.[159]

Wholeheartedness, Moral Self-Cultivation and Morally Responsible Agency

Perhaps it is not clear *prima facie* how Frankfurt's notion of volitional necessity fits into a project aimed at explicating the freedom-relevant conditions necessary for grounding morally responsible agency. I am asking what freedom-relevant conditions must be met in order for a person to be responsible for her *character*, not simply for particular actions. I ultimately make the case that a libertarian notion of free will is necessary to satisfy these conditions. However, the kind of libertarian position that I am articulating is a *modest* libertarianism that takes seriously certain insights of the compatibilist tradition. In particular, I look at the notion that an agent does not need access to alternative possibilities at every moment in order to be responsible for either a particular action or character. Furthermore, I argue that Kierkegaard's thinks that a person can cultivate her character freely and come to the point where she is, in a sense, unable to will otherwise and yet truly is free. Kierkegaard is worth quoting at length here:

> What Augustine says of true freedom (distinguished from freedom of choice) is very true and very much a part of experience—namely, that a person has the most lively sense of freedom when with completely decisive determination he impresses upon his action *the inner necessity which excludes the thought of another possibility*. Then freedom of choice or the 'agony' of choice comes to an end.[160]

In these two sentences from Kierkegaard's journals, he hints at a position that I argue at length in chapter four: the cultivation of character takes libertarian freedom (what

[159] Ibid., 105.

[160] Kierkegaard, *Journals and Papers* X4 A 177. Emphasis mine.

Kierkegaard here calls the freedom of choice) and the point of having those choices is to cultivate one's character so that the person is free to care wholeheartedly about what she wants to care about. In this state of wholeheartedness, the agent may no longer have real options in the way that she did while she was cultivating that state. In other words, although the freedom-relevant conditions necessary for cultivating morally responsible agency include access to alternative possibilities, an agent still can be responsible for her character even when she no longer has access to alternative possibilities as long as she has been the source of actions and made live choices that have led to her now formed state of character.

CHAPTER THREE

Fischer's "Semicompatibilism"

Introduction

Frankfurt argues that the only necessary agential properties required for responsibility ascriptions are the right mesh of higher and lower order desires at the moment of action—a time-slice mesh theory of moral responsibility.[1] In the second chapter I identified two key objections to Frankfurt's theory, the problems of infinite regress and authority. The authority objection claims that Frankfurt fails to give an account of the freedom or responsibility-granting nature of higher-order desires. This objection is the most important, and most common. Frankfurt appears committed to the idea that in order to be responsible for one's actions an agent must be the *source* of her actions, and this sourcehood distinguishes a responsible agent from an addict. However, Frankfurt's notion of second-order desires *alone* fails to account for the origin of this sourcehood. Frankfurt admits that even if a manipulator succeeded in giving an agent an entirely new character (which would surely include second-order desires), the agent still is responsible for the action that stems from that character. Therefore either Frankfurt's hierarchical model of the will needs to be modified to include some element(s) that will answer the authority objection, or needs to be discarded altogether. I think that discarding the hierarchical view is hasty, and that the hierarchical model (when properly modified) can solve certain problems that competing views fail to solve.[2] The two most

[1] Recall that I am arguing that free will ascriptions require that the agent take responsibility for the cultivation of both her character as a whole and for particular actions. This richer understanding of moral responsibility proves problematic for Frankfurt.

[2] I must hold off making this argument until the fifth and final chapter.

compelling responses to and modifications of Frankfurt's hierarchical model are articulated by Eleonore Stump and John Martin Fischer, philosophers who accept certain aspects of Frankfurt's theory of the will and responsibility while rejecting or modifying other aspects.[3] These modifications yield more coherent and systematic Frankfurt-style theories of the freedom-relevant conditions necessary for grounding moral responsibility.[4] In large part they succeed, yet both ultimately prove deficient. I think that a systematic articulation of Kierkegaard understanding of free will and moral responsibility provides contributions similar to Stump and Fischer without the correlating deficiencies. Kierkegaard's model yields an understanding of the freedom-relevant conditions of moral responsibility that avoids the objections to which Stump and Fischer fall prey, objections that will become clear shortly.

I argue that ultimately Fischer's modification to Frankfurt's theory reveals the most fundamental problem with both Frankfurt's and Fischer's understanding of the freedom-relevant conditions necessary or moral responsibility: the inability to differentiate between a morally responsible agent and an agent whose actions result from *responsibility-undermining external forces* such as manipulation. While Fischer offers the best articulation of semicompatibilism to date, he cannot offer a solution to this problem. So I find good reason to turn to incompatibilist accounts of the freedom-relevant conditions necessary for grounding moral responsibility to see if they fare better.

---

[3] Fischer does not directly answer the objection, but instead rejects Frankfurt's hierarchical model. However, he is still responding to this objection and he retains key elements of Frankfurt's project.

[4] Neither Fischer nor Stump refers to the necessary conditions for the cultivation of character, but I think that I can adapt their ideas, like Frankfurt's, to my project.

Fischer's Semicompatibilism

Fischer calls himself a "semicompatibilist" because he agrees with certain premises of classical compatibilism while rejecting (or remaining uncommitted) to others. He accepts "a basic commitment of the compatibilist…that not all causally deterministic sequences undermine freedom equally."[5] He also agrees that a certain kind of freedom—the freedom necessary for moral responsibility—is compatible with determinism. However, he is not committed to the "conditional analysis" notion of freedom that many classical compatibilists embrace.[6] Roughly, the conditional analysis view of freedom states that a certain kind of freedom is compatible with determinism, namely, the notion that the agent "could have done otherwise just in case he would have done otherwise, if he had chosen to do otherwise."[7] In other words, the agent was free to do what she wanted. Had she wanted to act differently she could have, but since she did not want to act differently than she did only one option was available to her. Fischer rejects the conditional analysis mainly because of the problems surrounding phobias, addictions, and so forth that have been introduced by Frankfurt, Watson, and others.[8] In an instance such as a phobia the agent could *not* have chosen otherwise even if she *wanted* to because the phobia blocks the alternative action.

---

[5] Fischer, "Compatibilism" in *Four Views on Free Will: Great debates in Philosophy,* Malden, (MA: Blackwell Publishing, 2007), 53.

[6] The conditional analysis is referred to in a number of ways in the literature. Kane (1996) refers to the conditional analysis of "can," while Timpe (2007) and others refer to subjunctive analyses of "could have done otherwise." Although there are many fine distinctions between the various versions of this argument, for my purpose all I need to show is that Fischer agrees with Frankfurt that conditional analyses of freedom (in whatever form) fail to account for the difference between free action and action that results from manipulation or phobia.

[7] Fischer, "Compatibilism," 49. Kane has an excellent overview of the various conditional analysis positions in *The Significance Of Free Will*, 51-60.

[8] Ibid., 50.

Fischer uses the example of Thomas to show why he thinks the conditional analysis of freedom is inadequate. Thomas had a traumatic encounter with a snake when he was a young boy. As an adult, he suffers from a pathological aversion to snakes. He cannot pick up or even come near a snake. Now Thomas is in a situation where a snake is on the ground in front of him, and he needs to walk near the snake to get to his child who is on the other side of the snake. Thomas cannot go around the snake; he cannot get that close because of his pathology. Were the conditional analysis of freedom true, then Thomas could walk around the snake if he wanted to. Of course he wants to walk around the snake because he wants to get to his child, but his phobia will not allow him. Fischer thinks that this example and numerous others show that the conditional analysis of freedom cannot account for situations that seem to render an agent "psychologically incapable of choice."[9] Fischer accepts that a number of revisions could be made to the conditional analysis, but at a point the "refined" conditional analysis is no longer helpful in seeking to resolve the original problem, namely, whether causal determinism is compatible with freedom.[10]

---

[9] Ibid., 50. Gary Watson argues that this is not actually an issue about freedom of the will, but about what he calls "normative competency" or the minimum proper psychology function that is necessary for an agent to make responsible choices. This claim is connected to Watson's correlating claim that "there is no such thing as failing to will; willing is necessarily successful." Watson thinks that manipulation of phobia prevents the agent from willing as such, it does not impair the agent's volitional faculty but instead bypasses that faculty. Although there is something to this point I believe that Watson is ultimately wrong, and that this is a reason to endorse the hierarchical view of the will that Watson rejects. In Fischer's instance what is impaired is the agent's ability to endorse the desire that he wishes to endorse. Thomas wants to go around the snake, and he even has he second-order volition to make this desire his will. But his phobia prevents him from doing making the desire to go around the snake his will. Perhaps Watson would respond that the only act of will that Thomas performs is the will not to go around the snake, and that act of will was successful. But this is an oversimplified view of the volitional states of agent's. Agents can have various desires that are reflectively endorsed but cannot be willed because of various phobias, manipulation, etc. So perhaps the debate comes down to semantics, but still remains an important debate…

[10] Ibid., 51.

Fischer holds to semicompatibilism because he thinks that humans do have a certain sort of freedom and control as agents—control necessary for moral responsibility—yet he does not believe that access to alternative possibilities is a freedom-relevant condition for moral responsibility. In other words, he thinks that a basic argument[11] for the incompatibilism between freedom and causal determinism is extremely plausible (though he remains unconvinced of its indefeasibility). Causal determinism states that at any given time the events of the past and the fixity of the laws of nature entail every future event.[12] Fischer agrees with van Inwagen's "consequence argument," which states that "[i]f determinism is true, then our acts are the consequences of the laws of nature and the events of the remote past. But it is not up to us what went on before we were born, and neither is it up to us what the laws of nature are. Therefore the consequences of these things (including our present acts) are not up to us."[13] Fischer articulates the consequence argument this way:

> If causal determinism obtains, then (roughly speaking) the past together with the natural laws entail that I act as I do now. So if I am free to do otherwise, then I must either have the power over the past or power over the laws of nature. But since the past and the laws of nature are "fixed"—for instance, I cannot now so act that the past would be different from what it actually is—it follows that I am not now free to do otherwise.[14]

What Fischer finds convincing about this argument is that if causal determinism is true, then the consequence is that humans lack a certain kind of freedom, namely the freedom

---

[11] Fischer tends to use the basic argument and the consequence argument interchangeably.

[12] Fischer, "Responsiveness and Moral Responsibility," in *My Way: Essays on Moral Responsibility*, 76.

[13] Van Inwagen, Peter, *An Essay on Free Will*, (Oxford : Clarendon Press, 1983), 16. Quoted in Kane, *The Significance of Free Will,* 44. Although van Inwagen gives three different versions of the consequent argument, he regards the three as versions of the same basic argument.

[14] Ibid., 77.

to do otherwise than they actually do. Yet he remains convinced that the best procedure is to follow common intuitions about clear cases concerning freedom and moral responsibility. One of our most basic intuitions is that a responsible person has a certain kind of freedom or control over her actions that a non-responsible (for instance manipulated or phobic) person lacks. The consequence argument does not mean that causal determinism is incompatible with moral responsibility if one can show that the kind of freedom or control necessary for moral responsibility is something other than the freedom to do otherwise. In Fischer's own words, "semicompatibilism is the claim that causal determinism is compatible with moral responsibility, quite apart from whether causal determinism rules out the sort of freedom that involves access to alternative possibilities."[15] Fischer thinks that freedom—defined as alternative possibilities—is incompatible with determinism, but he does not think that causal determinism undermines the freedom-relevant condition of *control* for moral responsibility.

Fischer's Response to Frankfurt

Fischer is influenced heavily by Frankfurt, particularly by Frankfurt's famous arguments against the "principle of alternative possibilities."[16] Fischer thinks that he can construct an *FSC* (Frankfurt-style case) that demonstrates that the freedom that is necessary for moral responsibility is compatible with determinism and therefore defeats (or at least makes a very plausible argument against) the direct argument for the incompatibilism of causal determinism and moral responsibility.[17] Fischer modifies

[15] Fischer, "Compatibilism," 56.

[16] These arguments are dealt with at length in the second chapter.

[17] Direct and indirect arguments are discussed in the following section.

Frankfurt's example of Jones and Black to say that Jones is going to vote and plans to vote for the democratic candidate, and Black plans to ensure that Jones does vote for that candidate. Fischer follows Frankfurt's lead in suggesting that Black will intervene if Jones attempts to vote otherwise and that Black's potential intervention does not take away Jones's moral responsibility in voting. Black never actually intervenes; he simply waits ready to intervene if Jones is going to vote for the republican candidate, and thus takes away his options. Fischer (like Frankfurt) thinks that most people intuitively judge Jones as responsible for the way that he voted, because he voted the way that he wanted to vote. Nobody made him vote that way; he was not coerced or manipulated in any way.[18]

Fischer thinks that *FSCs* like the one with Jones and Black make a persuasive argument that "it is plausible that alternative possibilities are irrelevant to ascriptions of moral responsibility."[19] What makes an agent responsible is quite possibly something that happens in the actual sequence (Jones's decision and subsequent action of voting for the democratic candidate) and has nothing to do with the alternative scenario (Jones makes a different decision and votes for the democratic candidate). In other words, Jones's responsibility comes from what he *actually* does and is able to do and has nothing to do with whether or not he is able to perform an action *different* from his actual action. For Frankfurt, the freedom-relevant condition that grounds ascriptions of moral responsibility is the right mesh between higher and lower order desires (not alternative possibilities); and, Frankfurt seems to think that this condition is compatible with causal

[18] Once again, the key distinction here is between free action and action that is the result of manipulation, phobia, etc. This distinction becomes very important later in the chapter.

[19] Fischer, "Frankfurt-Style Compatibilism" in *Contours of Agency*, edited by Sarah Buss and Lee Overton, (The MIT Press, Cambridge MA, 2002), 7.

determinism. Fischer agrees with Frankfurt that alternative possibilities are not necessary for freedom, and that the freedom relevant conditions for moral responsibility are present in the actual sequence and not the alternative sequence, but he disagrees with Frankfurt on the nature of the freedom-relevant condition.

Frankfurt argues that the freedom relevant condition for moral responsibility is having the right "mesh" of higher and lower order desires because this mesh demonstrates that the agent is the source of her actions in the necessary way. Fischer argues that mesh theories in general are inadequate. "The problem with such hierarchical 'mesh' theories, no matter how they are refined, is that the selected mesh can be produced via responsibility-undermining mechanisms…I believe that the problem with the hierarchical mesh theories is precisely that they are *purely structural* and *ahistorical*. It matters what kind of process issues in an action."[20] Later, he states that a mesh like this could be "produced by certain sorts of brainwashing or subliminal advertising…(or) direct stimulation of the brain." In either case the "manipulative history blocks responsible agency."[21] I look at an elaborate example of this below when I turn to Stump's critique of Fischer. The basic argument states that if all that is necessary to ascribe moral responsibility to an agent is the right mesh of desires or a certain configuration of mental states, then it seems entirely possible that an external agent could cause this mesh through manipulating the agent's brain.[22] The agent no longer wants what she wants to want, but wants what the external manipulator wants her (and causes

---

[20] Fischer, "Responsiveness and Moral Responsibility," 79. My emphasis.

[21] Fischer and Ravizza, *Responsibility and Control,* 197. Cited in Bratman, Michael E., "Fischer and Ravizza on Moral Responsibility and History," *Philosophy and Phenomenological Research* 61 (2000), 456.

[22] We saw in the previous chapter that Frankfurt even admits that this is the case!

her) to want. Fischer and Ravizza state that "[w]hen the harmony (or mesh) in the selected mental ingredients—whatever they are—is produced in these ways, the mechanism that issues in the relevant behavior is *not*, in an important intuitive sense, the *agent's own*."[23] It is not clear what Fischer's precise criticism is here, but he suggests that the manipulated person is not responsible for her actions because she is not the *source* of her actions or does not have the requisite *control* over her actions.[24] Fischer thinks that something more than a momentary mesh of desires is necessary for moral responsibility, and thus believes that despite the attractive features of Frankfurt's implicit semicompatibilism, Frankfurt's theory needs to be corrected in a fundamental way. Frankfurt needs an account of sourcehood and control that makes sense of the intuitive difference between a manipulated agent and one who acts freely on her own.[25]

Partly because Frankfurt lacks an adequate account of the freedom-relevant conditions of sourcehood and control, Fischer thinks that Frankfurt fails to give any kind of response to the direct argument for compatibilism (the argument that causal determinism is incompatible with moral responsibility apart from the issue of alternative possibilities). Further, Fischer thinks that an adequate response to the direct argument is necessary to show that semicompatibilism can provide the freedom-relevant conditions for moral responsibility. Therefore Fischer applies his own notions of sourcehood and

---

[23] Ibid.

[24] As stated in chapter 1, Frankfurt does agree that there is a source component necessary for grounding moral responsibility ascriptions, and he thinks that his mesh theory meets that requirement despite the fact that a manipulator could case that mesh.

[25] Frankfurt agrees that we need an adequate account of agency, and of "the difference between being passive and being active" in order to explain the compatibility of determinism and human agency fully. However, Fischer does not think that Frankfurt offers an adequate account of agency and control. This criticism of Fischer's becomes very important later, when I seek to show that Fischer's account is subject to the same criticism, which shows that there is no adequate articulation of sourcehood or ownership that is compatible with causal determinism. See Frankfurt's "Reply to John Martin Fischer," 29.

control to formulate an argument against the direct argument. I examine that argument below, but first I discuss the direct and indirect arguments for the incompatibilism of freedom and moral responsibility. I then discuss Fischer's understanding of the sourcehood and control components of moral responsibility, what he calls "guidance control."

Direct and Indirect Arguments for Incompatibilism

In their book *Responsibility and Control: A Theory of Moral Responsibility*, Fischer and his co-author Mark Ravizza attempt to show that an agent can be morally responsible for her actions even when those actions are causally determined. They think that Frankfurt successfully defeats indirect arguments for incompatibilism. The basic structure of an indirect argument is as follows: moral responsibility requires alternative possibilities, causal determinism rules out alternative possibilities, and therefore moral responsibility is incompatible with determinism.[26] However, indirect arguments are only one of two major clusters of arguments in favor of incompatibilism. Direct arguments for incompatibilism argue that causal determinism is *directly* incompatible with determinism. In other words, even if alternative possibilities are not necessary for responsibility ascriptions (which FSCs are designed to demonstrate), causal determinism still is incompatible with moral responsibility because moral responsibility requires something more than alternative possibilities, and this 'something more'—whatever it is—is also incompatible with causal determinism. Since Frankfurt style cases do not address the direct arguments, Fischer and Ravizza can defend semicompatibilism successfully only if

[26] Van Inwagen discusses various types of indirect arguments in *An Essay on Free Will*, see especially 56-105.

they defeat the direct argument for incompatibilism by showing that an agent can be responsible for her actions even if those actions are causally determined.[27]

Fischer and Ravizza develop three arguments that they think establish the compatibility of moral responsibility and causal determinism. The first argument attempts to defeat the indirect arguments for incompatibilism by utilizing Frankfurt style cases to show the compatibility of moral responsibility and causal determinism. They argue for this assertion because the freedom-relevant condition for moral responsibility is guidance control and not regulative control (I discuss these issues in the next section). The second argument attempts to show that direct arguments for incompatibilism fail as well. Finally, they attempt to show that in cases of manipulation a distinction must be made between the causal determination of an agent's mechanism by nonintelligent causes and the causal determination of an agent's mechanism by a manipulator. More specifically, they argue that in the case of nonintelligent causal determination the agent is still morally responsible (due to the guidance control that the agent maintains) whereas in the cases of manipulation (by another intelligent agent) the agent is no longer responsible because she no longer maintains the requisite control. I deal with each of these in turn after discussing guidance control.

Guidance Control

On the Fischer/Ravizza account, the freedom-relevant condition for grounding moral responsibility is "guidance control." In instances in which an agent exhibits guidance control her "actions flows from the agent's own, moderately reasons-responsive

[27] Stump thinks that direct arguments fail while indirect arguments succeed. I argue at length against the indirect arguments and Frankfurt-style cases in chapter four.

mechanism."[28] Fischer and Ravizza oppose guidance control to regulative control, or the agent's ability to regulate "which of two alternative and mutually incompatible states of affairs comes about."[29] Guidance control focuses on the actual sequence of the action (as opposed to the alternative or counterfactual sequence). An agent has guidance control when she is able to do what she wants without any special hindrances, and furthermore can do what she wants on *purpose* and not on accident. In a typical Frankfurt style case, the agent exhibits guidance control because her action (which occurs in the actual sequence) is the result of normal practical reasoning and deliberation. The counterfactual action in the alternative sequence in FSCs is caused by direct stimulation of the brain and does not result from the agent's normal reasoning.[30] Thus the basic feature of an agent with guidance control is an unhindered deliberative faculty that produces an action that the agent wished to perform.[31]

To illustrate guidance control, Fischer imagines a situation where a person, Jeff, is driving his car and wishes to make a left-hand turn. Assuming that the car is functioning properly (the steering column and brakes are working) and there are no external hindrances that keep Jeff from making a turn, he is able to apply pressure to the brake, turn his signal on, and turn the steering wheel, thus guiding his car to the left as he intends to do. Fischer thinks that this is the kind of control that humans have when they

---

[28] Fischer and Ravizza, *Responsibility and Control,* 207. Cited in Bratman, "Fischer and Ravizza on Moral Responsibility and History," 453.

[29] Glannon, Walter, "Responsibility and Control: Fischer's and Ravizza's Theory of Moral Responsibility," *Law and Philosophy* 18 no. 2 (1999), 198.

[30] Fischer, "Responsiveness and Moral Responsibility," in *My Way: Essays on Moral Responsibility*, 66.

[31] Here, as throughout this project (unless otherwise noted), I am using action in the broadest possible sense to include willings, intentions, physical actions, etc.

are morally responsible. We have the ability to guide ourselves intentionally down a certain path. Furthermore, it is what happens in this sequence, not the possibility of performing a different action, which is important for moral responsibility. To emphasize that it is what occurs in the actual sequence and not the alternative sequence that is relevant to responsibility ascriptions, Fischer alters the scenario so that something is wrong with the steering column in Jeff's car so that he cannot turn in any way but left. Fischer thinks that the fact that Jeff could not turn in any other direction than the direction that he wanted is not relevant since he was able to guide the car in the direction that he wanted to go. In other words, what is crucial about guidance control is "not that we have the ability to do otherwise, but that we acquire our reasons autonomously and act on them in an uncoerced and uncompelled way."[32]

Fischer and Ravizza give a detailed account of the nature of guidance control. Based on the definition of guidance control exhibited in instances when the agent's "action flows from the agent's own, moderately reasons-responsive mechanism," three main issues need explaining: the concepts of "mechanism," "reasons-responsiveness," and "ownership" or what makes that mechanism an agent's own. I will deal with each of these in turn.

Mechanism

Fischer and Ravizza's use of the word mechanism is in certain ways unfortunate, and is intentionally left largely undefined. Fischer and Ravizza stress that they do not mean to reduce agency to a mechanistic process. Instead, by mechanism they mean the

[32] Glannon, "Responsibility and Control," 188.

"faculty of practical reasoning and the process of practical deliberation"[33] from which the action flows and which can also include "non-reflective mechanisms of various kinds."[34] Furthermore, in this context, the usage of the word mechanism signals a shift away from an agent-based approach to moral responsibility. The emphasis in an agent-based approach "is on *who* the agent is and *what* the agent does (or is not able to do). In contrast, on a mechanism-based approach the emphasis is on the *process* leading to the action, or *how* the action comes about."[35] This distinction shows that Fischer and Ravizza agree that Frankfurt style cases provide useful insights. Specifically, the same agent with the same history can perform two different actions in the same circumstance, but the agent is responsible in one instance (the instance where the agent does what she wants without the intervener's manipulation) due to something that happens in the actual sequence leading up to the action, not due to access to alternative scenarios that she can actualize. Fischer states that

> [i]t seems to me impossible that there be cases in which there are two agents who perform actions of the same type as a result of exactly the same kind of actual causal sequence, but in which one agent is morally responsible for the action and the other is *not*. Differences in responsibility ascriptions must come from differences in the actual physical factors resulting in action; mere differences in alternative scenarios do not translate into difference in responsibility ascriptions.[36]

In other words, an agent's ability to do otherwise is not relevant to ascriptions of moral responsibility. What proves important is that a certain kind of deliberation took place, and that the agent has a certain kind of control over her actions. In other words,

---

[33] *Ibid.,* 190.

[34] Fischer and Ravizza, *Responsibility and Control,* 86.

[35] Glannon, "Responsibility and Control," 191.

[36] Fischer, "Responsiveness and Moral Responsibility," *My Way: Essays on Moral Responsibility*, 72.

according to Fischer and Ravizza what matters is the kind of mechanism that produced the action.

Reasons-Responsiveness

Fischer and Ravizza are among Frankfurt's critics who think that the right mesh alone cannot adequately differentiate responsible agents from non-responsible (i.e. manipulated) agents. Therefore Fischer and Ravizza mainly focus on differentiating responsible agents—ones who have the requisite control—from agents who are not intuitively responsible. Fischer and Ravizza argue that reasons-responsiveness and ownership can differentiate between these two by showing how the actions of responsible agents originate from them in ways that the actions of non-responsible agents do not. In other words, Fischer and Ravizza think that their understanding of reasons-responsiveness and ownership can account for the necessary grounding responsibility ascriptions of sourcehood and control.

Fischer and Ravizza try to clarify what makes for an agent's own responsible action. They argue that the agent must be responsive to reasons in her practical deliberation to distinguish her action from the action of agent's who are intuitively non-responsible, such as an agent who is subject to severe phobias or manipulation of the brain.[37] Fischer argues that

> the causal history of the action matters to us in making moral responsibility attributions. When persons are manipulated in certain ways, they are like marionettes and are not appropriate candidates for praise or blame. Certain factors issuing in behavior are, we understand intuitively, responsibility-undermining factors.[38]

[37] Here is where Fischer and Ravizza bring in the epistemic condition.

[38] Fischer, "Responsiveness and Moral Responsibility," 65.

One of the basic responsibility-undermining factors is the lack of a reasons-responsive mechanism. In its most basic form, reasons-responsiveness has to do with the connection between (i) what reasons are available to and recognized by the agent (*receptivity*),[39] (ii) the agent's reason(s) and choice (*reactivity*), and (iii) the agent's choice and action (*translation*).[40] Fischer and Ravizza understand that these connections can be strong or weak, or that the agent might recognize reasons and still fail to react to them, etc. Thus they distinguish between three levels or reasons-responsiveness, strong, weak, and moderate.

Strong and weak reasons-responsiveness. Strong reasons-responsiveness is displayed by an agent who properly responds to, reacts to, and translates reasons into action. Furthermore, "[s]trong reasons-responsiveness obtains when a certain kind *K* of mechanism actually issues in an action and if there were sufficient reason to do otherwise and *K* were to operate, the agent would recognize the sufficient reason to do otherwise and thus choose to do otherwise and do otherwise."[41] Fischer and Ravizza utilize the following example to illustrate the difference between strong and weak reasons-responsiveness.[42] Suppose Bob the college student has a sufficient reason to go to a movie, yet has an even stronger reason to stay home and write a paper that is due the following day. Since Bob has sufficient reasons for performing either action, he would

---

[39] This is an intentionally vague phrase that I will refine. Fischer and Ravizza are concerned that responsibility not be a purely internalist notion, but that responsibility have to do with responding in the proper ways to reasons that are presented to the agent by the situation. These are reasons that in at least some cases could be recognized by a third party.

[40] Fischer, "Responsiveness and Moral Responsibility," 67.

[41] Ibid., 67.

[42] Fischer and Ravizza, *Responsibility and Control,* 42. Cited by Glannon, "Responsibility and Control," 192.

display reasons-responsiveness in either case. However, if Bob goes to the movie he displays weak reasons-responsiveness; whereas, if he stays home to study he displays strong reasons-responsiveness. One important feature common to both strong and weak reasons-responsiveness is the ability to recognize and respond to a reason for action. Bob has good reason to perform either action, but a much better reason to stay home. Thus he is weakly responsive to reasons when he deliberately chooses to go to the movie. In earlier work, Fischer and Ravizza argued that weak reasons-responsiveness was all that was required for moral responsibility. But since weak reasons-responsiveness is consistent with unusual patterns of behavior, they now argue that moderate reasons-responsiveness is necessary for moral responsibility.[43]

Moderate reasons-responsiveness. Fischer and Ravizza claim that "a mechanism leading to action is moderately reasons-responsive if (i) it can recognize at least a significant range of reasons (including moral reasons) for and against so acting and (ii) it can adjust behavior in light of and in accordance with at least some of those reasons."[44] Moderate reasons-responsiveness is distinguished from strong reasons-responsiveness where there is a "tight fit" between sufficient reason and action.[45] Furthermore, Fischer and Ravizza state that an agent who exhibits moderate reasons-responsiveness will exhibit a pattern of regular receptivity but only weak reactivity to reasons.

> A person who acts on a regularly receptive mechanism must exhibit an appropriate pattern of reasons-recognition: he must see how the sufficiency of

---

[43] Glannon, "Responsibility and Control," 192.

[44] Bratman, "Moral Responsibility and History," 454. Here is another important difference between Fischer and Frankfurt. Whereas Frankfurt argues that the distinguishing mark of a person is the ability to form higher-order desires, Fischer thinks that at least one of the distinguishing marks of a person is the ability to recognize and respond to reasons when deliberating about what action to perform.

[45] Fischer and Ravizza, "Précis of Responsibility and Control," 443.

> reasons varies as we adopt different sets of values, beliefs, and desires, and he must show an appropriate understanding of how reasons connect with—and relate to—other reasons.[46]

Therefore on Fischer and Ravizza's account, to be ascribed moral responsibility the agent must display moderate reasons-responsiveness when she performs the action in question. Furthermore, they believe that an agent subject to manipulation or phobia cannot display the reasons-responsiveness necessary for moral responsibility. I return to this point below.

Ownership

Fischer and Ravizza insist that Frankfurt fails to differentiate adequately between an agent whose rightly meshed desires are the result of direct manipulation and the agent whose desires are not the result of direct manipulation. They argue that taking ownership for one's own mechanism is the necessary differentiating factor. Eleonore Stump describes the difference this way:

> They [Fischer and Ravizza] maintain that an agent is acting on a reasons-responsive mechanism that is his own if and only if the agent's history included his taking responsibility for that mechanism by taking responsibility for the actions that stem from it: 'The process by which an agent takes responsibility for the springs of an action makes them his own in an important sense.'[47]

Taking responsibility for a mechanism has three components. First, the agent must take a first-person perspective on his behavior. In other words, he must "see himself as the source of his behavior"[48] in at least the minimal sense that he sees that his "desires,

---

[46] Fischer and Ravizza, "*Responsibility and Control,*" 64. Cited in Glannon, "Responsibility and Control," 193.

[47] Stump, Eleonore, "Control and Causal Determinism," in *Contours of Agency*, 47.

[48] Fischer and Ravizza, *Responsibility and Control,* 210. Quoted in Stump, "Control and Causal Determinism," 47.

beliefs, and intentions result in actions and upshots in the world…"[49] Second, when the agent acts on this mechanism he must "accept that he is a fair target of the reactive attitudes [of others] as a result of how he exercises this agency in certain contexts."[50] Last, "the agent's view of himself as an agent and sometimes appropriately subject to the reactive attitudes [must] be grounded in his evidence for these beliefs."[51] If an agent meets these three conditions when acting, then she has acted on a mechanism that is her own. And if she has acted on a mechanism that is her own, then in Fischer and Ravizza's view she has displayed the control necessary for moral responsibility regardless of whether or not that action was causally determined.

Taken together, Fischer and Ravizza think that their understanding of guidance control, involving a moderately reasons-responsive mechanism of the agent's own, adequately responds to the objections to Frankfurt's implicit semicompatibilism. They argue that an agent is responsible for an action despite the truth of causal determinism so long as the agent properly responds to reasons and takes responsibility for the actions that have issued from the agent's mechanism (the agent's deliberative faculty) in the past. Furthermore, they believe that these characteristics adequately differentiate an agent who acts on her own, and thus responsibly, from one whose action results from responsibility-undermining features such as direct manipulation of the brain or compulsive phobia, and thus acts without responsibility.

---

[49] Fischer and Ravizza, *Responsibility and Control,* 208. Quoted in Bratman, "Fischer and Ravizza on Moral Responsibility and History," 454.

[50] Ibid., 211.

[51] Ibid., 213. Here is one place where Fischer and Ravizza's "externalism" is particularly evident. The agent must have a certain kind of connection to the world in order to be acting responsibly.

Direct Arguments for Incompatibilism

Fischer and Ravizza believe that *FSCs* adequately respond to indirect arguments for the incompatibilism of causal determinism and moral responsibility by showing that alternative possibilities are not necessary for ascriptions of moral responsibility. Therefore Fischer and Ravizza's arguments are aimed largely at the direct arguments for incompatibilism. They do not believe that Frankfurt's counterexamples adequately respond to arguments for the *direct* incompatibility of causal determinism and moral responsibility. Fischer and Ravizza summarize the most influential of the direct arguments, what they call the Principle of the Transfer of Non-Responsibility,[52] in the following way:

> If causal determinism is true, then there is some state of the world in the distant past *b* that is connected by the laws of nature to any action *A* that one performs in the present. But since no one (alive now) is even partly morally responsible for the state of the world *b* in the distant past, and no one is even partly morally responsible for the laws of nature that lead from *b* to *A*, it follows that no one is even partly morally responsible for any action *A* that is performed in the present. That is, if no one currently existing is morally responsible for the distant past, and no one is morally responsible for the distant past's leading to current actions (via natural laws), then no one is now morally responsible for current actions.[53]

In shorter form, if event *X* happened and no one is even partially responsible for *X*, and if *X* happens then *Y* happens and no one is even responsible for this (*X* happens then *Y* happens), then no one is even partially responsible for *Y*. To see the intuitive plausibility of this argument, let us consider some event that we would all agree that no one alive today is responsible for, the big bang for example. Given the truth of causal determinism, this event was the one link in a long chain of events that could not have been otherwise given the events that preceded it and the fixity of the laws of nature. Therefore the big

[52] Henceforth *TNR*.

[53] Fischer and Ravizza, *Responsibility and Control,*153.

bang is the first cause[54] of the assassination of JFK. If it is the case that no one is even partially responsible for the big bang, and it is also the case that no one is even partially responsible for the big bang beginning the causal sequence that leads to Lee Harvey Oswald assassinating JFK, then no one is responsible for the assassination of JFK. In other words, given the truth of causal determinism (the past (P) and the laws of nature (L) determine any and all facts about the world (F)) and the fact that no one is even partially responsible for this fact (P and L > F), applying *TNR* leads to the conclusion that no one is even partially responsible for any arbitrary fact about the world.[55] Therefore moral responsibility is directly incompatible with causal determinism without any reference to the existence of alternative possibilities.

Needless to say, Fischer and Ravizza do not accept this conclusion, which would disavow their semicompatibilism because they would be admitting that moral responsibility is not compatible with the truth of causal determinism. Their strategy attempts to show that moral responsibility is possible in a causally determined world. Therefore they counter the conclusion that no one is even partially responsible for any arbitrary fact about the world by referencing cases they refer to as *preemptive* or *simultaneous overdetermination*. For instance, in one of their counterexamples, which they call *Erosion*, Betty uses explosives at *T1* to cause an avalanche with the intention of destroying a camp at the base of a mountain at *T3*. However, if she had not used the explosives then natural causes (the erosion of a glacier) would have started an avalanche at *T2* that would have resulted in the destruction of the camp at *T3*. Fischer and Ravizza

[54] I am not using first cause in any technical Aristotelian sense here. I simply mean the first temporal cause, like the first domino that "causes" the rest of the dominoes to fall.

[55] This is Stump's basic argument in "Control and Causal Determinism." See especially 38-39.

argue that this scenario disproves TNR based on the following: first, no one is or ever will be responsible for the fact the glacier is eroding before Betty acts. Second, if Betty does not act, then the mountain would still erode and crush the camp, a fact for which no one is or ever will be responsible. Despite the fact that the destruction of the camp is *overdetermined* (it would happen whether or not Betty acts), Fischer and Ravizza argue that "Betty is responsible for this consequence insofar as she has guidance control of it."[56]

Fischer and Ravizza believe that cases of simultaneous overdetermination like *Erosion* show that an agent can be responsible despite that agent's lack of access to alternative possibilities, or what they refer to as guidance control. The agent, in this case Betty, was responsible: she did what she wanted to do, was at least moderately reasons-responsive and acting on her own mechanism, and was the cause of the actual consequence in question (the destruction of the camp at T3). Furthermore, she was not forced or manipulated into doing something against her will (more on that in a moment). Here we can see the similarities with the *FSCs* that Fischer and Ravizza also find convincing against indirect arguments for incompatibilism, for just as in *FSCs* against the indirect argument, in these counterexamples, the alternative scenario (the natural cause for instance) plays no role in the actual sequence. Thus Betty is responsible because she *determines* (is the source of) what happens in the actual sequence despite her lack of alternative possibilities.

Examples such as *Erosion* ultimately are attempts to show that an agent whose action is causally determined and therefore lacks alternative possibilities can still meet the necessary grounding conditions for moral responsibility in the actual sequence.

[56] Fischer and Ravizza, *Responsibility and Control*, 156.

Erosion proceeds by showing that even in instances where the outcome is determined (though in the alternative sequence no person would be responsible) someone is responsible in the actual sequence insofar as that person exhibits guidance control. These examples are an attempt to show that *TNR* is ineffective. Having shown its ineffectiveness, they open up the possibility of causal determinism's compatibility with moral responsibility.

Fischer and Ravizza anticipate a problem with this approach. In the alternative sequence of *Erosion*—where the avalanche causes the destruction of the camp without Betty's action, what they call an "ensuring condition" —the ensuring condition guarantees that outcome without the agent actually playing any role. The problem is that *Erosion* sidesteps the issue of the compatibility of causal determinism and moral responsibility. In the *Erosion* case, the ensuring condition does not affect Betty's behavior, thus her behavior might still be indeterministic despite her lack of alternative possibilities.[57] In other words, the outcome is determined because either Betty or the natural erosion will cause the avalanche that destroys that camp, but if Betty decides on her own to cause the avalanche her decision is not necessarily determined (only the outcome is determined). In an example such as *Erosion*, either Betty's action is determined or it is not. If her action is determined, then they have failed to construct a counterexample to *TNR* because they have not shown that nonresponsibility is nontransferable. On the other hand, if Betty's action is indeterministic, then this is not an instance of transferring nonresponsibility. In the latter case, Betty is the indeterministic initiator of her own action and therefore is not simply a link in a deterministic chain of

[57] Note that if Betty's behavior is indeterministic, then the example does nothing to show that causal determinism and responsibility are compatible.

events. Fischer and Ravizza admit that if causal determinism obtains in a case like *Erosion*, then "an Ensuring Condition is always efficacious in the actual sequence…"[58] But the point of examples like erosion is to show that an ensuring condition can guarantee that the action will take place without *causing* the action. They anticipate the objection that their counterexample does not actually show that responsibility is compatible with determinism. The ensuring condition is purely counterfactual; therefore, "[w]hat is needed is a counterexample in which an Ensuring Condition does play a role in the actual sequence, and the agent is still responsible."[59] Furthermore, Fischer and Ravizza acknowledge the worry that "although Frankfurt-style examples [such as *Erosion*] show that a person can be responsible in cases where an Ensuring Condition *only would come into play in some alternative sequence*, these examples do *not* say anything about an agent's responsibility in situations where an Ensuring Condition *plays a role in the actual sequence*."[60] The incompatibilist critic's worry is that determinism involves conditions that play a role in the actual sequence, thus undermining the agent's responsibility.

Fischer and Ravizza respond to this anticipated objection with some modified counterexamples, one of which they call "Joint Assassins." In the "Joint Assassins" case, two agents, Jack and Sam, simultaneously and independently shoot and kill the mayor. Fischer and Ravizza contend that even though Jack is not, and never will be responsible for the fact that Sam pulls the trigger at *TI*, and if Sam pulls the trigger at *TI*, then the mayor is shot at *T2*, Jack nevertheless is responsible for the consequence that the

---

[58] Fischer and Ravizza, *Responsibility and Control*, 160.

[59] Ibid.

[60] Ibid.

mayor is shot at *T2*. Jack is responsible because he independently pulled the trigger and his bullet would have killed the mayor at *T2* whether or not Sam also pulled the trigger.[61] In other words, Jack is responsible despite the ensuring condition which guaranteed that the outcome would have resulted whether or not Jack acted as he did. Fischer and Ravizza argue that

> the agent is responsible in these cases because 'on his path' he acts freely and the world is 'sensitive' to his action in just the same way it would have been, had the Ensuring Path not been present. That is, on the actual path that leads to the relevant outcome, the agent evinces guidance control; and this is so despite the presence of the Ensuring Path.[62]

Fischer and Ravizza argue that these counterexamples show that an agent can be responsible despite the inevitability of the outcome so long as the agent possesses a certain kind of control and the world is 'sensitive' to the agent's action.

Despite the ingenuity of counterexamples such as *Assassins* (they offer other similar examples), Fischer and Ravizza have managed to sidestep the issue at hand, namely, whether or not moral responsibility is compatible with causal determinism. More specifically, given the truth of causal determinism and the corresponding claim that no one is responsible for the past or the laws of nature (which, if causal determinism is true, cause all future events), the result is that no one can be held morally responsible for any action whatsoever. No one is responsible for the fact that events long in the past, combined with the laws of nature, determine future events. Fischer and Ravizza's example sidesteps the issue because it does not address whether or not Betty's or Jacks' action *itself* is causally determined and thus does not show that an agent is responsible for a causally determined action. In both cases, what is determined (or at least inevitable) is

[61] Ibid., 161.

[62] Ibid., 164.

the outcome, not the fact that a specific agent is the *source* or cause of that outcome. The camp will be destroyed and the mayor will be assassinated whether or not Betty and Jack choose to carry out their plans. Furthermore, the counterexamples do not address the third option: what if Betty and Jack simply decide not to carry out their respective plans? The camp still is destroyed and the mayor assassinated, but not as a result of Betty or Jack's actions. Therefore these counterexamples appear to allow for the possibility that, despite the inevitable outcome, Betty and Jack can be held responsible inasmuch as they can choose not to act. In both cases the agent's act itself is not determined, thus each agent has access to a morally relevant alternative possibility—the possibility of refraining from action.

The objection can be reformulated in the form of a dilemma: either the world of *Erosion* and Assassins is not causally deterministic (in which case Betty and Jack may act indeterministically and *TNR* still applies because Betty and Jack are in that case not links in a causal chain but are the initiators of their own actions) or the world of *Erosion* is causally determined—which it must be in order to be used as a counterexample to *TNR*.[63] In the latter case, Fischer and Ravizza must demonstrate precisely how the causally determined decisions of the agents in those scenarios are exceptions to *TNR*.

In their counterexamples to *TNR*, Fischer and Ravizza construct examples where there are two paths to the same outcome. In one of those paths the agent is responsible for the effect. Initially, this does not seem to help their case, because it does not address

[63] *TNR* is simply an extension of van Inwagen's formulation of causal determinism. If the unalterable events of the past and the unchanging laws of nature ensure a necessary future state of affairs, then the only way that agents in such a world can be morally responsible is if nonresponsibility is nontransferable. In other words, an agent must be able to act responsibly despite the fact that her actions are directly and deterministically caused by previous events and the laws of nature over which she has no control and was not responsible for.

whether or not Betty's decision is causally determined. What they have to show is exactly why it is that they think these kinds of examples are able to save causally determined actions from nonresponsibility. As we have seen, Fischer (as well as Ravizza) believes that the essential aspect of moral responsibility is something that happens in the actual sequence, namely, that the agent exhibits a certain kind of control. More specifically, if an agent has guidance control and a mechanism that is moderately responsive to reasons, then the agent is responsible for her actions even if she has not had access to alternative possibilities. The most pressing question, however, is whether one can possess the control necessary for ascribing praise and blame in a causally determined world.

Critiques of Semicompatibilism

Two major criticisms have been raised against Fischer's distinct brand of "soft" or "semi" compatibilism. The first focuses not on Fischer explicitly but on Frankfurt-style cases in general. In *FSCs* a counterfactual intervener (such as Black) is able to ensure that Jones votes in the way Black wants without actually *causing* Jones to act. Black accomplishes this by watching for a prior sign (perhaps an external physical sign like a blush, or perhaps an internal neurological one) that will indicate that Jones is going to vote a certain way (in this case the way that Black does not want him to vote). As long as Black does not see this sign he does not manipulate Jones. But surely there is a problem here; this problem is sometimes referred to as the Kane/Widerker objection.[64] The objection roughly states that in any Frankfurt style case either causal determinism is

[64] So called because both philosophers formulated similar objections at around the same time, Kane in *The Significance of Free Will* (especially 142-44) and Widerker in "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities," *The Philosophical Review* 104 no. 2 (1995), 247-261, and "Libertarian Freedom and the Avoidability of Decisions," *Faith and Philosophy* 12 (1995), 112-118.

presupposed, or the intervener has no way of knowing how Jones will vote until he actually does vote. Causal determinism seems presupposed because some prior sign is assumed which will reveal the way that Jones is going to vote before Jones himself is aware of his decision. In other words, because of the sign Black knows Jones's decision before Jones. Surely this assumes that some prior mental state or physical neurological state *determines* how Jones will vote. Thus, indeterminists cry foul. They find that the *FSC* can work only if causal determinism (the absolute link between prior sign and action) is true. In the next chapter, I look at this objection and the response of compatibilists in more depth.[65]

Stump, Source Incompatibilism,[66] and Manipulation Arguments

Putting aside for now whether or not Frankfurt-style cases presuppose determinism,[67] other potential problems with semicompatibilism or soft compatibilism arise.[68] On Frankfurt's hierarchical account, a manipulated agent cannot be differentiated from a free and responsible agent. Since Fischer thinks that a common human intuition is

---

[65] Stump agrees with Frankfurt and Fischer that Frankfurt style cases show that alternative possibilities are not necessary for ascriptions of moral responsibility, but here it is enough to point out that this is one potential flaw of Fischer's semicompatibilism that must ultimately be corrected.

[66] Source incompatibilists argue that causal determinism directly rules out moral responsibility because the agent would no longer be the source of her actions but instead external forces would be the source of her actions. They do not, however, believe that *FSCs* are necessary for moral responsibility [is this really what you want to say?]. Therefore they are sometimes referred to as "Frankfurt-style libertarians."

[67] This issue is discussed at length in the next chapter.

[68] Soft compatibilists are compatibilists like Fischer who argue that certain robust agent-relative features *not* present in manipulations cases are necessary for moral responsibility ascriptions. Hard compatibilists—represented below by McKenna's argument—assert that moral responsibility is possible without these robust features. In other words, whereas soft incompatibilists argue that the freedom necessary for moral responsibility is compatible with causal determinism, hard incompatibilists think that moral responsibility is compatible with causal determinism but that moral responsibility does not necessarily require any robust notion of human freedom.

that praise or blame cannot be ascribed to a manipulated agent, he argues that a more refined notion of sourcehood and control can account for this distinction. Fischer thinks that his notion of guidance control answers what I argue is the basic objection to Frankfurt—that he fails to distinguish properly between manipulated agents and morally responsible agents. In the end, Fischer's notion of guidance control fails to separate the instances of non-responsibility, like phobic behavior and manipulated behavior, from more clear-cut instances of responsibility. Derk Pereboom's argument provides a way of evaluating the success of Fischer's notion of guidance control.[69] Pereboom, a source incompatibilist, objects to soft or semicompatibilism by showing that, even in Fischer's refined view, Fischer has not successfully differentiated manipulated behavior from responsible behavior. *All* behavior in a deterministic world is a result of causes that are out of the agent's control, and control is what Fischer admits is necessary for moral responsibility. Pereboom is worth quoting at length here:

> In the deterministic view, the first and second-order desires and the reasons-responsive process that result in…[the agent's action] are inevitable given their causes, and those causes are inevitable given their causes. In assessing moral responsibility…[for this action] we wind our way back along the deterministic chain of causes that results in his reasoning and desires, and we eventually reach causal factors that are beyond his (the agent's) control—causal factors that he could not have produced, altered, or prevented.[70]

Pereboom provides the basic incompatibilist objection to determinism in this passage: the cause of the agent's action is out of the agent's control. For the agent to be responsible for the action in question, that agent must possess the requisite control over that action.

---

[69] Pereboom is a hard incompatibilist who thinks that incompatibilism is necessary in order for the kind of freedom to obtain that is required for moral responsibility, but does not think that this freedom is actually possible, and thus does not think that moral responsibility—at least as normally construed—is possible.

[70] Pereboom, Derk, "*Determinism* Al Dente," *Nous 29* no. 1 (1995), 23.

Incompatibilists do not think that an agent has the requisite control if the cause of the action is outside the agent. The action must originate from the agent; in other words, the agent must be the source of the action and cannot be if the action is causally determined. Essentially, this objection is another form of the direct argument and has close (informal) parallels to *TNR*. Notably, Pereboom makes no mention of alternative possibilities. He thinks that the mere fact that an action is causally determined means that the agent cannot meet any compatibilistic requirements for moral responsibility. I return to Pereboom's critique of semicompatibilism in the last section of this chapter.

Fischer must overcome two strong intuitions to demonstrate that moral responsibility is compatible with causal determinism. The first intuition is the *Alternative Possibilities* intuition (*AP*): alternative possibilities are necessary for moral responsibility, but they are clearly not accessible to an agent in a causally determined world. Therefore it appears that moral responsibility is incompatible with causal determinism. The second intuition he must overcome is the *External Force* intuition (*EF*): if causal determinism obtains, then the agent is pushed or compelled to act as if by an external force.[71] This push or compulsion appears similar to the way that a manipulated or phobic agent is pushed, as if by an external force. For instance, an agent whose action is caused by a scientist manipulating her brain is not responsible because an external force is directly causing her actions. The same is true of an agent with a severe

[71] Fischer recognizes this objection in a number of places, perhaps most clearly in his response to David Widerker where he states that "the incompatibilist believes that the laws of nature exert a 'push' or contain a kind of 'compulsion' that is incompatible with moral responsibility. On the direct approach, the push directly rules out moral responsibility…" Needless to say, Fischer does not think that this intuition of the incompatibilist is warranted, but he thinks that direct arguments bring out important features in the dialectical debate between compatibilists and incompatibilists. See "The Direct Argument: You say Goodbye, I say Hello," unpublished essay, http://experimentalphilosophy.typepad.com/2nd_annual_online_philoso/files/john_martin_fischer.pdf accessed June 21, 2010, 8.

phobia. The source of her actions is a force external to her own will. She does not wish to act as she does and is therefore not responsible. Therefore Fischer must show that the features that distinguish a responsible agent from a manipulated agent are both compatible with causal determinism and able to distinguish sufficiently between these two kinds of agents/actions. I deal with the *AP* intuition in the next chapter when I turn to source incompatibilism. In the rest of this chapter, I focus on responses to the way that Fischer deals with the *EF* intuition.

Despite their improvement on Frankfurt's account, Fischer and Ravizza do not succeed in constructing a Frankfurt-style compatibilism that avoids the most basic problem in Frankfurt. Frankfurt provides plausible arguments that alternative possibilities are not necessary for responsibility ascriptions. Yet, Frankfurt fails to provide an account of the agential properties that differentiate free and responsible actions from actions that result from responsibility-undermining external forces. Frankfurt does not account for the *EF* intuition, an intuition that is reinforced by *TNR*. *TNR* attempts to demonstrate that responsibility is not compatible with causal determinism under any circumstances. If an agent's actions are directly caused by a combination of events in the past and the laws of nature—neither of which the agent has any control over—then the agent's actions are pushed by external forces over which she has no control. Therefore that agent is not different in any fundamental responsibility-granting way from a manipulated agent. Fischer and Ravizza think that they can show that the *EF* intuition is unwarranted with their notion of guidance control and their arguments against *TNR*.

Stump (like other source incompatibilists) argues that Fischer and Ravizza fail to formulate a successful argument against *TNR* and the *EF* intuition. Stump examines Fischer and Ravizza's freedom-relevant conditions necessary for grounding moral responsibility—guidance control and all that it entails—to see whether they can be met in a causally determined world. She argues that if those requirements can be met in a causally determined world, then Fischer and Ravizza have made no headway in defeating *TNR*. Since *TNR* argues that no agent is responsible in a causally determined world because no agent is responsible for the ultimate causes of her actions and this nonresponsibility is transferable, whatever else you can say about the state of the agent when she acts in a causally determined world, you cannot say that she was responsible for her actions if *TNR* holds. Fischer responds that if the requirements for guidance control can be met in a causally determined world, then the result of *TNR*—that causal determinism is directly incompatible with moral responsibility—is no longer incontrovertible. First, I discuss Stump's argument against guidance control. Second, I evaluate Fischer's rejoinder. Finally, I argue that he fails to demonstrate the compatibilism of causal determinism and moral responsibility.

Fischer and Ravizza's conditions for control (such as reasons-responsiveness and ownership) do not address whether or not moral responsibility is actually compatible with determinism. Rather, Fischer and Ravizza's control conditions provide an account of the agential properties that they think are present in intuitive cases of responsibility but lacking in cases of actions caused by manipulation. If Fischer and Ravizza's conditions for guidance control can be met by a manipulated agent, then the *EF* intuition returns in full force since guidance control can be manufactured by responsibility-undermining

external forces. Then, the intuition generated by TNR—that moral responsibility is unattainable in a causally deterministic world—would remain and Fischer and Ravizza's arguments against *TNR* fail. Stump pursues this strategy.

Stump imagines a science fiction-like scenario where aliens invade the planet, but do not want humans to know about them.[72] In this scenario, an individual alien takes over an individual person (whom Stump calls Sam) by invading that person's consciousness. That person "has within himself not only his own consciousness but the master's as well."[73] The alien wishes to go undetected, so he makes sure that Sam maintains his personality and continues to go about his life in more or less the way that he normally would. But the alien can introduce new thoughts and desires in Sam's consciousness, as well as suppress or eradicate certain thoughts. Thus, the master can slowly reshape Sam's character without Sam's knowledge.[74] The alien master can get Sam to act in precisely the way that he would on his own (without the master alien), yet the alien master is ultimately controlling Sam's actions. Here is an instance of a causally determined action (Sam is not the ultimate cause of his action but instead a source outside of him is the cause). Yet Sam meets Fischer and Ravizza's criteria for reasons-responsiveness because he responds the same way he would under normal conditions.

Stump adds to the story the condition that the alien master has a conversation with Sam. The alien master uses arguments about the compatibility of causal determinism and moral responsibility to convince Sam that when he "acts under the control of the alien,

---

[72] This example is based on Robert Heinlein's *The Puppetmasters*. Stump, "Control and Causal Determinism," 47ff.

[73] Stump, "Control and Causal Determinism," 47.

[74] Ibid., 48.

Sam is as much an agent and as suitable a candidate for the reactive attitudes of others as he ever was in is unaffected state."[75] If the alien master can convince Sam that he is still in control of his actions in the way necessary for moral responsibility then Sam appears to meet the first two components for moral responsibility according to Fischer and Ravizza. He sees himself as an agent and as a fair target of the reactive attitudes of others. This leaves the third component. Sam's beliefs about himself must be grounded in appropriate evidence. This component might be a bit trickier, since clearly Sam's beliefs are false. But as Stump points out, many people (including philosophers) hold false beliefs, yet those beliefs still are based on the evidence available to that person. If Sam has been convinced by plausible arguments, and comes under the sway of a certain (false) perspective, Sam's beliefs were appropriately based on the evidence available to him. Fischer and Ravizza are not committed to the epistemic claim that moral responsibility requires justified true belief. Therefore, if Fischer and Ravizza are committed to a more moderate externalist position then "Sam takes responsibility for the mechanism on which he acts when he is controlled by the alien, and so this mechanism counts as his own, on Fischer and Ravizza's account." Furthermore, this mechanism is responsive to reasons, and thus Sam's actions would count as morally responsible actions.

Initially this might not appear to be a problem for Fischer and Ravizza, because after all, they are committed to the claim that moral responsibility is (in theory) compatible with determinism, and this story is an example of this claim. However, they are equally committed to the claim that their articulation of moral responsibility can distinguish between cases of manipulation and cases where the agent acts in a free and

[75] Ibid.

responsible manner. Stump's case of Sam demonstrates that a manipulated agent can be morally responsible on Fischer and Ravizza's account. One compatibilistic response that Stump anticipates is that although it may be true that no one is even partially responsible for the laws of nature *L* and an event in the past *E* causing a certain neural state in an agent *A*, agent *A* is morally responsible for a certain mental state that is correlated with the neural state; therefore, the agent is responsible for the mental state despite the fact that the neural state was causally determined. However, this response will still not ultimately work against direct arguments for incompatibilism such as *TNR*. Although the agent may well have caused his action via his mental state, and this mental state is correlated to a neural state that is located in his brain and that he is thus responsible for, it is still undeniable that if causal determinism is true, then the cause of his neural state is outside of him. The laws of nature and the events of the past causally determine that he has this neural state at this time, and that neural state causes his mental state, and, in turn, the cause of his action. The same transfer principle (*TNR*) applies to neural states as to mental actions. If he is not responsible for his neural state then he is also not responsible for his mental state either.

Michael McKenna presents an argument that is similar in many respects to Stump's. The upshot is that moral responsibility is "incompatible with causal determinism in all one-path cases."[76] McKenna argues that it is not sufficient to show that moral responsibility *might* obtain in the unique cases of simultaneous overdetermination (cases where two different causes will independently but simultaneously produce exactly the same outcome) since in all likelihood the vast majority of cases are *one-path* cases. McKenna adapts *TNR* so that responsibility cannot

[76] Fischer, "The Transfer of Nonresponsibility," 163.

be transferred in one-path cases where the outcome is causally determined. In those cases, the actual sequence of events that give rise to the action in question can be traced back to causally sufficient deterministic causes for which no one is or ever has been even partially responsible, and thus no one is even partially responsible for the result.[77] If responsibility cannot be transferred in the vast majority of cases, then the person is morally responsible only in rare instances of preemptive or simultaneous overdetermination. McKenna's point is that Fischer and Ravizza's arguments against *TNR* result in a theory in which persons are rarely (if ever) responsible. This is certainly not an intuitive conclusion or the conclusion that they wish to draw.[78] I will deal with Fischer's response to Stump and McKenna in the next section.

Fischer's Response to McKenna and to Stump's Manipulation Argument

Fischer does not agree with Stump's, McKenna's, or any other source incompatibilists' arguments in favor of TNR and against his notion of guidance control. In his essay, "The Transfer of Nonresponsibility," Fischer contends that responses to his argument by incompatibilists like Stump, Carl Ginet, and Michael McKenna ultimately fail to provide indefeasible arguments in favor of TNR. Fischer objects to these source incompatibilists, saying that they presuppose that causal determinism directly rules out the possibility of moral responsibility without giving arguments for that claim. Fischer believes that this is a non-starter. Source incompatibilists cannot simply assume that causal determinism is directly incompatible with moral responsibility. Instead they must

[77] McKenna, Michael, "Source Incompatibilism, Ultimacy, and the Transfer of Non-Responsibility," *American Philosophical Quarterly* 38 no. 1(2001), 45.

[78] At least this is a main point of McKenna's as Fischer understands McKenna. Fischer, "The Transfer of Nonresponsibility," 163.

give an argument backing that claim. But Fischer fails to realize that these source incompatibilists are dealing with Fischer's claim that his notion of guidance control can account for the difference between actions that result from responsibility-undermining external forces such as manipulation and actions that are free in the sense necessary for moral responsibility ascriptions. They succeed in showing that if causal determinism is true, then Fischer's notion of guidance control fails to distinguish free action from manipulated action. In order to show this failure, I will examine Fischer's criticisms of McKenna and Stump in detail.

Fischer responds to McKenna by imagining a world in which McKenna's reformulation of *TNR* is in effect. In this deterministic world, most people would not be responsible for their actions, because their actions occur in the "one-path". However, in the rare instance of overdetermination (a "two-path" case), an agent *would* be responsible for her actions. This might seem odd rather than detrimental to McKenna's argument, but Fischer reminds us that the point of *TNR* is to "generate incompatibilism about causal determinism and moral responsibility…"[79] Fischer argues that if, in the world of McKenna's modified *TNR,* agents are *sometimes* responsible for their behavior (despite the truth of causal determinism), then "we do not (yet) have the incompatibility of causal determinism and moral responsibility—only the incompatibility of causal determinism with moral responsibility for *some* behavior."[80] If this holds true for McKenna's argument, then *TNR* is in peril.

---

[79] Ibid.

[80] Ibid., 164, emphasis in original.

While Fischer is right that McKenna allows that "in a deterministic world, an agent is morally responsible *only* in cases of simultaneous overdetermination,"[81] McKenna should not have conceded this point. Just before this concession, McKenna notes that despite the fact that causal determinism does not rule out overdetermination,

> if determinism is true, then the manner in which the facts of the past and the laws of nature entail one unique future is *not* analogous to the manner in which one set of independently existing causally sufficient conditions…ensure a subsequent event also ensured by *some distinct set* of independently existing causally sufficient conditions…[rather] the pertinent facts (consisting in the deterministic order of things) are not *independent* of an agent's reasons for action, *they constitute them!*[82]

This statement indicates that McKenna should not have allowed that agents can be responsible for actions in the rare instance of simultaneous overdetermination. Even in those instances, the agent's actions still are the direct result of historical facts and the laws of nature. The *Erosion* example helps illustrate this point. Not only is the *universal* consequence of Betty's action inevitable (i.e. that the camp will be destroyed at *TI*) but the *particular* consequence is inevitable as well. Betty—not the avalanche or any other event—causes the avalanche, and if Betty's world is causally determined then following McKenna it is correct to say that the deterministic order of things *constitutes* Betty's reasons for action. Fischer must directly address whether or not an agent is responsible if that agent's actions are the result of causal determinism. He fails to demonstrate that this can be the case.

Pereboom presses this point about the relationship between causal determinism and moral responsibility when he states that "in assessing moral responsibility…we wind our way back along the deterministic chain of causes that results in his [the agent's]

---

[81] McKenna, "Source Incompatibilism," 46.

[82] Ibid., 45.

reasoning and his desires, and we eventually reach causal factors that are beyond his control…"[83] One upshot of Pereboom's comment is that whether an action traces back to natural or agential causes, the bottom line is that all actions trace back to causal factors beyond the agent's control. I find that in the deterministic world that Fischer envisions, it is not the case that overdetermined actions are somehow free while others are not. In a deterministic world *both* the counterfactual ensuring condition for the universal consequence *and* the actual sequence that results in the particular consequence can be traced back to causal factors beyond the control of the agent who brings about the particular consequence. A past event (such as the big bang perhaps) eventually branches out into a number of future events, such as the avalanche that occurs from erosion, as well as Betty's setting off dynamite which simultaneously causes the avalanche. It is certainly not the case that the natural cause of the avalanche can be traced back to deterministic causes while Betty's action cannot. Thus Betty is not responsible for her action since the consequent of the special cause is inevitable.[84] The claim that she is responsible requires that she will be responsible whether or not there is an ensuring condition. The same factors should be considered even if she is not considered responsible; the presence of an ensuring condition should not bear on intuitions. I grant credence to Fischer's objection to McKenna's reformulation of *TNR* because of McKenna's unfortunate admission. But Fischer's objection misses McKenna's larger point, to which I return below.

---

[83] Pereboom, "*Determinism* Al Dente," 23.

[84] Fischer certainly agrees, and I believe that McKenna does as well, based on what I take to be the larger point of his essay.

Fischer's response to Stump reiterates a problem that he has with McKenna's modification of *TNR*. Fischer claims that both McKenna and Stump are presupposing that causal determinism rules out moral responsibility instead of giving *arguments* for precisely why this is the case. As Fischer puts the point, "[a] proponent of the incompatibility of causal determinism and moral responsibility is not entitled to help himself…to the claim that causal determinism crowds out moral responsibility."[85] If the incompatibilist did proceed in this way, then a compatibilist like Fischer has only two possible counterarguments. The first would involve agreeing with Stump and McKenna that causal determinism rules out moral responsibility. Or, second, he must simply state that causal determinism does not rule out moral responsibility and construct examples based on that supposition. The former route would amount to a concession and not an argument; the second strategy would be question-begging. Obviously, neither procedure helps to advance the argument. Fischer is certainly right to dismiss any argument that claims to show that moral responsibility is incompatible with causal determinism, yet presupposes the claim that moral responsibility is incompatible with determinism. However, this is not what Stump is doing.

Fischer thinks that two-path cases like *Erosion* show precisely what grounds moral responsibility. Fischer thinks that Betty is morally responsible "because no *uncontroversially* responsibility undermining factor impairs (or in any way affects) her deliberations, her formation of an intention, and her action in accordance with it."[86] In other words, since Betty exhibited guidance control that includes deliberation and responsiveness to reasons, she is responsible for her actions whether causal determinism

[85] Ibid., 168.

[86] Ibid., 167. My emphasis. Of course the debate

obtains or not. Since Fischer does not think that causal determination is an *uncontroversially* freedom undermining condition, he thinks that Stump and McKenna's arguments fail. He claims that they focus only on the issue of causal determinism and do not say exactly why causal determinism rules out moral responsibility. But Stump's response to Fischer does not presuppose the conclusion that causal determinism is incompatible with moral responsibility. Instead, she closely examines Fischer's requirements for morally responsible action and takes seriously the fact that Fischer is trying to differentiate morally responsible action from action that is the result of manipulation, phobia, or some similar agency-compromising external force. She concludes that if all of Fischer's requirements for morally responsible action—i.e. guidance control—can be met by an agent who is manipulated into action and is not intuitively responsible for her actions, then she has demonstrated that causal determinism is directly incompatible with moral responsibility.

I believe that taking Stump's argument one step further and bracketing out the question of determinism completely will help demonstrate the incompatibility of determinism and moral responsibility without presupposing that claim. This strategy focuses on whether or not Fischer's requirements for guidance control actually succeed in differentiating morally responsible action from action that is the result of manipulation. Stump's argument shows that Fischer does not succeed, since a manipulated agent can exhibit all the necessary freedom-conferring properties entailed by guidance control. Fischer needs to demonstrate what exactly his notion of guidance control accomplishes since this control can be produced by manipulation—what he agrees is an uncontroversial responsibility-undermining external factor. On its own, guidance control does not

adequately differentiate morally responsible action from action that is the result of manipulation.

Fischer attempts to respond to Stump's manipulation case in a slightly different way by claiming that her "Puppetmaster" case "is off the mark because it employs an overly broad notion of mechanism-individuation." Fischer and Ravizza understand a mechanism more narrowly.[87] Fischer's distinction is far from clear, but it seems to go something like this: if the mechanism referred to is practical reason or deliberation in general, then it is defined too broadly, because practical reason, in general, is not the mechanism that gives rise to a particular action. According to Davenport's interpretation of mechanism individuation, it is an "intentional explanation" that helps explain or make intelligible a particular action by referencing specific beliefs, motives, considerations, urges, etc.[88] If this understanding of mechanism individuation is right, then Fischer perhaps is claiming that in Stump's "Puppetmaster" scenario the alien master is manipulating Sam's entire reasoning faculty, and not a particular mechanism, and thus Sam is no longer a candidate for guidance control. Fischer and Ravizza's "Judith I" case attempts to demonstrate this difference between broad and narrow mechanism-individuation. A neuroscientist implants a device to stimulate Judith's brain so that Judith has an irresistible urge to punch Jane. Judith does punch Jane as a direct result of this neuroscientist's device.[89] Fischer and Ravizza claim that Judith is not responsible for her actions for two reasons. First, "the mechanism leading to the action is not

---

[87] Fischer, "Responsibility and Manipulation," *The Journal of Ethics* 8, no. 2 (2004), 153.

[88] Davenport, "Moral Sanity and Weakness of the Will," *The Journal of Ethics,* 6 no. 2 (2002), 241.

[89] Fischer and Ravizza, *Responsibility and Control,* 231.

moderately reasons-responsive…(because) Judith would strike Jane, no matter what kinds of reasons to refrain were present."[90] Fischer and Ravizza claim that since the actual mechanism that gives rise to the action in question is caused by an external source—the manipulator—and cannot be otherwise, the mechanism is not reasons-responsive. Of course, Fischer and Ravizza claim that this is not a claim about alternate possibilities, but about the agent's responsiveness to reasons in other possible worlds. Thus the problem is not that Judith lacks access to alternative possibilities, but that she would not act differently no matter what reasons where present. Second, Fischer and Ravizza state that Judith is not responsible for her action. Despite the fact that she may have taken responsibility or ownership for her faculty of practical reason, she has not taken responsibility for the particular mechanism that issues in her urge and subsequent action, namely, hitting Jane.[91]

Fischer responds to Stump by saying that "the account of manipulation only works, if it works at all, if one holds fixed the actual kind of brain manipulation, when one holds fixed the kind of mechanism that actually operates."[92] Again Fischer's comments are far from clear, but comparing Stump's "Puppetmaster" case with Fischer and Ravizza's "Judith I" case yields a better understanding of Fischer's critique of Stump. Stump's "Puppetmaster" case is unique in that it is not an instance of global manipulation, in which the agent is created and determined entirely by neuroscientists, nor is it an instance of the manipulation of a particular mechanism like "Judith I." On the one hand, it is closer to cases of global manipulation since the alien is able to control any

---

[90] Ibid., 232.

[91] Ibid.

[92] Fischer, "Responsibility and Manipulation," 153.

"intentional explanation" in Sam, instead of one particular urge (like the urge to punch a certain person). On the other hand, the "Puppetmaster" case is different from global manipulation cases because the alien master that controls Sam's actions makes sure that Sam acts in accord with his prior personality, practical reason, desires, etc. (without creating that personality). Furthermore, the alien master convinces Sam that he is responsible and should take ownership of his actions even when his actions are caused by external sources. Finally, the alien master makes sure that when "under the control of the alien, Sam does A, it is also true that if there had been reason sufficient for Sam in his uninfected state to do not-A, the alien would have brought it about that Sam in his uninfected state did not-A."[93] Sam remains reasons-responsive. Therefore, Stump argues that Sam is responsible on the Fischer-Ravizza model despite the fact that he is manipulated in a way that is intuitively responsibility undermining.

Perhaps Fischer means to say that the difference between these two cases is that Sam's entire person—his thoughts, beliefs, practical reason, urges, and all the springs of his action—is under the direct control of the alien whereas Judith retains all of those elements save for one particular urge that is manipulated by the neuroscientist. Thus Judith remains on the whole a responsible agent whose other mechanisms or intentional explanations are responsive to reasons and are her own in the Fischer-Ravizza sense of ownership. Although Sam is responsive to reasons in a way consistent with his personality, and takes ownership of his actions, all of this is directly caused by an external agent in a way similar to that in which Judith's urge to hit Jane is caused by the neuroscientist. Sam's entire agency is manipulated, whereas only one particular

[93] Stump, "Control and Causal Determinism," 48.

mechanism of Judith's is manipulated. This appears to be the critique buried in Fischer's cryptic comments.

Fischer's critique of Stump is not ultimately successful. It is true that Stump does not employ as narrow a notion of mechanism-individuation as Fischer does, but the fact remains that Sam meets all of Fischer's requirements for guidance control, yet his actions are a direct result of responsibility undermining manipulation. Fischer might counter that for an action to be an agent's own it must stem from a fixed mechanism, but in the case of Sam the mechanism changes and therefore is no longer his own. But how has the mechanism changed when Sam's actions remain largely consistent with his uninfected state? He acts in ways consistent with his pre-infected character. If he does act differently from his uninfected self he believes that these actions are his own, and that he is morally responsible. The only other difference between Sam and Judith is that Sam is *wholly* manipulated (in a way that is consistent with his character and in such a way that he still views himself as a responsible agent) while Judith has only *one aspect* of her agency manipulated (in a way that is inconsistent with her character). Moreover, Judith might deny responsibility for hitting her friend Jill. She might claim that she did not feel like she was herself when she hit Jill or that she did not really want to hit Jill or some other similar reasons. Thus one might argue that Judith is not responsible because she does not see herself as responsible (for that particular act); whereas, Sam is responsible because he sees himself as responsible.

Fischer does not respond in this manner, and for good reason. Fischer cannot admit that Sam is responsible without giving up his crucial point that guidance control can distinguish between responsible action and action that is the result of responsibility

undermining manipulation. Whether or not the mechanism is individuated in the correct way, the fact remains that Sam fulfills all the requirements for guidance control; yet, his behavior clearly results from manipulation that Fischer would surely agree is responsibility undermining.

The comparison between the Stump's "Sam" and Fischer and Ravizza's "Judith I" cases yields one further insight. As previously noted, Fischer claims that Stump is simply presupposing the incompatibility of causal determinism and moral responsibility. This comparison shows that even if we bracket out the issue of causal determinism the result is still highly problematic for Fischer's project. The only remaining reason that Fischer can plausibly invoke against Sam's moral responsibility is that his actions are not his own, but directly caused by another agent in a way similar to Judith's desire to hit her friend. The cases of Sam and Judith are both instances of non-responsible action. The only characteristic that links the two cases is that both Sam and Judith act in a way that they might not otherwise act due to causal force outside the agent. Therefore Stump's example succeeds on two levels. First, if the issue of causal determinism is bracketed out, her example succeeds in exploiting one feature of Fischer's guidance control, namely, its inability to distinguish between responsible action and action resulting from responsibility undermining manipulation. Second, this also shows that the real difference between a responsible act and a manipulated act is that the responsible act stems from the agent's own mental states, and this cannot be the case if these states are directly caused by an external agent. That is, a particular *sense* of "the agent's own" is required for responsibility.

Many philosophers would point out that the most common objection to soft or semicompatibilistic accounts of moral responsibility is that if causal determinism is true then any agent's action can be traced back to causes that lie outside the agent and thus beyond the agent's control. The intuitive reaction this generates is that if the ultimate cause is outside the agent then the agent is not responsible for that action. This reaction deals with the origination of causes. Therefore if causal determinism is true then no one is morally responsible for anything because no one has the requisite control. Fischer has yet to give an adequate response to this objection.

Pereboom's Manipulation Argument

Pereboom begins by emphasizing origination (a sourcehood claim that acts that result from responsibility-undermining external forces do not originate from the agent in the same way that a responsible agent is the originator of her actions) to stress a fundamental problem with compatibilist notions of sourcehood and control. Pereboom highlights a major problem for soft compatibilist theories like Fischer's by generalizing "from bizarre cases to those coming closer to a normally determined agent (to) suggest that…there are not relevant differences between a manipulated agent and a determined one."[94] Fischer argues that we do not ascribe moral responsibility to an agent who is being manipulated (or, more controversially, has an addiction or phobia) since a manipulated agent lacks the requisite control. The agent is not the cause (or originator) of her actions. She does not wish to perform her actions; rather, they are being caused, in a way, by the manipulator (or by the phobia). Pereboom argues that if the world is deterministic and the causes of an agent's action are outside of that agent, then that

[94] McKenna, "A Hard-line Reply to Pereboom's Four-Case Manipulation Argument," *Philosophy and Phenomenological Research* LXXVII no. 1 (2008), 145.

agent's actions are not fundamentally different from that of someone who is being manipulated. Therefore the agent does not meet the requirements for moral responsibility.

Pereboom's "manipulation argument" is an attempt to generate the intuition that no matter what compatibilist friendly agential properties are present, causal determinism *always* undermines responsibility ascriptions. The argument consists of a series of cases in which the agent is manipulated or controlled by an outside force, a force that progressively increases as the stories unfold. In Case 1 the agent, Plum, is a normal human being except that he was created by neuroscientists who can directly manipulate him. They can bring it about that he has the right mesh of desires, the right reasons, and even a character that is consistent with this reasoning, desires, and actions.[95] In other words, the neuroscientists can bring it about that the agent meets whatever requirements a particular brand of soft compatibilism demands. Pereboom realizes that a semicompatibilist might object that the neuroscientists in this scenario are unrealistically close in proximity to the agent, so he revises the scenario.

In Case 2, the agent is a normal human except that she was created by neuroscientists who cannot directly manipulate her, but have programmed her to be a rational egoist so that when she finds herself in certain circumstances she is causally determined to undertake a certain reasons-responsive process (the one that a perfectly rational egoist would) and to possess the mesh of desires that results in her particular action. This scenario is an improvement over the first because it puts some distance between the agent and her external causes. Once again, Pereboom argues that this is an

[95] *Ibid.,* p. 24. Fischer does not have much of a rejoinder here as I will demonstrate. It is interesting to note that Frankfurt does not even attempt a rejoinder. As long as we have the right mesh of desires, on his account we meet the basic freedom-relevant moral responsibility requirements.

instance of nonresponsibility because the agent's actions are ultimately determined by factors beyond her control. One could object that this is not how things actually happen (although this may not be that far off in the future and it may well be true if a certain kind of God exists). One might also find this instance problematic due to the fact that it is directly caused by an external agent.[96] Due to these potential objections, Pereboom revises the scenario once again.

In Case 3 the agent is a normal human being except that he is determined by the "rigorous training practices of his home and community to be a rational egoist."[97] This training happens so early in his life that he cannot prevent or alter it. Thus he is still causally determined to have the mesh of desires and the reasons-responsiveness that results in this action. Pereboom recognizes that a compatibilist might argue that in this instance the agent is actually responsible for his actions. Pereboom's rejoinder is simple. In order to argue that this agent is responsible when the first two were not, we must be able to identify the features present in the third scenario that are absent from the first two scenarios.

Just in case the compatibilist has a sound rejoinder, Pereboom constructs one final scenario. In Case 4, determinism (the unalterable natural laws and historical facts) is true, and the agent is a rational egoist but an otherwise ordinary human raised in typical circumstances. Despite the fact of determinism (the external cause of all the agent's actions), the agent's action in the actual sequence (to use Fischer's term) appears to stem

[96] Ibid.

[97] Ibid.

from a reasons-responsive mechanism and the right mesh of desires.[98] After all, according to Fischer and Frankfurt's arguments, despite physical determinism one agent can have a moderately reasons-responsive mechanism or the right mesh of desires and thus be responsible in a way that another agent who lacks these features is not. The agent's action in the final scenario, just as in the first three scenarios, results from a deterministic causal process that traces back to factors that are beyond her control. Pereboom wishes to press this question: what responsibility grounding element(s) are present in the final scenario that is (are) not present in the other scenarios? The one distinguishing feature of scenario four is that the external cause (physical determinism) is not another agent. Then, again, the first three cases could be reformulated "so that the determination is brought about by a spontaneously generated, mindless machine…"[99] Even if the scenarios were reformulated in this way, the intuition that the agent is not responsible in the first three scenarios would persist. If those intuitions fade due to the substitution of a mindless machine for another agent, then the question falls again to the compatibilist. Why would it matter whether or not the cause is a person? A common and persistent intuition is that if a person is directly manipulated by someone or something out of his control then she is not responsible. If this intuition is legitimate, then the final scenario contains the same fundamental problem as the first. Pereboom argues that the intuition behind all of these scenarios is that humans do not believe that an agent is responsible for the agent's actions if the agent's "action results from a deterministic

[98] Ibid., 25.

[99] Ibid.

process that traces back to factors beyond his control."[100] Unless the agent is the ultimate source of the actions in question, the agent cannot be held responsible for those actions.

McKenna's Response to Manipulation Arguments

Fischer needs help if he is going to defeat the manipulation arguments of Stump and Pereboom, and Michael McKenna provides an alternate compatibilist-type response to Pereboom's manipulation argument. McKenna suggests that Pereboom's manipulation argument (which McKenna calls a "generalization strategy") does not *decisively* demonstrate that a causally determined agent fails to fulfill the requirements for moral responsibility. He states that "Pereboom is not entitled to presume that in Case 1 Plum is not free or morally responsible."[101] McKenna bases this claim on his own generalization strategy. McKenna reverses the order of the cases and argues that "it is not evident that Plum in Case 6 (Pereboom's Case 4 where causal determinism is true) is not free and morally responsible even though he acts in a world in which causal determinism is true."[102] McKenna believes that by reversing the order of the generalization strategy we will arrive at a clash of intuitions.

Many people's stronger intuition is that Plum is not free or morally responsible in Case 1 because his mental states and actions are directly determined by another agent. Others claim that Plum is free and morally responsible in Case 4 since he has the right mental states as well as a coherent history and knowledge of acting in morally relevant ways. If these are both strong and legitimate intuitions, then depending on which case

---

[100] Ibid.

[101] McKenna, "A Hard-line Reply," 153.

[102] Ibid.

comes first, very different arguments are generated. McKenna begins with the final case where the agent is a normally functioning human being in a causally deterministic world. He argues that the agent is morally responsible in that case since he has the necessary "compatibilist-friendly agential structures (CAS)."[103] Since Case 3 (where the agent is reared in a cult-like environment designed to produce an agent with very specific beliefs and practices) does not look all that different from Case 4, and the agent still has the necessary agential structures, then she plausibly can be held morally responsible as well. By the process of generalization, McKenna concludes that the agent in Case 1 may be morally responsible as well, since he has the same necessary agential structure as the agent in Case 4.

There are three possible responses to McKenna's critique of Pereboom's manipulation cases. The first is the one that Pereboom himself makes to an earlier article of McKenna's:

> McKenna's considered view is not that we should focus solely on the agential properties, but rather that in assessing the four-case argument, one could legitimately draw greater attention to them, and that this will elicit the intuition that Plum is responsible—certainly in Case 4, but even, for example, in Case 2. At the same time, he allows that drawing greater attention to the hidden causes and their deterministic nature could occasion the intuition that Plum is not morally responsible. But given that each of these two strategies is equally legitimate, the result will be a stalemate. In response, I advocate drawing equal attention to the sorts of agential properties that typically serve as a basis for ascribing responsibility, and to the hidden causes and their deterministic nature by way of the four cases, and then let the intuitions fall where they may…I still have a strong intuition that Plum in Case 4 (and certainly in Case 2) is not morally responsible. To be sure, others may have opposing intuitions. My four case argument will not have persuasive force against them, and I have no conclusive argument to show that they are unreasonable in their reactions.[104]

---

[103] Ibid., 142.

[104] Pereboom, Derk, "Defending Hard Incompatibilism," *Midwest Studies in Philosophy 29*, 22-23 (web version). Accessed on February 2, 2010, http://www.uvm.edu/~phildept/pereboom/Replytocritics.pdf. Quoted in McKenna, 154.

On the one hand, Pereboom's cases draw attention to the source of an agent's actions. In all these cases, he concedes the potential discovery of hidden causes for the agent's action. Since Pereboom thinks that moral responsibility can only be ascribed in cases in which the source of action is under the agent's control, he concludes that perhaps no agent can be ascribed moral responsibility.[105] On the other hand, Pereboom agrees with McKenna that reversing the order and utilizing his own generalization strategy might also generate the intuitive conclusion that a determined agent can meet the requirements for moral responsibility. Pereboom concludes that we have a stalemate even though his intuitions are that the agent in all four cases is not morally responsible for his actions.

McKenna does not claim that his reverse generalization method yields a compatibilist view of moral responsibility, but he does think that the burden of proof is back on the incompatibilist. Since Pereboom offers no counter-argument, McKenna believes that he has made an important point against the incompatibilist.[106] McKenna is mistaken in this assessment, because at best all he has managed to do is put the burden of proof *back* on the compatibilist. McKenna takes his distinction between hard and soft-line responses to manipulation cases from a distinction that Robert Kane makes in *The Significance of Free Will*. Kane calls hard compatibilists (like the later Frankfurt) one who takes the 'hard' path by arguing against the commonly held intuition that globally manipulated agents are not morally responsible. 'Soft' compatibilists (like Fischer) do not think that globally manipulated agents are morally responsible, and they seek to make

[105] This is Pereboom's "hard incompatibilism." Thus his manipulation argument is only designed to show that the compatibilist-friendly agential structures for moral responsibility are not enough.

[106] McKenna, "A Hard-Line Reply to Pereboom's Four-Case Manipulation Argument," 154.

a distinction between manipulation and mere causal determinism.[107] Furthermore, soft compatibilists like Fischer only seek to show that Pereboom's generalization strategy does not work in Case 2 since the agent does not meet the necessary requirements for moral responsibility. The soft compatibilist agrees with the intuition that a globally manipulated agent is not morally responsible and therefore is not going to accept that the agent in Case 1 is responsible. Compatibilists like Fischer have spent many years beginning with "clear" cases (such as cases of global manipulation) and then attempting to distinguish a properly morally responsible agent from a manipulated one. Pereboom (like Stump) begins with this premise and proceeds to argue that whatever responsibility conferring agential properties the compatibilist ascribes to an agent, those same properties can also exist in a *globally manipulated agent*. For this reason, the debate is not furthered by saying that a globally manipulated agent can display the necessary requirements for moral responsibility when that agent has developed exactly the requirements to distinguish a manipulated agent from a responsible agent.[108] The burden of proof still rests with McKenna.

McKenna admits that he has not actually disproved manipulation arguments like Pereboom's but simply put the burden of proof back onto Pereboom the incompatibilist. However, since the history of the debate shows that the compatibilist retains the burden of proof, he only has two options. First he could try to disprove, not simply call into question, Pereboom's manipulation argument. Since McKenna admits that he cannot

---

[107] Kane, *The Significance of Free Will,* 67-68. Cited by McKenna, "A Hard-Line Reply," ftn 2, 143.

[108] For this reason it makes sense that soft compatibilists like Fischer attack Case 2 and not Case 1. Fischer agrees with the incompatibilist that the agent in Case 1 is not responsible since he believes that this is a clear case of responsibility-undermining manipulation.

disprove Pereboom's argument, I see only one remaining option for him. He must, following Kane's distinction, take the true "hard" compatibilist route and show that a globally manipulated agent can meet the requirements for moral responsibility. McKenna hints at this route at the end of his article when he quotes an important comment of Frankfurt's:

> A manipulator may succeed, through his interventions, in providing a person not merely with particular feelings and thoughts but with a new character. That person is then morally responsible for the choices and the conduct to which having this character leads. We are inevitably fashioned and sustained, after all, by circumstances over which we have no control. The causes to which we are subject may also change us radically, without thereby bringing it about that we are not morally responsible agents. It is irrelevant whether those causes are operating by virtue of the natural forces that shape our environment or whether they operate through the deliberate manipulative designs of other human agents.[109]

McKenna returns full circle to the argument with which he began. The only way to reject Pereboom's manipulation argument conclusively (and truly put the burden of truth back on the incompatibilist) is to claim that globally manipulated agents can fulfill all the necessary requirements for moral responsibility. McKenna realizes this, but also seems to realize that demonstrating this is a very difficult task. He cites the example of people making radical changes in life for reasons that seemed to them beyond their control, as well as moral luck cases of Nagel and Williams, but these do not amount to arguments of any substance.[110] Therefore, in the end, McKenna's arguments demonstrate that Pereboom's manipulation arguments actually succeed. They put the burden of proof

---

[109] Frankfurt, "Reply to John Martin Fischer," in *Contours of Agency*, 27-28.

[110] I find the examples of people undergoing radical changes due to reasons apparently beyond their control quite fascinating, but do not think that these examples make any headway at all towards showing that causal determinism is compatible with the requirements for moral responsibility much less that globally manipulated agents are morally responsible. That a person feels compelled to act in a certain way by reasons that were not under her direct control does not imply that she is being manipulated or even that she is determined. Ferreira's notion of a "critical threshold" discussed in chapter four is an interesting way to deal with cases like these in a libertarian context.

back on the compatibilist. Whatever requirements a (soft) compatibilist deems sufficient for conferring moral responsibility are attainable by a globally manipulated agent; therefore these requirements are not sufficient for *distinguishing* moral responsibility from manipulation. Finally, a hard compatibilist route encounters a fundamental difficulty. Since the kind of freedom that a hard compatibilist claims is compatible with determinism is the same kind of freedom that a globally manipulated agent possesses, this is no longer "freedom" in any intuitive sense of the word, but in a different way entirely. Therefore, if McKenna chooses to take this hard compatibilist route, then he will no longer be a *compatibilist*. He will not be arguing that *freedom* is compatible with determinism, but that the *necessary responsibility-grounding agential structures* are compatible with determinism. So McKenna (or any compatibilist) can at best claim that whatever agential properties are required for moral responsibility are displayed by the manipulated agent, but this is a very different and difficult argument to make. Regardless, McKenna's argument does not assist compatibilists like Fischer or the early Frankfurt. In the end, manipulation arguments such as those of Pereboom and Stump retain a leg up in the debate because Fischer and his cohorts have failed to demonstrate that moral responsibility is compatible with determinism.

Fischer's Final Challenge

Fischer ends his earlier book *The Metaphysics of Free Will* by stating:

> Even if there is just one available path into the future, I may be held accountable for *how I walk down this path.* I can be blamed for taking the path of cruelty, negligence, or cowardice. And I can be praised for walking with sensitivity, attentiveness, and courage. Even if I somehow discover that there is but one path into the future, I would still care deeply how I walk down this path. I would

> aspire to walk with grace and dignity. I would want to have a sense of humor. Most of all, I would want to do it my way.[111]

He recently states that "I can walk down a path where, unknown to me, there is a counterfactual intervener whose presence ensures that I do not have genuine, robust alternative responsibilities."[112] We can gather from these comments (and from compatibilism generally) that Fischer thinks that a person can be held morally accountable for her actions even if there were a *real* (not merely counterfactual) intervener present for the entirety of that person's life. After all, if causal determinism obtains, then no person ever has or ever will have a real live option; her own character has been formed by factors ultimately out of her control. But humans certainly *think* that they do have options; and, they live lives accordingly (this is after all part of the motivation for Fischer's notion of guidance control), and hold people accountable as if they really did have more than one option available to them.

Fischer admits here that an entire life can be lived authentically, creatively, and responsibly despite the fact that all of the "choices" that the agent had may not have been choices. Ironically, this claim meshes well with Frankfurt's response to Fischer's unique style of compatibilism. Fischer thinks that a person's history is one of the relevant conditions necessary to ground moral responsibility. Fischer is building on a common objection to Frankfurt, namely, that higher-order desires do not belong to the agent in a responsibility-granting way. Fischer argues that instead of a hierarchical model of the will, what is necessary to claim that certain desires belong to the agent is a mechanism

---

[111] Fischer, *The Metaphysics of Free Will: An Essay on Control*, Aristotelian Society series, v. 14 (Cambridge, Mass: Blackwell, 1994), 216. Cited in Fischer, "A Framework for Moral Responsibility," p. 27.

[112] Fischer, "A Framework for Moral Responsibility," in *My Way,* 28.

that is both moderately reasons-responsive and historically constituted. In other words, the agent must have taken responsibility in the past for actions that stemmed from that mechanism. But as Stump indicates, both of these elements can exist in a person whose character and choices are wholly determined. Fischer thinks that this simply shows a stalemate in the debate. Incompatibilists like Stump think that we cannot take responsibility for a self that is wholly determined because we cannot be the source or originators or our action. Fischer, by contrast, thinks that determinism does not necessarily undermine any of the characteristics—such as creativity, intentionality, and origination—that libertarians and compatibilists alike value. Fischer points out two different notions of origination at work here; source compatibilism and source incompatibilism. The former is compatible with determinism while the latter is not. He pushes back at the libertarian by asking why libertarians think that only their brand of origination—one that is not compatible with determinism—is sufficient to ground moral responsibility. At this point, I have sought only to show why Fischer's notion of origination is insufficient: it fails to distinguish properly acts that result from responsibility-undermining external forces such as manipulation from acts where the agent is intuitively morally responsible because that agent has the requisite control. In the final chapter I will flesh out a Kierkegaard-style libertarian notion of origination and show how it can ground moral responsibility in the specific ways that Fischer's notion of origination cannot.

Conclusion

Fischer's semicompatibilism is a definite improvement on Frankfurt's largely unsystematic understanding of freedom and moral responsibility. Fischer claims that

moral responsibility—and not the freedom to do otherwise—is compatible with determinism. This point clarifies Frankfurt's project. Furthermore, his additions of reasons-responsiveness and an historical element mitigate some of the major problems with Frankfurt's hierarchical volitionalism. However, Fischer's more systematic approach reveals deep problems for any soft compatibilist theory of moral responsibility, including Frankfurt's. Fischer provides the most compelling and comprehensive compatibilist theory of moral responsibility, yet he has not accomplished what he intended. He has not differentiated cases of responsibility-undermining external forces such as manipulation from cases of responsible agency. Thus Fischer does not provide Frankfurt with any additional arguments for the compatibility of determinism and moral responsibility. All is not lost, however, for Fischer or Frankfurt. They both have important insights into the freedom-relevant conditions for moral responsibility. Frankfurt's initial insight that an agent's reflective endorsement of her desires is fundamental to the free and responsible action of a person remains largely unchallenged by Fischer. Fischer simply raises objections against the hierarchical view and offers his own historicist view as an alternative. There is no reason to abandon this hierarchical view if these objections are answerable from within that view, which is exactly what Stump seeks to do. Therefore I will now turn to Stump's Frankfurt-style libertarianism.

CHAPTER FOUR

Stump's "Source Incompatibilism"

Introduction

Stump and Pereboom's manipulation arguments show sufficiently that Fischer fails to provide a convincing argument for the compatibility of causal determinism and moral responsibility. However, Stump and Pereboom agree with Fischer that Frankfurt's chief insight remains valid: alternative possibilities are not a relevant explanatory condition for moral responsibility. Stump is another philosopher who agrees with certain aspects of Frankfurt's theory of freedom and moral responsibility while disagreeing with other aspects. Like Fischer, she agrees that Frankfurt's counterexamples against the principle of alternate possibilities succeed in showing that moral responsibility does not require alternative possibilities. She also agrees with Fischer that Frankfurt's hierarchical or structuralist model is inadequate for grounding moral responsibility because it lacks a reason or knowledge component.[1] Unlike Fischer, she does not reject the structuralist model in favor or an historicist model; instead, she modifies Frankfurt's structuralist model by adding a reasons or knowledge component, also referred to as an intellect component. This component utilizes insights from Aquinas to answer concerns that second-order desires are not responsibility-conferring. She argues that this knowledge component makes a higher-order desire an agent's own and therefore meets the freedom-

[1] As I pointed out in chapter one, not only does Frankfurt's account lack a reasons component, he explicitly rejects that any such component is necessary for moral responsibility. Fischer of course provides a reasons component but it is difficult to see how this component helps in a deterministic context.

relevant conditions for moral responsibility. Finally, Stump disagrees with both Fischer and Frankfurt that the freedom necessary for moral responsibility is compatible with determinism. Thus she is a Frankfurt-style incompatibilist or what she at one point refers to as a "modified libertarian."[2]

Stump argues against the intuitions of leeway incompatibilists like van Inwagen who believe that

> moral responsibility requires the ability to do otherwise. Since, however, I (Stump) agree with van Inwagen that moral responsibility does entail indeterminism, I will also argue that an agent who acts indeterministically need not have alternative possibilities for acting open to her. I share the second of van Inwagen's intuitions, that an agent's being causally determined is incompatible with her having the ability to act otherwise; but I will argue that an agent's not being causally determined, her acting indeterministically, isn't sufficient for her to have that ability, and so it isn't necessary that an agent who acts indeterministically have alternative possibilities for her action.[3]

Stump is an incompatibilist who thinks that the freedom-relevant conditions for attributions of moral responsibility are *incompatible* with determinism. She also thinks that causal determinism is false, thus allowing that the necessary freedom-relevant conditions are met in cases of morally responsible actions. Her modified libertarianism is a version of "source incompatibilism," an increasingly popular position among incompatibilists.[4] Source incompatibilists stress that the *primary* explanatory condition for moral responsibility is the causal history of the action, which must be

---

[2] Stump, "Augustine on Free Will", in *The Cambridge Companion to Augustine*, ed. Norman Kretzmann and Eleonore Stump, (Cambridge: Cambridge University Press, 2001), 125.

[3] Stump, "Alternative Possibilities and Moral Responsibility: The Flicker of Freedom." *The Journal of Ethics*, 3, no. 4, (1999), 299.

[4] The term originates from Michael McKenna, "Source Incompatibilism, Ultimacy, and Transfer NR." Timpe argues that Stump is a prime example of a source incompatibilist in his "Source Incompatibilism and its Alternatives," *American Philosophical Quarterly* 44.2 (2007): 143-155.

indeterministic.[5] Source incompatibilists argue that the causal history of the action must be indeterministic because responsibility ascriptions require that the agent be the initiator or source of her action. If something external to the agent is the source of the action in question (which would be the cause if causal determinism obtains), then the agent would not be the source and would not be responsible for that action. Furthermore, if no external force is the primary cause of the agent's action, then that action is indeterministic.[6]

Source incompatibilists are distinguishable from more traditional "leeway incompatibilists." Leeway incompatibilists agree that the causal history of the action is important because it must be indeterministic. However, they stress that the causal history is indeterministic precisely because an "indeteterministic history is required to secure alternative possibilities."[7] In other words, source incompatibilists may allow that alternative possibilities are sometimes available to agents, but are not necessary for moral responsibility; leeway incompatibilists argue that alternative possibilities are either the main *grounding condition* for moral responsibility or a necessary part of the *grounding mix* of freedom-relevant conditions. In sum, both source and leeway incompatibilists hold that "an agent acts with free will, or is morally responsible for an act, only if the act

---

[5] Pereboom, "Source Incompatibilism and Alternative Possibilities," in *Moral Responsibility and Alternative Possibilities*, ed. David Widerker and Michael McKenna, 197.

[6] The assumption here is that at least some of a normally functioning agent's actions are not determined *internally* either. In other words, this is a way of distinguishing free and responsible action from actions that are the result of addiction, compulsion, manipulation, etc. Thus source incompatibilism is amenable to the claim that not all actions that flow from the agent as source are free and responsible actions.

[7] Pereboom, Derk, "Source Incompatibilism and Alternative Possibilities," 186. Pereboom coined the term "leeway incompatibilists."

is not *ultimately* causally determined by anything outside of the agent;"[8] but, only leeway incompatibilists hold that alternative possibilities are a *necessary* freedom-relevant condition for moral responsibility ascriptions.

This chapter proceeds in three movements. In the first section, I examine the Thomistic roots of Stump's source compatibilism. In the second section, I examine her arguments against the principle of alternative possibilities. In the final section of the chapter I criticize Stump's source compatibilism. I decisively demonstrate that the claim that indeterminism is a necessary grounding condition for moral responsibility is inseparable from the claim that alternative possibilities are also a necessary grounding condition for moral responsibility.[9] Furthermore, I find no reason to discard Frankfurt's structuralist approach to moral responsibility when augmented properly. This examination of Stump's particular brand of source incompatibilism reveals that the freedom relevant conditions necessary for grounding moral responsibility include a hierarchical notion of freedom supplemented with a reasons or knowledge component. To satisfy the freedom relevant conditions for moral responsibility this hierarchical account must be incompatible with determinism because the agent (and not some external force) must be the source of her actions and must have access to alternative possibilities.

---

[8] Stump, "Augustine on Free Will," 125. Quoted in Timpe, "A Critique of Frankfurt-Libertarianism," *Philosophia 34* (2006), 192.

[9] I am in a way agreeing that we should "say goodbye to the direct argument" because I agree that ultimately APs are part of the grounding mix of moral responsibility. However, I still think that Fischer is correct to point out that the sourcehood and control components are in theory separable even for a leeway or wide source incompatibilist. The reason that the direct argument is still valuable for the leeway incompatibilist is that whereas the indirect argument states that causal determinism indicates a blockage of alternative paths that limits the agent's ability to guide herself along her chosen path, the direct argument states that causal determinism results in a push or compulsion that also limits the agent's ability to guide herself along her chosen path. So though I agree that in the end these two arguments are intertwined, they can be pursued separately in a dialectically useful way.

Stump's Thomistic Theory of Free Will and Moral Responsibility

Stump remarks that Frankfurt's hierarchical theory of freedom is new in its formulation yet the concept itself is quite ancient. "It is a commonplace of medieval philosophy that the higher faculties of human beings are characterized by reflexivity."[10] For instance, she claims that Aquinas holds that "all the higher powers of the soul…are able to act on themselves."[11] Stump's understanding of free will draws partially on Augustine and especially on Aquinas. She distills her interpretation of these thinkers over several articles.[12] Stump contends that both Augustine and Aquinas are difficult to interpret because contemporary scholars assume that they have to fit into one of our modern conceptions of free will and moral responsibility, primarily compatibilism or incompatibilistic libertarianism. However, she argues that a close examination of Augustine's and Aquinas's views of freedom and responsibility yields a view that does not fit cleanly into either the common understanding of compatibilism or of libertarianism.

---

[10] Stump, "Augustine of Free Will," 126. Reflexivity is a crucial part of Frankfurt's hierarchical description of the will. What defines a person, according to Frankfurt, is the ability to engage in self-directed activity, where activity is primarily mental actions. So on Frankfurt's account, the various mental faculties (or intentional mental states) —such as the will—can act on themselves and thus reflexively.

[11] Stump, "Freedom: Intellect, Action, and Will," *Aquinas*, 5 (online version). Accessed online at http://sites.google.com/site/stumpep/onlinepapers.

[12] See especially "Freedom: Intellect, Action, and Will," in *Aquinas* (London: Routledge, 2003), http://sites.google.com/site/stumpep/onlinepapers; "Sanctification, Hardening of the Heart, and Frankfurt's Concept of Free Will". *The Journal of Philosophy.* 85, no. 8: 395-420; "Augustine on Free Will," *The Cambridge Companion to Augustine*, (Cambridge, UK: Cambridge University Press, 2001), 124-147; "Persons: Identification and Freedom," *Philosophical Topics* 24 no.2 (1996): 183-214. Note that I am not interested in debating Stump's interpretation of Aquinas or Augustine. The only way to explain fully Stump's view is by understanding her interpretation of these thinkers. One might argue that her interpretation is wrong, but it is of no matter here because these interpretations still represent her view and that is all that I am ultimately interested in.

Stump argues that one reason why Aquinas's view is so difficult for moderns to categorize is that the contemporary free will discussions assume that "human freedom ultimately is or depends on a property of just one component of human mental faculties: the will, and freedom most fundamentally consists in the will's ability to act autonomously in general and independently of the intellect in particular."[13] Aquinas, however, maintains that free choices stem from the proper interaction of two mental faculties, the *intellect* and the *will*. In other words, Aquinas builds an intellectual component into his account of human freedom.

First, Aquinas thinks that the will is not neutral but inclines or bends towards goodness. The will is like an appetite, a hunger for goodness. The will is primarily a *volitional* capacity and not a *rational* capacity. Therefore Aquinas holds that the will does not apprehend or make judgments concerning goodness because that is the intellect's function.[14] However, the will does naturally desire goodness. Thus Aquinas calls the will a "rational appetite." On this oversimplified picture, it appears that no one would ever choose anything bad. The intellect perceives some act *X* as good, then the will, which is a natural appetite for goodness, chooses to perform act *X*. Furthermore, Aquinas sees the intellect as the final cause of human action since "apprehending or judging things as good is the business of the intellect."[15] Despite the fact that the will is an appetite for goodness and that the intellect is the final cause, the will still retains a certain amount of power over human action. "…[T]he will exercises some degree of

---

[13] Stump, "Freedom: Intellect, Action, and Will," 1.

[14] Ibid., 2.

[15] Ibid.

efficient causality over the intellect. In some circumstances, it can command the intellect directly to adopt or reject a particular belief. It can also move the intellect by directing it to attend to some things and to neglect others, or even to stop thinking about something altogether."[16] So, the will can command itself in two ways: indirectly by commanding the intellect to stop thinking about one thing and perhaps focus on something else instead; and directly by willing to will something.[17]

In Aquinas's conception, the will has limited *efficient* causality (limited because the will cannot always redirect the focus of the intellect); but the intellect has *final* causality. Despite the will's (limited) ability to redirect the focus of the intellect, the will can never will something other than what the intellect presents under *some description as good*.[18] Thus there is an *intellectual* component to free action; in fact, for Aquinas this intellectual component is primary. An act is not free unless the agent *cognizes* the act as good under some description. An agent can never will something without *perceiving* that thing to be good. Of course, a person can talk herself into thinking something is good and later consider the action not good; but, Aquinas's point is that humans do *perceive* that the action is good or they would not perform that action.[19] Stump argues that

---

[16] Ibid., 3.

[17] I.E. the will can want to want what it wants.

[18] Ibid., 5. The qualification of "some description" is important because it indicates why agents don't always will the good. The intellect can give a description of an action or end as good that really isn't good so that will wills it even though it really isn't good. This will become important in what follows.

[19] A simple example can illuminate this idea of perceiving an action as good. Bill is on a diet and knows that too much sugar can give him a headache, yet he also really loves bread pudding. When the server brings the dessert cart by Bill's table, upon seeing the bread pudding Bill desires it because it looks good to him; the initial impression that Bill has of the cake is that it is desirable. Despite the fact that Bill knows that he is on a diet and that the bread pudding is not a good idea, his may choose to act in accordance with his initial perception of the bread pudding. According to Aquinas he could not make this decision if he did not perceive the pudding as good in some sense.

Aquinas holds that "[b]ecause God has created the will as a hunger for the good, the will by nature desires the good. And whatever is good to such a degree and in such a way that a person cannot help but see it as good, the will of that person wills by natural necessity. One's own happiness is of this sort…"[20] Thus if something is clearly and overwhelmingly perceived as good by the agent, the agent cannot help but will this thing. Two points are important here; first, that this intellectual component is quite strong for Aquinas, so strong in fact, that willing does not happen without the agent believing that the object of her will is good under some description. Second, the will does not always have a choice concerning what it is able to will. If the intellect presents something as *overwhelmingly* good to the will, the will *must* will that thing. This suggests a compatibilistic interpretation of Aquinas. He appears to reject a leeway condition: Aquinas thinks that human action can be free as long as the will wills what the intellect presents; alternative possibilities are not necessary for free action. Stump thinks this is a hasty conclusion, however, and proceeds to show that Aquinas is a kind of libertarian despite his insistence that free action does not require alternative possibilities.

Stump develops three main ideas from her understanding of Aquinas. First, she argues that free human action has an intellectual component. Free action must stem from the proper interaction of the intellect and will. The agent must *perceive* the action as good under some description in order to *will* that action. Second, Stump believes that the will is compelled to will something only if it is presented to the will by the intellect as overwhelmingly and incontrovertibly good (thus eliminating the *AP* condition). This second point ties to the third point, which is crucial. Based on Aquinas's account, Stump argues that the source of the action must be inside the agent for that action to be free. She

[20] Ibid.

quotes Aquinas as saying that "an act of the will is nothing other than an inclination which proceeds from an interior cognizing principle…but what is compelled or violent is from an extrinsic principle."[21] Aquinas's meaning is not immediately clear. Stump glosses Aquinas, saying:

> If something extrinsic to the agent were to act on the will with efficient causation, then the tie of the will to the intellect, from which acts of the will get their voluntary character, would be broken, and so the act of the will wouldn't be voluntary—or to put it more nearly as Aquinas seems to think of it, in such a case it wouldn't be a real act of the will at all.[22]

Aquinas believes that in order for an agent to act freely and responsibly, the action must proceed from the proper interaction of that agent's intellect and will. If either one of these faculties or the interaction between the faculties is compromised, then the agent's action is no longer free. For instance, if someone (or thing) forces or manipulates the agent to act in a way that the agent would otherwise not act, then the agent is not acting freely or responsibly. Furthermore, if the agent's intellect is so severely compromised that the agent's will no longer will's what that agent's intellect perceives as good under some description, then the agent is no longer free and responsible. For instance, Aquinas does not believe that Satan can directly cause an agent to sin.[23] Apparently he believes that the only way for Satan to directly cause another agent to sin would be either for

---

[21] Aquinas *ST* IaIIae.6.4. Quoted in Stump "Freedom: Intellect, Action, and Will," 10.

[22] Stump "Freedom: Intellect, Action, and Will," p. 10. By efficient causation both Aristotle and Aquinas mean what we would simply call causes. For instance, Aristotle says that "the source of the primary principle of change or stability is a cause" and this is what we usually call the efficient cause. Thus in response to the question "Whether the intellect is moved by an extrinsic principle," Aquinas quotes Aristotle saying that "the appetible object is a mover not moved, whereas the will is a mover moved." (*Physics* II.3, 194b30). In other words, because the will can only will what is presented to it as good under some description by the intellect, the intellect's perception of the object is the final mover (Aristotle's "mover not moved"), while the will as a mover moved by the intellect only has efficient causal power. In order for the agent to be responsible, both of these elements must be present.

[23] See Aquinas, *On Evil*, trans. by Richard J. Regan, ed. by Brian Davies, (Oxford: Oxford University Press, 2003), 151.

Satan to take over the agent's intellectual faculty (as in the case of demon possession) or force the agent to act in the same way that a thief with a gun forces his victim to act.[24] In the former case it is no longer the *agent's* perception of the object that is the final mover but an external agent, while is the latter it is not the agent's will strictly speaking that is the efficient cause of the action but the will of the robber.[25] Thus agent's are not free and responsible in instances where an an extrinsic principle either forces the agent to act or severs the proper tie between intellect and will.

Stump argues that this final point is very important. Aquinas thinks that the voluntary action must result from the proper interaction of the agent's intellect and will, and that if an extrinsic principle (in other words not the agent's perception of the object) acts on the agent's will with efficient causation then this tie is broken. In other words, the causal sequence must initiate in a special way from within the agent, and therefore Stump concludes that Aquinas is an incompatibilist, specifically a source incompatibilist. Pereboom's critique of soft compatibilism locates a similar issue and helps make sense of Stump's assessment of Aquinas:

> In the deterministic view… [the causes of the agent's action] are inevitable given their causes, and those causes are inevitable given their causes. In assessing moral responsibility…[for this action] we wind our way back along the deterministic chain of causes that results in his reasoning and desires, and we eventually reach causal factors that are beyond his (the agent's) control—causal factors that he could not have produced, altered, or prevented.[26]

---

[24] Unless of course the thief with the gun simply gives the agent the excuse to do what he already wanted to do, but in all other (usual) cases this compulsion undermines the agent's freedom and responsibility.

[25] Of course strictly speaking the agent does will the action that he performs, but this is not what he wants to do all things considered, and he might consider his action morally abhorrent and not normally engage in that activity. Even Frankfurt considers such an coerced action unfree and thus not a responsible action.

[26] Pereboom, Derk, "*Determinism* Al Dente," 23.

According to Pereboom, if causal determinism is true, then the ultimate causes of an agent's actions stem from extrinsic principles and not intrinsic principles. In a deterministic world the agent does not control how her actions come about or what actions come about. Those actions result from causes ultimately outside the agent's control. Although Aquinas does not address the problem in the same way and does not refer to causal determinism, he has the same basic objection as Pereboom. So according to Stump, Aquinas—like Pereboom—is a source incompatibilist.

Source compatibilism can be understood in comparison to traditional or common libertarianism. It includes two main claims:

> (L1) an agent acts with free will, or is morally responsible for an act, only if the act is not causally determined by anything outside the agent;
> and
> (L2) an agent acts with free will, or is morally responsible for an act, only if he could have done otherwise.[27]

These two claims comprise the common libertarian (sometimes referred to as "leeway") position. L2 provides the strong control component of alternative possibility (or avoidability). And L1 provides the strong sourcehood component[28] about the causal history of the action.[29] However, Stump is among a growing number of incompatibilists (some of whom are libertarians) who accept the strong sourcehood component (L1) while rejecting the strong control component (L2). They are incompatibilists because despite

[27] Stump, "Augustine on Free Will," 125. Stump uses this formulation in a number of places and credits Van Inwagen and Widerker among others.

[28] I refer to this as a strong sourcehood component because soft compatibilists like Fischer stress a source component (the mechanism must be the agent's own in such a way that it cannot be the manipulated agent's own) but this source component is weak because it is still compatible with determinism.

[29] Note that like many other Frankfurtians, she does not make the stronger claim that alternative possibilities are logically impossible, but instead the weaker claim that they are not necessary for moral responsibility. Therefore even if an agent sometimes has genuine metaphysical access to alternative possibilities, this is not a necessary freedom-relevant condition for ascribing moral responsibility.

the fact that they reject L2, they still think that L1 is true and that it is incompatible with causal determinism. The strong sourcehood component (L1) is actually a two-part claim, one part about the causal history of the action and the other about the origination or source of the action. Pereboom is a source incompatibilist who argues that an agent is not morally responsible for an action if that action traces deterministically back to "causal factors beyond the agent's control." The core of this claim about causal history is an intuition about origination. He argues that "[i]f an agent is morally responsible for her deciding to perform an action, then the production of this decision must be something over which the agent has control, and an agent is not morally responsible for the decision if it is produced by a source over which the agent has no control."[30] Pereboom, like Stump and Fischer, is expressing the claim that for an agent to be responsible for her action she must have the requisite control. Unlike Fischer however, Pereboom and Stump think that the control necessary for ascriptions of moral responsibility is incompatible with causal determinism.

Since L1 is primarily a claim about the casual history of an act, and causal history is related to but does not necessarily include the notion of control, it is necessary to add a third claim about free will and moral responsibility with which many (but perhaps not all) libertarians would agree. This claim could be formulated in a number of ways, but Stump, again following Aquinas, formulates the third claim this way:

> (L3) an agent acts with free will, or is morally responsible for an act, only if her own intellect and will are the sole ultimate or first cause of her act.[31]

---

[30] Pereboom, "Source Incompatibilism and Alternative Possibilities," *Moral Responsibility and Alternative Possibilities*, edited by David Widerker and Michael McKenna, 186. The very language of control that Pereboom employs here is problematic. If sourcehood is really about control then it is difficult to see how an agent can be the source of her actions in the requisite sense without the assertion that alternative possibilities are at least a part of the grounding mix of moral responsibility ascriptions.

[31] Stump, "Augustine on Free Will," 126.

Stump claims that Aquinas is a source incompatibilist because he thinks that an action only satisfies the freedom-relevant conditions for moral responsibility when the action cannot be produced by an extrinsic principle (a claim focusing on causal history—L1). That condition alone is not sufficient for moral responsibility. The cause of an action could be internal to the agent without the agent having the control necessary for moral responsibility (such as in the case of a phobia where the agent cannot control her action due to the presence of the phobia). Therefore Stump emphasizes that Aquinas also requires that the agent's own intellect and will be the sole ultimate cause of the action.

Stump's Thomistic source incompatibilism requires more than the addition of L3—which requires that morally responsible action include that the source of the action is located in the agent's intellect and will—because she claims that her Thomistic libertarian position differs from the traditional libertarian positions such as van Inwagen's.[32] Her brand of libertarianism, derived from Aquinas, accepts only L1 and L3 but rejects L2. She claims that Aquinas is an incompatibilist who rejects alternative possibilities as a freedom-relevant condition necessary for grounding moral responsibility. She summarizes Aquinas' view as follows:

> For Aquinas, human freedom is vested [in] human cognitive capacities and in the connection of the will to those capacities. As long as human acts originate in those faculties, those acts count as free, even if the agent couldn't have done otherwise in the circumstances or the act of the will is necessitated by natural inclinations of the intellect and the will. On Aquinas's account, the causal chain culminating in a free mental or bodily act cannot originate in a cause extrinsic to the agent just because it must have its ultimate source in the proper functioning of the agent's own intellect and will.

[32] Many libertarians accept all three propositions. Timpe refers to these libertarians as "Wide Source Incompatibilists."

> What is sufficient for libertarian free will, then, on Aquinas' account, is that the ultimate source of an action be the agent's own will and cognitive faculties.[33]

She points to passages where Aquinas argues that an agent can will what she wills necessarily. For instance, "when what is willed is so altogether good that the intellect can't find any description under which to present it as not good—as in the case of happiness"[34] the agent wills necessarily. On the surface, this might appear to conflict with the earlier point that an extrinsic principle cannot act on an agent's will with efficient causation because the agent's intellect and will must be the source of voluntary acts. Recognizing this problem, Aquinas distinguishes between two different classes of necessary willings: 1) instances where the agent wills something necessarily due to coercion, and 2) instances where the agent wills something necessarily due to natural inclination, like the natural inclination that he thinks all humans have towards their own happiness. The necessity of coercion is precisely the kind of causation that Aquinas claims is incompatible with voluntary action, such as the earlier example of a manipulator directly causing an agent to will something without the consent of that agent's intellect. In the latter case, the necessity of natural inclination, willing is still voluntary because the volition results from the proper interaction of the agent's intellect and will. Aquinas cites the example of the redeemed in heaven who can no longer will evil because they have such a clear intellectual view of the good that "their intellects can no longer find descriptions under which to present as good things that really are evil."[35] Clearly,

---

[33] Stump "Freedom: Intellect, Action, and Will," 32-3.

[34] Ibid., 25.

[35] Ibid., 26.

Aquinas does not think that the requirements for moral responsibility include having access to alternative possibilities at the time of willing. Stump concludes that alternative possibilities are not necessary for moral responsibility and therefore *PAP* is false and L2 is unnecessary.[36]

Stump's Frankfurt-Style Libertarianism

On the surface, Stump's Thomistic understanding of free will may seem not to have a lot in common with Frankfurt's understanding. Aquinas is an incompatibilist who stresses that the right connection between intellect and will is required for free action. Conversely, Frankfurt is a compatibilist who clearly downplays any intellectual component in morally responsible action. However, Stump (following her interpretation of Aquinas) agrees with Frankfurt that alternative possibilities are not a necessary grounding condition for moral responsibility. Furthermore, and once again following Aquinas, she thinks that the mental powers of a human being are able to act on themselves, i.e., the intellect can cognize itself and the will can will to will. She builds on both fundamental aspects of Frankfurt's account. However, Stump thinks that

[36] I think that this conclusion is hasty on Stump's part. That Aquinas does not require that *every* action have alternative possibilities in order to be a morally responsible action does not imply that he thinks that alternative responsibilities are not necessary at all for moral responsibility. One might argue that the redeemed can no longer will anything other than the good because their character has been formed over time, through their own free actions in which they did have access to alternative possibilities; and, that freely formed character, for which they are responsible, dictates that they will only the good, which is after all what they want. This "tracing" notion of moral responsibility—where in order for an action to meet the freedom-relevant conditions for moral responsibility the agent must either have access to alternative possibilities at that moment *or* at some moment(s) is the past that was character forming and dictated the present action—is similar to Kane's position as well as Kierkegaard's; and, based on Stump's exegesis of Aquinas, I see no reason why it might not apply to Aquinas as well. However, her interpretation of Aquinas is plausible and I am ultimately interested in her Thomistic view; so I will not challenge her exegesis.

Frankfurt does not adequately argue either of these points. In other words, if Frankfurt's account is taken on its own, serious objections remain.

One such objection is that Frankfurt's arguments against *PAP* fail for two fundamental reasons. First, Frankfurt's examples appear to presuppose causal determinism.[37] How else can the intervener know that Jones is going to vote Republican? The intervener, Black, sees some prior sign (whether external or internal) and knows based on that sign which way Jones will vote. Black can know "if x (Jones blushes) then y (Jones votes Republican)" only if x and y are part of the same causal chain, and as soon as one event happens it is inevitable or determined that the other event will happen. Some may dispute that this is an instance of determinism, but it fits with Robert Kane's definition of determinism, which is representative of the most common contemporary view:

> An event…is determined when there are conditions obtaining earlier (such as the decrees of fate or the foreordaining acts of God or the antecedent causes plus the laws of nature) whose occurrence is a sufficient condition for the occurrence of the event. In other words, it *must* be the case that, *if* these earlier determining conditions obtain, then the determined event will occur.[38]

Black can *know* that Jones will or will not vote republican if some prior sign obtains such that "*if* these earlier determining conditions obtain, then the determined event will occur." If it is not the case that once the prior sign occurs Jones will necessarily perform a certain action, then there is no way for Black to know whether to implement his neural inhibitor

---

[37] Some might argue that all that is needed for Frankfurt cases to work is a probabilistic or locally reliable connection between the sign and Jones's voting, and therefore the universal claim of causal determinism is not required. I will argue below that I think that this view is mistaken and that the only way for the case to do the work that Frankfurt wants it to do is if the connection is absolutely reliable or deterministic.

[38] Kane, *A Contemporary Introduction to Free Will*," (Oxford: Oxford University Press, 2005), 5-6. Quoted in W. Matthews Grant, "Can Libertarians Hold that Our Free Acts are Caused by God?" *Faith and Philosophy 27:1* (2010): 25.

to ensure that Jones votes the way Black wants. In other words, it appears that if causal determinism is not true, then Jones can perform the prior sign (that is supposed to tell Black how Jones is going to vote) and still vote either way, because the occurrence of the prior sign is not a "sufficient condition for the occurrence of the event" (again, assuming now that causal determinism is false).

In addition to agreeing that alternative possibilities are not necessary for responsibility ascriptions, Stump also offers a modified Frankfurt-style hierarchical account of the will. Her modifications deal with one of objections that have been leveled against Frankfurt's hierarchical account. Stump puts the primary objection this way:

> …there is what has come to be called 'the problem of authority.' What is there about second-order desires, or any higher-order desires, that makes them authoritative for a person? Why should one set of desires be any more authoritative than any other? Why suppose that it is constitutive of the nature of a person to have desires about desires? Gary Watson puts the point this way: 'Since second-order volitions are themselves simply desires, to add them to the context of conflict is just to increase the number of contenders; it is not to give a special place to any of those [desires] in contention.'"[39]

Stump agrees that this is a serious problem with Frankfurt's account, but she does not think that is it damning of *all* hierarchical accounts. Utilizing insights from Aquinas she attempts to articulate an account of personhood that shows why higher-order desires are authoritative and thus freedom-granting in a way that first-order desires are not.

Stump's Solution to the Hierarchical Problem

Stump argues that Frankfurt's hierarchical account of identification and Fischer and Ravizza's attempt to provide an alternative compatibilist account of moral responsibility fall prey to the same basic problem. These accounts focus on only one

[39] Stump, "Persons: Identification and Freedom," 184.

aspect, or mental capacity, of human agency. Frankfurt is a volitionalist who thinks that freedom is a feature of the will alone. Alternatively, Fischer and Ravizza attempt to answer the authority problem by locating freedom in the intellect or what they call the agent's reasons-responsive mechanism. Neither of these approaches adequately addresses the problem of authority.[40] However, according to Aquinas's account, freedom is a systems-level feature that emerges from the proper and dynamic interaction of the intellect and the will, in which the intellect takes the primary role.[41] Stump argues that only an understanding of human freedom that incorporates both volitional and intellectual elements can avoid the problem of authority and thus provide the freedom-relevant conditions for moral responsibility.

To understand Stump's response to Frankfurt, Frankfurt's most recent and succinct articulation of identification proves helpful. Frankfurt describes identification as the process where the agent first temporarily suspends or brackets his relationship with a desire, then reflects on the desire, forms another desire or attitude towards that desire (a second-order desire), and finally either identifies herself with or alienates herself from that desire. When an agent identifies with a desire, she endorses that desire and accepts it as her own. "Higher-order desires provide backing for, or withhold backing from, first-order desires. That is, they draw the agent to get behind his first-order desires or then draw him to put himself against those desires."[42] For Frankfurt "[t]he higher-order

---

[40] More importantly, neither of these approaches can adequately distinguish between free and responsible actions and actions that result from responsibility-undermining external forces such as manipulation.

[41] *Ibid.,* 194.

[42] Frankfurt, "Reply to Michael E. Bratman," in *Contours of Agency: Essays on Themes from Harry Frankfurt*. Cambridge, Mass: MIT Press, 2002, p. 87. This response answers critics like Bratman who examine what Frankfurt means when he says that an agent "evaluates" or "reflects on" a desire before

attitudes that are formed in processes leading to identification involve 'evaluations' only in the sense that is strictly value-neutral."[43] One may think that, at the very least, an agent who reflectively endorses and identifies with a desire is taking that desire as a *justifying* reason for action. In other words, the agent has a desire to X, and on reflection thinks that action X is justified (for any reason, moral or otherwise). She then decides that she wants to want to X, thus identifying with her desire to X. Although Frankfurt admits that this does happen, he does not think that this kind of identification is necessary for freedom of the will or moral responsibility. "For someone to identify with a desire means merely that—for whatever reason, or for no reason whatever—he joins himself to the desire and accepts it as his own."[44]

In order to better understand what Frankfurt means by "value-neutral" evaluation, it is helpful to invoke Charles Taylor's distinction between strong and weak evaluation. In his essay "What is Human Agency?" Charles Taylor begins by assuming Frankfurt's distinction between first and second-order desires. He agrees that the reflective evaluation of our desires is a fundamental part of what it means to be human, but makes a further distinction between strong and weak or qualitative and quantitative evaluation (and argues that strong evaluation is the defining mark of human agency). One difference between these two kinds of evaluation is that weak evaluation is often concerned with *outcomes*, whereas strong evaluation is concerned with the *quality* of our

---

she identifies with that desire. Most critics assume that evaluation involves assessments of value, including but not necessary limited to moral value.

[43] *Ibid.* Frankfurt regrets his use of 'evaluation' in his earlier work for precisely this reason. He does not think that an authoritative higher-order desire must be the result of judging the desire to be good. Charles Taylor's distinction between weak and strong evaluation is helpful here.

[44] Frankfurt, "Reply to Michael E. Bratman," 89.

motivation.[45] The most fundamental difference is that strong evaluation is concerned with qualitative distinctions in *worth*. In other words, some of our second-order evaluations are concerned merely with distinctions of preference, such as feeling like having chocolate cake for dessert instead of pumpkin pie. There is no real difference in value or quality here; the decision is merely a result of reflection on one's desires and the subsequent decision to endorse the stronger desire. Thus, according to Taylor weak evaluations are not necessarily quantitative; they cannot always be weighed or calculated. The options may have "desirability characterizations," but I choose one over the other just because "I feel like it" (as in the case of choosing cake over pie).[46] Strong evaluation, however, makes a "qualitative characterization of desires as higher and lower, noble and base, and so on."[47] Though Frankfurt's notion of care involves strong evaluation, his general theory of identification (which is theory of the freedom-relevant conditions necessary for grounding moral responsibility) only requires weak evaluation. In this weaker form of identification, the agent wants to do X, reflects on the desire to do X, and the mere desire is taken by the agent as a sufficient reason to perform X. Due in part to this notion of caring and in part to his careless use of the term "evaluation," many critics have assumed that Frankfurt intends a stronger form of identification. The critics assume that when the agent wants to do X, and on reflection thinks that the action is

---

[45] Taylor, Charles, "What is Human Agency?" 16.

[46] Ibid., 17.

[47] Ibid., 23. Owen Flanagan sums up Taylor's ideal of the strong evaluator as someone who is "concerned that her desires, commitments, plans and character satisfy high ethical or spiritual standards. What makes her evaluation strong is that she engages in systematic moral inspection." Owen Flanagan, *Self-Expression: Mind, Morals, and the Meaning of Life* (New York: Oxford University Press, 1996), 205. Quoted in Peter Mehl, *Thinking through Kierkegaard: Existential Identity in a Pluralistic World* (Urbana: University of Illinois Press, 2005), 4. I believe that Peter Mehl is right when he argues that Taylor's concept of the strong evaluator aptly sums up Kierkegaard's "normative conception of personhood."

justified (because it is valuable, a moral obligation, etc.), then the agent has a second-order desire that provides sufficient reason to act. However, Frankfurt's notion of identification does not utilize a stronger form of evaluation but rather the weaker form that requires only that the agent reflects on and identifies with the desire.

Understanding that Frankfurt is utilizing a weaker sense of identification illuminates Stump's criticism.[48] She claims that Frankfurt "takes freedom to be a feature of the will," but what she means by this is not entirely clear. Based on her Thomistic modifications of Frankfurt's hierarchical structure, her basic criticism is that Frankfurt does not think that the intellect plays a primary role in freedom or moral responsibility. She is right, but this point requires nuance. Frankfurt clearly *does* think that freedom requires reflection or evaluation and that these are not functions of a person's will but of the intellect. The intellect does play a role in freedom and moral responsibility on Frankfurt's account. However, that role is severely minimized, due to the weak sense of evaluation involved in Frankfurt's notion of identification. In Frankfurt's view, a person can have second-order volitions without the intellect "approving" of the desire, in the sense of considering the desire as good under some description. The intellect does not have to think that the desire is good or justified in any sense; rather, it just has to think that that desire is *desirable*. In other words, the agent just has to think (perhaps rather covertly) "do I want to eat the cake?" and answer "yes, I like that desire and I want to want to eat the cake." To put the point another way, the will is not *constrained* in any way by the intellect. This differentiates Frankfurt from Stump and Fischer who both

[48] Though this is anachronistic since Stump was unaware of these clarifications when she wrote her articles, I think that she anticipated this articulation of identification. I will give a fuller treatment of this issue in chapter four that will show that Frankfurt uses the weaker sense of identification for much of his career including when Stump formulated her critique.

argue that the will is constrained by the intellect when making free and responsible choices. So, I take Stump to mean that the intellect should have a constraining role, and that the basic problem with Frankfurt's theory of freedom and responsibility is that he takes freedom to be primarily a feature of the will. Frankfurt still finds cognitive elements in free and responsible choices, but the cognitive elements take a backseat to the volitional element and in no way can constrain or be a final cause of the will.[49]

Stump's understanding of free and responsible willing begins with Frankfurt's basic hierarchical distinction between first and second-order desires. She adds first and second-order volitions to the mix where volitions are effectual desires.

> If an effective desire is one which moves the agent all the way to action, then an effective second-order desire is one which moves the agent all the way to the action of making the corresponding first-order desire his will. So a second-order desire constitutes a second-order volition only if it is an effective desire and the agent has a first-order volition corresponding to it.[50]

Thus a person can have conflicting first and even second-order desires, but not volitions.[51] Even with this clarification, the fundamental problems remain. First, how does an agent come to identify herself with certain desires while alienating herself from others? And second, why are second-order desires authoritative—i.e. freedom and responsibility granting—in a way that first order desires are not? Stump thinks that the solution lies in giving the intellect a stronger role without eliminating the fundamental role of the will. She argues, following Aquinas, that the intellect is the final mover of the

---

[49] We will return to this issue of which faculty is primary and whether the intellect constrains the will in free and responsible choices in Kierkegaard's understanding in chapter five.

[50] Stump, "Sanctification, Hardening of the Heart and Frankfurt's Concept of Free Will," in *The Journal of Philosophy* LXXXV, no. 8 (1988): 401.

[51] Perhaps it is possible to conceive of a situation where a person has conflicting first-order volitions so long as the volitions aren't such that they both must be actualized at the same time and in the same way but yet are still conflicting. But I don't think that this is the norm.

person. Also, since the will is a natural appetite for goodness it can only will something the intellect perceives as good in some sense.

This Thomistic account is both similar to and fundamentally different from Frankfurt's conception of the will. For instance, Frankfurt thinks that self-preservation is "solidly entrenched in our human nature" and therefore "we cannot help caring about avoiding crippling injury and illness, about maintaining at least minimal contact with other human beings, and about being freed from chronic suffering and endless stupefying boredom."[52] These deeply entrenched desires generate reasons to act, not because the agent upon reflection thinks that the desires are good, but just because human beings naturally desire those things. The formation of authoritative second-order desires can occur simply by means of the agent's reflecting on those desires and adopting them as her own. The agent could reflect on this desire and consider it good and then adopt it as her own for that reason, but this strong evaluative aspect is not necessary on Frankfurt's account. Therefore even though Frankfurt would agree with Aquinas that human beings *qua* human beings have certain natural desires, he does not think that these natural desires are a reason to give the intellect any fundamental role in forming second-order desires.

Stump thinks that giving the intellect a primary role in the formation of second-order desires will solve the problem of authority. To give the intellect this necessary role, Stump revises Frankfurt's account in the following way:

> An agent has a second-order volition V2 to bring about some first-order volition V1 in himself only if the agent's intellect at the time of willing represents V1,

[52] Frankfurt, *Taking Ourselves Seriously and Getting it Right*, 38. Frankfurt is careful to point out that he does not think that even these desires are logically necessary or necessarily constraining, but that it is just the way it is for most human beings and therefore someone that does will their own destruction or pain, for instance, just doesn't make any sense to the rest of us. We cannot even make a good argument against those people. All we can do is point out that their desires appear absurd to the rest of us.

> under some description, as the good to be pursued. A second-order volition, then, is a volition formed as the result of some reasoning (even when the reasoning is neither rational nor conscious) about one's first-order desires.[53]

Whereas Frankfurt thinks that second-order desires can be formed as the result of weak evaluation (mere approval), Stump believes that second-order desires or volitions are the result of strong evaluation which means that the intellect must perceive the first-order desire or volition as good in some sense. As an example Stump considers the case of an incontinent person, Irene, who has conflicting first-order desires about standing up in favor of the minority opinion at a faculty meeting. On the one hand, Irene wants to stand up for what she believes is right; but, on the other hand, she has the conflicting desire to vote with the majority due to anxiety over the potential backlash. She gives in to her anxiety and votes with the majority. According to Stump, on Aquinas's account, Irene forms this volition to vote with the majority because in some swift calculation (perhaps not fully conscious) she decides that it would be good for her. In the moment she views not standing up for the minority view as 'good' in some prudential sense, though, at the same time, her intellect judges the failure to stand up for what she believes is right not to be morally good.[54] Stump thinks that this second-order volition to give in to the majority is authoritative and freedom granting in a way that a mere first-order desire is not. Her second-order volition develops as a result of intellectual reflection on what would be *good* to do (even though this reflection is on what's good to do in that particular moment), instead of a simple reflection on which desire is more *intense* or more *desirable*

---

[53] Stump, "Sanctification, Hardening of the Heart and Frankfurt's Concept of Free Will," 400.

[54] Stump, "Persons: Identification and Freedom", 201-02.

(which is all it takes to form an authoritative second-order volition on Frankfurt's weak identification account).

Stump also thinks that her account provides a second form of identification that is lacking in Frankfurt's. The motivation behind the hierarchical account is that when the agent identifies with that particular desire she makes it her own while the rejected desires are *not* her own. Stump argues that a desire that results from a person's intellect deeming it as good under some description makes that desire authoritative and freedom granting. That desire is connected to the entire person (both intellect and will) in a way that second-order desires are not on Frankfurt's voluntaristic account. Irene's potential reaction to her decision illuminates this point. On reflection, she might be upset at her decision. She realizes that she should not have let her anxiety overcome her and should not let prudential concerns outweigh larger moral considerations. She retrospectively repudiates as not good the will not to stand up for her beliefs. As a result of this reflection she now has a second-order desire for a will that wills to stand up for what she believes is right.[55] Therefore, Irene "is to be identified with her second-order volition [the post-reflection will to stand up for her beliefs], as she is not to be identified with her first-order volition, because her second order volition derives from the all-things-considered judgment of her intellect in a way that her first-order volition does not."[56] Stump is saying that because the intellect is prime mover of the person and a person can never will anything unless the intellect perceives that action as good under some description, the all-things-considered-judgment of a person constitutes a person in a way

[55] Ibid., 203.

[56] I think that Stump's analysis of Irene is problematic, and I will return to it at the end of the chapter.

that a snap decision does not. Importantly, though Irene's first decision was the result of a snap judgment, it is still authoritative and freedom granting. This decision depends on the judgment of the intellect and therefore is also connected up to the entire person (intellect and will) in the way that Stump thinks is necessary for identification to avoid the problem of authority.

On the whole, Stump's Thomistic account of identification is an improvement over both Frankfurt's and Fischer's. She correctly argues that any account that focuses primarily on the will or the intellect is lacking since it does not reflect the judgment of the entire person. Critics like Watson and others rightly note that a second-order desire that is simply the result of the agent reflecting on her desires and deciding which one is more desirable does not help demonstrate why some actions are free while others are not. This weak form of second-order desires and volitions at most shows that the person with higher-order desires wanted what she wanted more than the person with only first-order desires and volitions. It does not show that the person with higher-order desires and volitions is any more *free* or *responsible* than the person lacking higher-order desires. On the other hand, Fischer rightly notes that an intellectual component like reasons-responsiveness is necessary for a person's will to be her own. But, as I argued in chapter two, despite his improvement on Frankfurt, Fischer's account is unable to distinguish free and responsible action from action that results from responsibility-undermining external forces such as manipulation. Cases of manipulation help isolate the characteristics of actions that are not the agent's own but are caused by a source outside the agent. Differentiating manipulation from free actions shows what makes an action the agent's own in the sense required for moral responsibility and the freedom necessary for moral

responsibility. Since Fischer's notion of guidance control fails to make this distinction, he also fails to show that an account that focuses primarily on intellectual faculties can provide the conditions necessary for ownership or identification. Stump's theory is therefore an improvement. It gives proper place to the intellect and will in free and responsible action. Furthermore, whereas Fischer favors an historicist account because of the problem of authority, Stump shows that a properly articulated hierarchical account can solve the problem of authority.[57]

Stump's Arguments Against PAP

One of the main components of Stump's source compatibilism is her rejection of a strong control (alternative possibilities) condition for free and responsible action. This rejection rests on her own particular articulation of Frankfurt-style counterexamples. She takes these counterexamples as sufficient to show that an agent can be morally responsible without having access to alternative possibilities. Many philosophers have challenged the validity of *FSCs*. Stump thinks that Frankfurt's own articulation of the counterexamples, as well as the counterexamples offered by other compatibilists such as Fischer, imprecisely articulate the mechanism utilized by the manipulator in those examples.[58] Stump offers her own *FSC* that utilizes a clear mechanism and is thus not open to some of the same criticisms as other *FSCs*. Though I agree that her example is an improvement over many other examples, she still fails to show that alternative possibilities are not necessary for moral responsibility.

---

[57] In the next chapter, I will address problems with Stump's account, namely that it puts too much emphasis on the intellect in free and responsible action.

[58] Stump, "Alternative possibilities and Moral Responsibility: The Flicker of Freedom," 303.

All *FSCs* have a similar structure: an agent performs an action in circumstances that would incline most people to conclude that the agent is responsible for that action.[59] However, in the scenario a manipulator, unknown to the agent, insures that the agent performs the action that the manipulator wants her to perform. For instance, the manipulator, Black (a liberal neuroscientist), wants the agent, Jones, to vote for the democratic candidate in the election. Black has a device that allows him to monitor Jones's internal decision-making mechanism. If Jones votes for the liberal candidate on his own, Black simply will let him perform that action on his own. However, if Black sees that Jones is going to vote for the republican candidate, he stimulates his brain so that Jones votes for the democratic candidate. In all *FCSs* the manipulator does not act on the agent because the agent—on his own and for his own reasons—does what the manipulator wants him to do. Jones votes for the democratic candidate and all Black does is monitor Jones's brain activity. This scenario generates the intuition that Jones is responsible for his actions because he did what he wanted on his own, without the intervention of the manipulator. But the crux of the argument for the compatibilist is whether or not Jones has access to alternative possibilities in this scenario. Frankfurt and his followers assume that Jones does not have such access. If Jones is going to vote for the republican candidate then Black will stimulate his brain so that he votes for the democrat instead.

Critics note a number of problems with the early formulations of FSCs. Fischer articulates the main objections in terms of a dilemma: "the Frankfurt-type stories presuppose either that causal determinism is true, or that it is false. If the former, then the claim that the relevant agent is morally responsible is question-begging, and if the latter,

[59] Ibid.

then the claim that the agent lacks alternative possibilities is false."[60] Laura Ekstrom expresses the first half of the dilemma in this way: if causal determinism is true, then the laws of nature and the past are a sufficient explanation for Jones's actions. Jones's perception of his actions is irrelevant, because he is *pushed* to act in a certain way. So whether or not determinism is true *is relevant* to the question of Jones' moral responsibility. If causal determinism is true then the question is whether or not an agent whose actions are caused by external forces can be morally responsible, not whether or not alternative possibilities are necessary for moral responsibility.[61] The second horn attacks the notion that a prior sign can help Black accurately predict Jones's behavior. If indeterminism obtains, then even at the point of the sign, the future will unfold indeterministically, and there is no way of knowing for *certain* how Jones will act. At the very least, this leaves room for Jones to *begin* to make the choice or form the intention to make the choice. One could argue that this beginning to make a choice is the minimum alternative possibility necessary for grounding moral responsibility.[62] Arguments that utilize this latter strategy are known as "flicker of freedom" arguments because a small amount of freedom remains available to the agent.[63] Thus the challenge is twofold: either create an FSC that operates in an indeterministic world, or create an example that is able to sidestep the issue of causal determinism while making the argument that it is intuitively plausible to ascribe moral responsibility to an agent's actions even if she lacks

---

[60] Fischer, "Frankfurt-Style Compatibilism," in *Contours of Agency,* 4.

[61] Ekstrom, Laura Waddell, "Protecting Incompatibilist Freedom," *American Philosophical Quarterly* 35 no. 3 (1998): 284-85. Cited in Fischer, "Frankfurt-Style Compatibilism," 5.

[62] Fischer, "Frankfurt-Style Compatibilism," 5.

[63] Timpe gives Fischer credit for coining the phrase "flicker of freedom." Timpe, *Free will: Sourcehood and its Alternatives*, London: Continuum. (2008), 55.

alternative possibilities. Fischer focuses on the latter strategy, and Stump finds these strategies lacking in precision and needing revision.

Fischer identifies four different flicker of freedom strategies against Frankfurt style cases. All of these strategies argue that the counterfactual intervener does not completely eliminate alternate possibilities; and, the ineliminable alternative possibility, or "flicker of freedom," is necessary for grounding moral responsibility.[64] The first strategy (and the one that Stump is mainly concerned with) states that the agent gives some indicator that lets the manipulator know the agent is about to act in the way that manipulator does not want her to act, and that the agent retains the freedom to initiate the alternative sequence even if she cannot complete that sequence. Advocates of this strategy argue that the agent's ability to initiate this alternative sequence is the minimum requirement for moral responsibility. Since FSCs do not eliminate this small but real alternative possibility, they do not prove that moral responsibility is ascertainable apart from alternative possibilities.[65] One of Fischer's favorite examples of guidance control illustrates the flicker of freedom objection well. In the example, a person drives a car that can only turn left. As long as the person intentionally turns left and is physically able to carry out her intention, she has guidance control but not regulative control. She does not have regulative control since if she tries to turn right the steering apparatus will fail (thus she can guide the car in the direction she wants it to go but cannot regulate whether it goes one direction or the other). Fischer acknowledges that this case is problematic if

---

[64] Fischer, "Responsibility and Alternate Possibilities," *My Way: Essays on Moral Responsibility*, 41.

[65] This strategy seems to assume that the world is indeterministic because the agent can either blush or not blush for instance.

used as an FSC. Clearly the person driving the car can form the *intention* to turn right, and an advocate of the need for regulative control for moral responsibility can argue that so long as a person can form an intention to do otherwise she is morally responsible for her action.

Fischer's own early FSCs did not make clear how the manipulator was able to intervene in Jones's actions. Fischer attempts to rectify this problem by reconstructing the Frankfurt scenario. The revised version states that "there is some sign or indication that would precede the initiating action and could be read by the counter-factual intervener…" and this prior sign in the revised FSC will reveal how the agent will act *before he forms the intention*.[66] This prior sign in a typical FSC might be external (such as a blush) or internal (such as the firing of certain neurons); either way, Fischer stresses that the prior sign happens before Jones forms a conscious intention or makes a conscious choice, thus eliminating the agent's intentional initiation of an alternative sequence. Despite this revision, the claim can still be made that Jones retains some level of freedom. If the world is not causally determined, then he may either blush or not blush, and therefore he still retains the freedom to initiate an alternative sequence which libertarians claim is necessary to ascribe moral responsibility to the agent.[67] This freedom is controversial, since it is very difficult to make the argument that something like blushing is under one's voluntary control. Nonetheless, instead of eliminating this possibility of the initiation of an alternative sequence, Fischer asserts that whatever

---

[66] Ibid., 48.

[67] There is also the related objection that I will focus on later, namely, that FSCs assume that if a certain action happens in the agent (whether physical or neurological) then the manipulator knows with absolute certainly that the agent will perform a certain action. In other words, FSCs assume a world in which if A happens then B must happen necessarily.

alternative possibility remains in the FSC is not sufficiently "robust" (Fischer's term) to ground moral responsibility.[68] He thinks that the only alternative possibility left for the agent in his FSC is the mere possibility of unintentional or involuntary behavior. This behavior is not enough to ground moral responsibility and "…there is a crucial difference between the ability to do otherwise and the mere possibility of something different happening."[69]

Fischer takes it that his revised FSC is sufficient to counter the early objections, but the best objections (such as those listed above) have been in response to his revisions. The fact remains that Fischer has not devised a way to eliminate all alternative possibilities without presupposing causal determinism. Many incompatibilists argue that as long as some kind of leeway remains then a viable argument can still be made that alternative possibilities are part of the grounding mix for responsibility ascriptions. Stump realizes that Fischer's *FSC* remains open to this criticism. She revises Fischer's example to clarify how the mechanism works. She thinks that doing so will help her to defeat the flicker of freedom objection.

In her example, a neurosurgeon named Grey wishes to ensure that Jones votes for the republican candidate in the upcoming election.[70] Grey has a powerful neuroscope that can both observe and bring about the firings of neurons that will in turn correlate with Jones's acts of will. Grey is able to observe that every time Jones votes for a Republican candidate, there is a complete neural sequence in his brain that includes at or

---

[68] Ibid., 6.

[69] Ibid., 8.

[70] The example that follows is found in Stump, "Alternative Possibilities and Moral Responsibility: The Flicker of Freedom," 303-305.

near the beginning the firing of neurons a, b, and c. Whereas, when he votes for a Democratic candidate the neural sequence includes the firings of neurons x, y, and z. If Grey observes the firings of neurons x, y, and z, he can interrupt that sequence and keep Jones from voting for the democratic candidate, and initiate the firings of neurons a, b, and c insuring that Jones votes for the Republican candidate. However, if Grey observes the firings of neurons a, b, and c, he does nothing and allows Jones to vote for the republican candidate. In this case, Jones votes for the person that he wants without any intervention from Grey, but seemingly without an alternative possibility as well. Yet intuitively, Jones is morally responsible because he voted the way he wanted.

To assure that Grey is an improvement upon Fischer's *FSC,* Stump specifies two presuppositions related to the philosophy of mind that she is employing in her *FSC*, both of which Frankfurt and Fischer neglect. First, she notes that she is presupposing that "the mind is at least implemented in the brain and that therefore there is some correlation between mental states and neural states." Furthermore she claims that "there is a strong connection between a mental act or state and a neural state" while leaving the precise nature of that connection "general and vague."[71] This first presupposition takes Cartesian dualism out of the equation. According to at least the most extreme versions of Cartesian dualism, agents are disembodied souls and no necessary correlation obtains between the mental/spiritual and the neural/physical. Her *FSC* does not work against extreme Cartesian dualism, but it does press anyone whose philosophy of mind assumes a strong connection between the neural states and mental states/acts to admit that in theory

[71] Ibid., 305.

it might be possible to disrupt the neural sequence and prevent the mental state/act from happening.

The second presupposition of Stump's revised *FSC* is that "the correlation between a mental act or state and the firings of neurons is a one-many relation."[72] By stating that the "mental act or state and the firings of neurons is a one-many relation" she claims that a *complete sequence* of neural firings is necessary for a mental state to emerge. In other words, if the neural sequence that gives rise to a particular mental act is cut off, the result is not a truncated or different mental act but no mental act at all.[73]

> To suppose that there could be some sort of mental act, truncated, incomplete, or otherwise defective, when there is no completed neural sequence correlated with that mental act, is to accept some version of Cartesian dualism. It is to suppose that there can be a mental act without there being a completed neural sequence correlated with that mental act.[74]

This point is crucial to Stump's argument against a flicker of freedom in *FSCs*. It allows her to argue that Jones's mental act can be prevented by Grey without Jones having any awareness that he has done anything other than what he wanted to do. Jones will not stop and think to himself "for a moment there I thought that I was going to vote democrat but I must have changed my mind." Jones will have no sensation of almost acting differently, no premonition that he was about to vote differently; he will have no thoughts whatsoever about alternative actions because the neural sequence was interrupted and an alternative sequence will have no possibility for initiation.

---

[72] Ibid., 306.

[73] Ibid., 307.

[74] Ibid.

Stump concludes that "[t]hese presuppositions together have the result that there is no act of will in an agent unless and until the correlated sequence of neural firings in that agent's brain is completed."[75] However, Stump's argument actually rests on three ideas, the most crucial of which she never makes clear. The first two are her admitted presuppositions, namely the correlation between neural sequences and mental states/acts and the one-many relationship between mental acts/states and neural sequences. Stump concludes from her two presuppositions that alternative paths available to the agent can be blocked. The agent cannot have an act of will without the corresponding *completed* neural sequence. This means that the neurosurgeon, Grey, must be able to tell that Jones is going to perform a particular action based on the firing of a few neurons. Stump points out that, for simplicity's sake, the firing of neurons a, b, and c is *only* associated with neural sequence *R* (the sequence that results in voting for the republican candidate) and that particular act of will; while x, y, and z is *only* associated with neural sequence *D* (the sequence that results in voting for the democratic candidate) therefore Grey will never be confused about which act of will is about to happen in Jones. Stump also stresses that this connection need not be law-like, but a high level of consistency over a certain period of time will allow Grey to predict accurately which act of will is about to happen in Jones. Even if we grant Stump this high level of consistency, surely room still remains for variation in what complete neural sequence will emerge from the firings of neurons a, b, and c. Given the complexity of the brain, I see no way that anyone could know with the certainty necessary that Jones is going perform a certain mental act based on the firing of only the first few neurons unless a law-like connection *is assumed* between the

[75] Ibid.

firings of just a few neurons and an entire neural sequence, or a law-like connection is assumed between particular neural sequences and corresponding mental acts. I flesh out this criticism in the section below, but it is necessary to mention that Stump must also presuppose this law-like connection between neural sequences and mental states/acts, and this will prove damning for her revised FSC.

Stump argues that the flicker of freedom strategies rely on one of two features of *FSCs*. The first feature found in all *FSCs* is that the manipulator must respond to something in the agent, and whatever the neuroscope (or other device) responds to in the agent, the agent must retain the freedom to do or not do that to which the device responds. For instance, if the trigger is the intention to vote, then the agent retains the ability to form that intention or not to form that intention. But if the prior sign is forming an intention (or something similar), then the manipulator detects another mental state before he intervenes. A mental state is something of which the agent is aware and, I think, robust enough to ground responsibility ascriptions. Stump agrees with leeway incompatibilists like Widerker that these *FSCs* seem to assume covertly that the agent is causally determined to act as she does in the actual sequence and that this sidesteps the main issue at hand, namely, whether an agent is responsible for a causally determined action.[76]

Stump spends more time with a second objection that Fischer, following van Inwagen, labels the "essentialist principle of event-individuation" objection. The event-individuation objection states that when an agent performs an action, she causes a specific event to occur. According to the strong essentialist version of event-

[76] Ibid., 312.

individuation, "all the actual causal antecedents of a particular event are essential to it."[77] Since all the causal antecedents are essential to the event, two distinct sequences could emerge from manipulation cases. In the first, the agent causes the actual sequence in which the agent does what she wants to do on her own, without any intervention from the manipulator. This sequence is distinct from the alternative sequence where the manipulator's intervention is a "causal antecedent" of the event.[78] Stump argues that an important assumption is at work in this strategy, namely, that in the actual sequence the agent is performing the action on her own. The action in the counterfactual sequence appears to be the same action. Yet according to the event-individuation objection, it is *not* the same action. Therefore the assumption is that the action of *doing X on one's own* is a distinct action from *doing X*.

> They [defenders of flicker of freedom strategies] must take the victim's doing W-on-his-own as something which the victim does, and they must suppose that doing W-on-his-own is not identical to doing W. Both these assumptions are necessary to their case. If doing W-on-his own weren't an action the victim does, then there wouldn't be something the agent does in the actual sequence but omits to do in the alternative sequence, as the flicker of freedom proponents argue. And if doing W on-his-own weren't different from doing W, then what the victim does in the actual and the alternative sequence would be identical, and the victim wouldn't have alternative possibilities available to him.[79]

Stump goes on to argue that in this alternative scenario the agent could rightly claim that she could not have done otherwise. Also, the agent would not be open to the suggestion that she had the alternative possibility available to her to *do the act on her own* instead of

---

[77] Fischer, Responsibility and Alternative Possibilities," 42.

[78] I find it illuminating that Fischer never responds explicitly to this objection, but instead seems to assume that no matter how the events may in fact be different, this difference is not robust enough to ground moral responsibility. In response to the claim that this difference is robust enough to ground moral responsibility, he argues that a prior sign can be developed that would precede the initiating action of the agent and therefore even if the events are different, the agent would not play any conscious role in which event comes about. This latter strategy is exactly what Stump is attempting to do.

[79] Stump, "Moral Responsibility and Alternative Possibilities," 314.

simply *doing the action*. In other words, if doing action W is what the agent wants to do, then *doing the action on one's own* does not seem like an alternative possibility to *doing the action*.

Stump thinks that this argument sufficiently shows that no viable alternative possibility exists in her *FSC*. Just in case the leeway incompatibilist insists that doing an action on one's own is distinct from simply doing an action, she goes a step further and attempts to construct an *FSC* that purports to prevent even this alternative possibility. If doing an action on one's own is a distinct mental act, then, on her account, it still must be preceded by a neural sequence since all mental states/acts are preceded by the correlated neural sequence. She concludes that this neural sequence can be interrupted by the intervener just like the original neural sequence in *Grey*. Why she bothers to make this point is unclear, since it merely reiterates her basic argument. This reiteration does show that she has full confidence in her view that mental states/acts always are preceded by neural sequences or that those neural sequences can be interrupted and even manipulated in such a way that all alternative possibilities are eliminated.

Sourcehood and Origination

Before considering critiques of Stump's revised *FSC* in particular and her source incompatibilism more generally, it is important to keep some distinctions in mind. Stump thinks that causal determinism is incompatible with moral responsibility because if causal determinism is true then the agent's own intellect and will cannot be the *ultimate* source[80]

[80] I am using ultimate source to indicate that the agent contributed some important causal factor(s) to the production of the action that cannot be traced back to casual factors outside of that agent's control. Furthermore, this agent-produced causal factor was a part of the specific cluster of causes that gave rise to the action in question.

or cause of that agent's action but rather that action is *entirely* the result of causal factors that are beyond the agent's control. She endorses something similar to this formulation of the ultimacy argument:

> (1) If determinism is true, then no one is the ultimate source of one's acts.
>
> (2) One is morally responsible for one's acts only if one is their ultimate source.
>
> (3) Therefore, if determinism is true, no one is morally responsible for one's acts.[81]

Furthermore, notice that for this formulation to do the work that Stump needs it to she must go beyond the weaker view of determinism as the view that every event is caused, and embrace the stronger view that every event has a *nomically sufficient condition*. Widerker defines a nomically sufficient condition in the following way: "an event *E* occurring at time *T* has a *nomically* sufficient condition if and only if the proposition that *E* occurs at *T* is entailed by the laws of nature and some proposition *p* that expresses the state of the world at some time prior to *T*.[82] If the agent's action has a nomically sufficient condition, then the agent contributed nothing to the action that was not directly traceable to casual factors (the laws of nature and some the state of the world at some point in the past) outside of the agent's control. Therefore Stump thinks that *in order for the agent to be responsible the agent must contribute something to the action that is not the direct result of the laws of nature and some prior state of the world*, and furthermore

---

[81] Widerker, David, "Farewell to the Direct Argument". *The Journal of Philosophy* 99:6 (2002): 321.

[82] Ibid., 323.

this action must be something that the agent does on purpose and that stems from that agent's intellect and will.

Stump's source incompatibilist emphasizes that the agent must make a causal contribution to the action that is not the direct result of the laws of nature and the events of the past. She is claiming if the agent makes this kind of contribution, and if that contribution stems from the agent's intellect and will, then the action *originated* in the agent in the sense necessary for moral responsibility. But there are other forms of sourcehood or origination as previously discussed in chapters 1 and 2. Stump is a source incompatibilist, while Fischer (and I argue Frankfurt as well) is a source *compatibilist*. Fischer also believes that the necessary conditions for morally responsible action include the agent being the source of the action and not a freedom-undermining external cause such as manipulation of the brain or compulsive phobia. Furthermore, the agent must possess a kind of control, control that is unavailable to an agent who's volitional and evaluative faculties are compromised in certain ways. But he does not believe that this kind or origination is incompatible with determinism. Finally, leeway incompatibilists (or traditional libertarians) believe that sourcehood is a necessary component of free and responsible action, but that this must be coupled with an alternative possibilities or avoidability component. I have already demonstrated how Fischer's source compatibilism fails to provide the freedom-relevant conditions necessary for grounding moral responsibility. In the final sections of this chapter I will demonstrate why Stump's source incompatibilism also fails to meet these conditions.

Critique of Stump's Revised FSC

I think that there are two basic approaches that one can take to attempt to show that Stump's source incompatibilism fails to provide the freedom-relevant conditions necessary for grounding moral responsibility. The first approach is to argue that Stump's revised *FSC*, *Grey*, fails to show that alternative possibilities are not a necessary part of the grounding mix for moral responsibility. Stump argues that the causal chain leading to the agent's action must be indeterministic for the agent to be the cause of the action, and not a responsibility-undermining external force.[83] However, her description of the connection between neural and mental states does not allow for the indeterminism she requires for moral responsibility. Furthermore, she claims that morally responsible actions must stem from the agent's own intellect and will. However, one could argue that to make this claim she must admit that the actions in the actual and counterfactual sequences stem from different intellects and wills; therefore, performing an action differs from performing an action on one's own in a morally relevant way. I treat each of these objections in turn.

Stump argues that mental states/acts are directly correlated to and result from (completed) neural sequences. Since the intervener disrupts the neural sequence that leads to the mental act, the agent does not have access to alternative possibilities. Furthermore, there is not even a "flicker of freedom." According to Stump, that flicker would have to be located in an intention, effort, or some other mental act. Since the intervener is disrupting a neural sequence and not a mental act, Stump concludes that her

[83] As specified in previous chapters, "external force" refers not to the force being located outside of the agent, but outside of the agent's *will*. A phobia, which is certainly internal to the agent, is external to the agent's will insofar is it not a desire/belief/etc. that the agent *identifies* with.

revised *FSC* eliminates even the smallest flicker of freedom from *Grey* and demonstrates that moral responsibility is intuitively plausible without alternative possibilities as a part of the grounding mix.

The first problem I find with this argument is that Stump argues that the world must be indeterministic for the agent (and not a responsibility-undermining external force) to be the source of the action, in the sense necessary for moral responsibility.[84] If the world must be indeterministic for the freedom-relevant conditions to obtain, then two things must be true. First, the indeterminism must be located in the causal process leading to the agent's action. Second, the agent must *cause* the action; i.e. the cannot be nomically sufficient conditions for the agent's action. To meet the first criterion the break in the causal chain can either happen somewhere *outside* the agent (i.e. either the natural laws are indeterministic or the facts of the past can be otherwise), or somewhere *inside* the agent. While indeterministic natural laws are possible, a mutable past is highly unlikely. If the natural laws are indeterministic and the break in the causal chain is located (only) in the indeterministic laws, the requirement that the agent, and not an external force, cause the action will not be met. Indeterministic causal laws might only result in the agent's actions being random. According to Stump's definition of responsibility-grounding agent causality, these laws render the agent nonresponsible as she is not the cause of the action. Therefore two things must be true on Stump's account to make responsibility ascriptions. First, the indeterminism must be located not only in the physical world, but also in the agent. Considering Stump's own understanding of

[84] This ties together the source and control requirements. For the agent to be the source of the action, the agent must cause the action in a manner distinguishable from an external force causing the action. If the agent causes the action in this manner, then the agent must have a level of control over her action.

mind/body issues, it is plausible to locate that indeterminism in the agent's brain. Second, since the motivation for locating the indeterminism in the agent's brain is to assure that the agent, and not some external force, is the cause of the action, the agent necessarily must have a level of *control* over the action.

These two claims have an important bearing on Stump's revised *FSC* and her claim that the agent's mental-state-causing-neural-pattern can be interrupted by an external agent. If the locus of the requisite indeterminism is in the brain, then it *cannot* be assumed that the firing of a few neurons (no matter how many or how complex) will *necessarily* be followed by one particular neural sequence. Kane establishes this point with his definition of determinism as "conditions obtaining earlier…whose occurrence is a sufficient condition for the occurrence of the event." If the firing of neurons a, b, and c necessarily leads to completed neural sequence R, then, like other *FSCs*, "the sign causally determined the action, or if it were associated with some factor that did, the intervener's predictive ability could be explained."[85] If the prior sign (in this case the firing of neurons a, b, and c) is causally connected to the completed neural sequence *R* in a deterministic way, then Stump's scenario has failed before it got off the ground because she assumes that free and responsible action is incompatible with determinism. On the other hand, if we grant Stump's claim that 1) free and responsible action is incompatible with determinism and 2) that the agent must have the requisite causal control to be responsible for an action (meaning that the action must stem from that agent's own intellect and will), then I conclude that two important points are implicit in Stump's modified libertarianism. First, indeterminism must obtain in the brain to break the causal

[85] Pereboom, "Hard Incompatibilism," in *Four Views on Free Will*, 90.

chain from external sources to the internal sources that ultimately cause the action. Second, the agent must be able to initiate the performance of the action that she wants to despite the fact that the indeterminism present in the brain would seem to allow for multiple actions.

If indeterminism is present in the brain but the agent's own intellect and will are the ultimate sources of morally responsible acts, then the agent must retain the requisite control and ability to intentionally bring one particular mental state out of many possible states, thus resolving the indeterminism.[86] Furthermore, this control might come in one of two basic forms. Either 1) the agent has the ability to initiate and, perhaps, alter the neural sequences that gives rise to mental states/acts, or 2) the agent can have simultaneous neural sequences (such as when a person has conflicting impulses about how to act in a particular situation) and can choose which action and which neural sequence will come to fruition.[87] I set aside option 2 (simultaneous sequences) for the moment to focus on option 1.[88]

In option 1, the agent is able to initiate and/or alter the neural sequences that give rise to mental states/acts. If the agent has the requisite control over her actions and therefore can control whether or not the neural sequences that give rise to acts of will happen or not, then the intervention by the neurosurgeon will not eliminate all potential alternative possibilities from the scenario. Jones (and not an external force) gets to

---

[86] Kane uses this phrase to refer to the way that responsible agents move from an internal indeterministic state where multiple outcomes are possible to controlling the initiation of one of those possible outcomes, thereby resolving the indeterministic tension.

[87] Kane proposes something much like this, and Kierkegaard's notion of the divided will—if located in the brain—could also take a similar form.

[88] I argue in chapter five that both of these options are a part of the grounding mix necessary for moral responsibility

decide whether or not he is going to vote Republican. Stump emphasizes that "the essence of freedom is that the agent's own mental faculties, her intellect and her will, are the ultimate sources of any free act, and not something outside the agent."[89] If Jones's action is to be a free and responsible one, he must be in control and this act of will must proceed from his own intellect and will (as Stump insists it must if it is a morally responsible act). That is, he must play a *causal role* in initiating the neural sequence that gives rise to the mental state/act to vote Republican. And if Jones plays some causal role in initiating the neural sequence that gives rise to the act of will to vote Republican, then *this* action is the locus of the alternative possibility that the intervener cannot eliminate without going back further in the causal history that leads up to Jones's decision. Stump claims that mental states/acts do not happen apart from corresponding neural sequences. Therefore, Grey does not have to interrupt a mental act of any sort. He can 1) accurately predict a completed neural sequence that will result in a specific act based on the firings of a few neurons (it may be more than a few neurons, but he must be able to intervene before the sequence is completed) and 2) interrupt that sequence without Jones being aware that anything has happened to him. Stump also claims that in order to be responsible Jones must be the ultimate source of his actions, but apparently he can be the ultimate source of his actions without controlling whether or not a particular neural sequence that causes a mental state/act takes place within him. Stump is apparently accepting the contradictory claim that although Grey's action must result from the proper interaction of his own intellect and will, Grey can be considered morally responsible

[89] Stump, "Libertarian freedom and the principle of alternative possibilities," 80. Cited in Timpe, "A Critique of Frankfurt-Style Libertarianism," *Philosophia* (2006), 198.

whether or not he controls his own intellect. Both of these claims cannot be true. I give what I take to be the proper resolution below.

Kevin Timpe provides a related objection. He notes that if we imagine a situation where a counterfactual intervener in an *FSC* causes an agent to make a decision, then surely the agent's intellect and/or will is no longer the ultimate cause of that action.

> If the intervener were to cause the agent to do an action either through bypassing her intellect or overriding her will, then the action would not be the one that originated in the agent's own intellect and will. In one sense then, the action in the two sequences would not be the same; the action in the actual sequence would be an act originating in the agent's own intellect and will, while the act in the alternate sequence would be one that originated in the intervener's intellect or will.[90]

Stump claims that her *FSC* is an example of an agent being morally responsible despite the fact that the agent cannot perform a different action. She also claims that for this to be the case both sequences (actual and alternative) must contain essentially the same action (doing act A on one's own is not distinct from simply doing A). But as Timpe points out, surely these actions are not numerically identical since they are products of different intellects and wills.[91]

> If the intervener's intervention bypasses the agent's own intellect and will in the alternative sequence, then the alternative sequence will contain a different action. Yet, in order for an *FSC* to be an example of a situation in which the agent is morally responsible without having the possibility of performing a different action, both sequences must contain the same action…(therefore) when Stump argues that that *FSCs* are effective in eliminating alternative possibilities for action, she cannot mean that the same action, narrowly conceived as originating from a specific intellect and will, occurs in both sequences.[92]

---

[90] Timpe, "A Critique of Frankfurt-Libertarianism," 198.

[91] Ibid.

[92] Ibid., 199.

Stump cannot have it both ways. If a free and responsible act must stem from the agent's own intellect and will, then doing an act on one's own is a distinct action and therefore there remains an alternative possibility in her revised *FSC*.

Timpe provides one final objection against Stump's revised *FSC*. He grants that perhaps she can formulate an *FSC* that overcomes the previous objection, but a fundamental problem remains. Stump's own evaluation of flicker strategies shows that she agrees with flicker strategists that there is a difference in the actual and the alternative sequence, but that

> the difference is not a difference between different actions the agent does, as the flicker of freedom proponents suppose. Rather, the difference has to do only with *how* the agent does what he does…[the agent] is causally determined to an act of will *W* in the alternative sequence, but not in the actual sequence. In the alternative sequence, the ultimate cause of what the victim wills is the intervener; in the actual sequence, it is the victim himself.[93]

Stump later adds that

> although it is not up to the victim whether or not he does the act in question, it is evident that the *mode* of the action is up to the victim…The one and only act open to the victim can be caused by the intervener or brought about by the victim of his own accord, and which of these modes is the one by which the act is done depends on the victim.[94]

This point is crucial. Stump admits that although she does not think that the actions in the two sequences are numerically distinct, the two difference sequences feature distinct *modes* of action. Furthermore, she agrees that the difference in mode is crucial for grounding ascriptions of moral responsibility. However, she still insists that this difference in the mode of action is not an adequate defense of the principle of alternative

---

[93] Stump, "Moral Responsibility and Alternative Possibilities," 322. Quoted in Timpe, "A Critique of Frankfurt-Libertarianism," 199.

[94] Stump, "Moral Responsibility Without Alternative Possibilities," 151. Quoted in Timpe, "A Critique of Frankfurt-Libertarianism," 199.

possibilities. Based on her own presuppositions and her insistence that a morally responsible action be performed in the right way, Timpe concludes that Stump must agree with a reformulation of the principle of alternate possibilities:

> an agent is morally responsible for doing an action *A* at time *t* only if she has alternative possibilities regarding the mode of action at time *t*.[95]

In this revised version of the principle of alternative possibilities, the victim does not control *whether* or not action *A* occurs, but she does control the *mode* in which that action occurs. Thus, the victim still has an alternate possibility available to her that is necessary to determine moral responsibility. The victim can either perform the action on her own (as a result of her own intellect and will) or the intervener can act on her, and become the ultimate source of her action. Timpe concludes that even though Stump denies the earlier flicker of freedom strategies, "her insistence on the importance of the mode of the action can be understood as a different version of the flicker strategy."[96] Therefore, Stump is a kind of flicker strategist even though she denies the necessity of the possibility of distinct actions in the actual and alternative sequences. She still must agree that some morally relevant alternative possibilities are available to the victim in Frankfurt-style cases.

The conclusion is that Stump's revised *FSC*, as ingenious as it may be, suffers from the same essential problem as all the others. In fact, the problem is compounded by her own requirements for morally responsible action. If indeterminism obtains in her scenario, then, at least, it seems impossible for intervener to predict which completed neural sequence would result from the firings of a few neurons. Furthermore, if indeterminism obtains and the agent must be the ultimate source of her actions, then that

---

[95] Timpe, "A Critique of Frankfurt-Libertarianism," 200.

[96] Ibid.

agent must be able to control which neural sequence is initiated. In this case, the necessary intervention is pushed back further into the agent's causal history making the neurosurgeon Grey's success at removing morally relevant alternatives impossible.

Timpe's Critique of Stump

Despite Stump's improvements on Frankfurt's hierarchical model, her acceptance of his arguments against PAP yields a "libertarian" account fraught with problems. Stump argues that "flicker of freedom" strategies are largely unsuccessful against *FSCs*. Some of these flicker arguments make the case that in the actual sequence the agent, Jones, is doing the action A *on his own*. In the counterfactual sequence, in which the intervener stimulates Jones's brain to guarantee that Jones does action A, Jones is no longer doing action A *on his own*, but as a result of the intervention. Yet, doing the action on one's own is something that many libertarians and compatibilists agree is necessary for moral responsibility. Libertarians argue that this alternative possibility—the possibility of either doing the action on one's own or *not* doing it on one's own—is the freedom relevant condition for ascribing moral responsibility. Stump denies that this is the case. She argues that the only way that this is a true alternate possibility is if doing an action on one's own is a numerically distinct action from merely doing the action. In other words, action O—doing action A on one's own—must be a different action from action W—doing action A; and, Stump does not think that this is the case.[97] She thinks that if the *FSC* is reversed, then it produces counterintuitive results. For instance, if a scenario is constructed in which the actual sequence results in action W, and the alternative sequence results in action O, that scenerio is clearly false. She states that "in

[97] Timpe, "A Critique of Frankfurt-Libertarianism," 196.

the analogue the agent has alternative possibilities for action."[98] She presses the point by arguing that if the agent in the analogue case were asked after the fact if she could have done otherwise, she would justifiably claim that she could not. She could not have acted differently because in the actual sequence she performed the action. So, no difference seems apparent between performing the action in the actual sequence and performing the action on her own in the alternative sequence.

Stump's insistence that action *W* is not numerically distinct from action *W* is peculiar. She asserts that what makes an action free is that it is ultimately caused by the agent's own intellect and will, and is not the outcome of a causally deterministic sequence that began outside of the agent. Yet, as Timpe observes, Stump is presented with a dilemma; either she must renounce her own incompatibilist libertarian position, or admit that the actual sequence is significantly different from the alternative sequence of an *FSC*. She must admit that the freedom-relevant condition in a typical *FSC* is the agent's ability either to be the source of her own action, or the intervener to be the source of her action.

Source Incompatibilism and the Direct Argument

Timpe's argument that the agent (victim) in *Grey* has the option to either act on her own or allow the intervener to be the source of the action and that this is the morally relevant alternative possibility available in the scenario is admittedly a difficult argument to sustain. After all, the agent in the scenario is not aware of the presence of an intervener, and therefore it is difficult to conceive of how any alternative possibility is available to the agent. I think that this is not the right counter to Stump or to

[98] Ibid.

incompatibilist *FSCs* in general. After all, one could continually press back against the particular features of the *FSC* claiming either that the *FSC* cannot work because it does not have the necessary predictive power or that some sort of morally relevant alternative possibility remain, and then the proponent of the *FSC* could continually refine the scenario *ad nauseum*, but this seems to result, as Fischer likes to say, in a dialectical stalemate. Instead, I think that the traditional libertarian should examine the direct argument that the source incompatibilist utilizes. The direct argument states that moral responsibility is directly incompatible with casual determinism. Many proponents of the direct argument (such as Stump) think that moral responsibility is incompatible with determinism not indirectly via alternative possibilities, but directly because determinism (the fixity of the laws of nature and the unchangeable past) makes it impossible that the agent be the source of the action in the way required for moral responsibility. In other words, moral responsibility is not possible in a deterministic world, but yet does not require alternative possibilities. The question to press against source incompatibilists here is this: what exactly is available to an agent (who does not have access to alternative possibilities) in an indeterministic world that is not available in a deterministic world? Or put differently, why does it matter that the world is indeterministic?

One way to pursue this question is to reconsider Stump's case of Irene. Why does Stump claim that Irene is responsible for her weak-willed action? After all, Stump does not want to say that she is *not* responsible. Gary Watson—critiquing his own earlier critique of Frankfurt's notion of identification where Watson claimed that what made a volition authoritative is that it stems from a person's evaluative system, or the agent's all-

things-considered judgment of what is best to will in a particular situation[99]—states that "[w]hen it comes right down to it, I might fully *'embrace'* a course of action I do not *judge* best; it may not be thought best, but is fun, or thrilling; one loves doing it, and it's too bad it's not also the best thing to do, but one goes for it without compunction. Perhaps in such a case one must see this thrilling thing as good, must value it; but, again, one needn't see it as expressing or even conforming to a general standpoint one would be prepared to defend."[100] This sounds much like Irene's weak-willed action. It is based in a snap judgment that Irene later repudiates, but yet in the moment she does it because she wants to and sees it as good under some description, which could be the same thing as seeing it as fun or thrilling or reputation-saving. In other words, Stump seems to agree with Watson that the reason that Irene is responsible is that her volitional and intellectual/evaluative faculties have *not* been compromised in freedom-undermining ways but rather Irene was able to make an evaluative judgment and her will followed suit. But whereas Watson seems to think that compatibilistic scenarios are possible where the agent's volitional and evaluative faculties are not compromised and therefore the agent is morally responsible, Stump wants to go a step further and say that if determinism is true then Irene's action would *not* have been free. Stump must think that Irene would not be responsible in a deterministic would because somehow the truth of determinism *alone* compromises Irene's volitional and intellectual/evaluative faculties in freedom-undermining ways. But how would this be the case? In a deterministic scenario every event is caused by and traceable to a prior event. In other words, Irene's decision *E* is

[99] "Watson, Gary. 1975. "Free Agency." *The Journal of Philosophy*. 72 (8): 215.

[100] Watson, Gary. "Free Action and Free Will," *Mind* 96 (1987): 150.

determined "if and only if the proposition that *E* occurs at *T* is entailed by the laws of nature and some proposition *p* that expresses the state of the world at some time prior to *T*."[101] But how does this fact alone—the fact that Irene's decision is entailed but the fixed laws of nature and unchangeable past—compromise Irene's volitional and/or evaluative faculties? She is not directly manipulated by some neuroscientist. She is not subject to any pathological phobias. She appears to reason normally and have the ability to draw her own conclusion.

To push to point, consider a modified example of Irene, $Irene_2$. $Irene_2$ sits in the faculty meeting and is disgusted by the treatment of minorities that she has both seen and personally experienced. She would really love to point out some of the atrocious actions she has seen her collogues commit. However, $Irene_2$ is also full of vainglory, wanting nothing more than to achieve all the status and glory that academia can provide. Knowing that maintaining a good standing in the department is essential for her continued success, she decides to vote with the majority despite the fact that she knows that on one level what she is doing is morally wrong. She knows she should stand up for the mistreated minorities (including her). After the meeting she experiences rather strong feelings of regret, but her regret at not having willed the good is *never* effectual. She never *repudiates* the will not to stand up for her beliefs, and she never forms an effectual higher-order desire to that effect. Finally, imagine that Irene's world is nomically deterministic. Her action is entailed by the laws of nature and some event in the past. But she also is not manipulated or pathological.

The pressing question for a source incompatibilist like Stump is this: Is $Irene_2$ morally blameworthy for what she did? As a proponent of the direct argument—the

[101] Widerker, "Farewell to the Direct Argument," 323 (fn. 17).

argument that causal determinism is directly incompatible with moral responsibility—Stump *must* answer this question negatively, because no agent is responsible for causally determined actions. Opponents (such as traditional compatibilists who affirm a conditional interpretation of could have done otherwise) will argue that *Irene*$_2$ *is* blameworthy, since she knew that she was acting wrongly and *could* have avoided acting as she did (if she wanted to), while hard compatibilists can claim that she is responsible because her volitional and evaluative faculties were not compromised in any freedom-undermining way. It seems clear that the only response Stump can make is that in the deterministic world under consideration, *Irene*$_2$ is not responsible because she could *not* have avoided voting with the majority.[102] In other words, it is not because her cognitive or volitional faculties are compromised in a responsibility-undermining way, or because she could have chosen to do otherwise had she really wanted to do otherwise, but because her world is nomically deterministic and therefore she lacks alternative possibilities that *Irene*$_2$ is not morally responsible. The result is that the plausibility of the direct argument "depends on the traditional assumption that determinism rules out avoidability."[103] Therefore the direct argument can only be successfully used in conjunction with the indirect argument, resulting in the severe weakening of the source incompatibilist's strongest argument.

[102] The structure of this argument is very similar to the argument that Widerker makes in "Farewell to the Direct Argument," 322.

[103] Widerker, 317.

Leeway Incompatibilism

If the critique of Stump's source incompatibilism is correct, then she must admit a form of *leeway incompatibilism*. Source incompatibilists, like Stump, and leeway incompatibilists agree that "causal determinism would be sufficient for the lack of moral responsibility." However, leeway incompatibilists also claim that moral responsibility requires alternative possibilities.[104] Timpe persuasively argues that not just Stump, but all source incompatibilists, should admit that alternative possibilities are necessary for ascribing moral responsibility. A number of (source) incompatibilists argue that alternative possibilities are not necessary for ascribing moral responsibility, despite affirming that causal determinism would rule out moral responsibility. Rather, they agree with Stump that what moral responsibility requires is that the agent be the ultimate source of the action. Also like Stump, source incompatibilists tend to accept the efficacy of *FSCs*. They posit that in a typical *FSC*, the agent, not the intervener, produces the action. Since the agent is the ultimate source of the action, she can be held responsible despite lacking alternate possibilities.[105]

A number of leeway incompatibilists agree that the source requirement is also necessary for ascribing moral responsibility. In fact some, for instance Robert Kane[106] and Derk Pereboom,[107] argue that although alternative possibilities are necessary to ascribe moral responsibility, the ultimacy or source condition is in some sense the more

---

[104] Timpe, "Source Incompatibilism and its Alternatives," 142.

[105] Clarke, Randolph, *Libertarian Accounts of Free Will* (New York: Oxford University Press, 2003), 11. Quoted in Timpe, "Source Incompatibilism and its Alternatives," 142-43.

[106] See Kane's *The Significance of Free Will.*

[107] See Pereboom, Derk, *Living Without Free Will*, (Cambridge, U.K. ; New York : Cambridge University Press, 2001).

important or fundamental issue.[108] For this reason, Timpe divides up the libertarians who affirm the fundamental importance of the source condition into Wide and Narrow Source incompatibilists. Narrow Source incompatibilists think that "the proper kind of source (necessary for moral responsibility) doesn't require alternative possibilities at all."[109] Wide Source Incompatibilists think that, although the source requirement is more fundamental, alternative possibilities are necessary to ascribe moral responsibility. Timpe argues that Wide Source Incompatibilism is the more plausible position, because Narrow Source Incompatibilists agree that in order to be morally responsible, the agent must be the ultimate source of her actions, and that this is incompatible with determinism. But what makes an agent the source of her action in a way that is incompatible with causal determinism? With Stump, it seems reasonable to conclude that an action for which the agent is the ultimate cause is an action that was one possibility among others (at least in a typical *FSC*). In these instances, the agent's causing her own action is different from another person's ensuring that the agent act in that way through stimulating her brain. As Timothy O'Connor puts it, "the significance of such alternatives (whether they are robust or mere 'buds') lies in their being indicators of the self-determination manifested by one's actions, which is necessary for responsibility."[110] That is to say that "the presence of any alternative possibilities is a sufficient condition

---

[108] Timpe, "Source Incompatibilism and its Alternatives," 152.

[109] Ibid., 153.

[110] O'Connor, Timothy, "Freedom With a Human Face," *Midwest Studies in Philosophy* 29 (2005), 209ff. Quoted in Timpe, "Source Incompatibilism and its Alternatives," 154.

for the falsity of causal determinism, which according to incompatibilists of all stripes is itself a necessary condition for moral responsibility."[111]

Timpe's basic argument proceeds as follows: both Narrow and Wide Source Incompatibilists agree that nothing about the agent's past (that the agent did not have a role in creating) nor the laws of nature can determine an action for which the agent is morally responsible. For the agent to be morally responsible, the agent must be the "difference-maker." In other words, the reason that this particular action happened instead of a different action traces back to something that the agent did on purpose that is not the direct result of the laws of nature and the events of the past. Again, both Narrow and Wide Source Incompatibilists would agree with this point. However, for the agent to be the difference-maker, "there must be more than one future that is compossible with the conjunction of the past (or those parts of the past that were not themselves determined by the agent) and the laws of nature."[112] More than one path that the agent can actually take must be available, given the truth about the past and the laws of nature. Thus, the agent must have alternative possibilities. This means that the ultimacy condition that Narrow Source compatibilists insist on entails an alternative possibilities condition; without alternative possibilities the agent cannot fulfill the ultimacy condition for moral responsibility.[113] It is important to note here that Timpe does not think that the alternative possibility has to be "robust." The alternative possibility merely has to be present because the action must both be undetermined and compossible with the past and the laws of nature. Thus more than one option must be available, however small.

---

[111] Timpe, "Source Incompatibilism and its Alternatives," 155.

[112] Ibid., 157.

[113] Ibid., 158.

Conclusion

Taken together, these arguments show sufficiently that Stump's "modified" or "Frankfurt-style" libertarianism is misguided. She insists that an agent must be the source of the action to be morally responsible. She also insists that in order to be the source of her action, an agent's action must be undetermined; therefore, she must admit that more than one action is compossible with the past and the laws of nature—that is, there is at least a weak alternative possibility condition for moral responsibility.

This critique does not mean that all is lost for Frankfurtians. One can accept Frankfurt's hierarchical structure, as Stump does, with certain modifications. First, an emphasis needs to be placed on the history of the agent's moral development, as Fischer argues. Second, for an agent to be responsible, she must be the ultimate source of her actions, as Stump and other source incompatibilists argue. For the agent to be the source of her actions, she also must have access to at least some morally relevant alternative possibilities, adding a weak alternative possibility (or strong control) condition for moral responsibility. Finally, both Stump and Fischer are right that ascriptions of moral responsibility must be grounded in an epistemic condition or knowledge component. Kierkegaard has all of these elements in his notion of freedom and moral responsibility. These elements allow him to address a major issue that Frankfurt cannot. Namely, Kierkegaard can speak to how an agent can cultivate her moral character. Frankfurt has no way to address this issue. Thus, I turn to Kierkegaard's understanding of the freedom relevant conditions necessary for grounding moral responsibility as well as his understanding of moral self-cultivation.

CHAPTER FIVE

A Kierkegaardian Account of the Freedom-Relevant Conditions Necessary for Morally Responsible Agency

Introduction and Structure of the Chapter

In this chapter I argue that Kierkegaard's picture of human agency and the will provides the necessary freedom-relevant responsibility-grounding conditions that Frankfurt' account (and all the preceding accounts) lack. Kierkegaard has a hierarchical view of the will that is in some respects similar to Frankfurt's, but with a (modest) knowledge component (what I call his motivational cognitivism) that both Fischer and Stump agree is essential in order for the agent to be considered the source of her actions. Kierkegaard agrees with Fischer and others that the causal history of the action is essential for differentiating free and responsible action from action resulting from manipulation. Kierkegaard believes that any account of the ethical life—the life of a person who takes responsibility for the formation of her character—must be *historical* in nature. As I will show in examples from Kierkegaard's pseudonym Judge William, the aesthete A is building a history through his choices whether he is aware of it or not. Judge William argues that A's choices form his character and solidify his self, which in turn contributes to his future actions. Though A believes that he is subject to necessity and therefore refrains from choosing, Judge William will not let A off the hook. Instead the Judge deems A responsible for making character-forming choices even as he refuses, in a sense, to choose. Therefore Kierkegaard articulates a view of freedom and

responsibility that is hierarchical *and* historical. Frankfurt's early account of free will and moral responsibility includes an ambiguity as to whether and in what way an historical element is necessary. Later comments about a manipulator providing the agent with an entirely new character and that agent being responsible for the actions that stem from that new character clearly indicate that Frankfurt believes that an agent can be morally responsible *without* any historical element.[1] However, his notion of care and love requires that the agent maintain her cares over time, and for Frankfurt love is one of the highest expressions of human freedom. I argue that in this one fundamental area where Frankfurt needs the correction that Kierkegaard can offer, and that Kierkegaard's combination of a hierarchical account of the will and an historical account of agency and responsibility is uniquely suited to this task.

Another similarity between Kierkegaard and Frankfurt is that they "[b]oth are skeptical towards the Kantian idea of founding morality in the laws of practical reason. They both deny that the concerns, which shape our lives, could simply be validated by subject independent values."[2] In other words, both believe that the agent's cares and concerns play a fundamental role in practical reason and normativity, and that nothing is normative for the agent *independent* of that agent's cares and concerns. Furthermore, they agree that "[c]hoice is occasioned. It requires that there are alternatives for someone and thus that these are seen as being alternatives. Such a perception requires a background of care for activities and relationships within and between which

---

[1] See Frankfurt's "Response to John Martin Fischer" in *Contours of Agency*, 27-31.

[2] Pillar, Christian, "Morality's Place: Kierkegaard and Frankfurt," *Revista Porteguesa de Filosofia 64* (2008), 1208.

'alternatives' could arise."[3] The self is not able to create itself from nothing by making choices that are completely abstracted from concrete reality or the agent's facticity. Free and responsible choices are made within a framework of beliefs, cares, concerns, etc.

Kierkegaard is distinguishable from Frankfurt by his commitment to a kind of libertarianism. A close reading of certain passages in Kierkegaard's writing reveals that he is committed to an alternative possibilities condition as necessary for moral responsibility. However, even Kierkegaard's libertarian commitments reveal important commonalities with Frankfurt. Like Frankfurt, Kierkegaard rejects the implausible strong libertarian position that "the will is by its nature so free that it can never be constrained."[4] Furthermore, both thinkers hold to a "voluntaristic account of practical normativity."[5] Frankfurt and Kierkegaard agree that in the final analysis morally responsible decisions stem primarily from the will and not from the intellect.[6] Whereas Frankfurt is a strong (almost irrational) voluntarist who believes that "[t]here can be no rationally warranted criteria for establishing anything as inherently important"[7] and "…it is possible to ground judgments of importance (normative judgments) only in judgments concerning what people care about,"[8] Kierkegaard is less voluntaristic, and allows an important, albeit not

---

[3] Weston, Michael, "Kierkegaard: The Literature of Freedom," in *Kierkegaard and Freedom*, ed. James Giles, 175-75.

[4] Descartes, *Passions of the Soul*, I, art. 41. Quoted in Frankfurt, "Concerning Freedom and the Limits of the Will, 71.

[5] Frankfurt, *Taking Ourselves Seriously and Getting it Right*," 105, footnote 1.

[6] Note that this is related to but not the same as Pojman's account of volitionalism in his *Logic of Subjectivity*. I also reject Pojman's analysis of Kierkegaard as a direct volitionalist who thinks that "the action by which a belief is formed as a basic action which can simply be willed." However, the formation of beliefs is not a primary aspect of the freedom-relevant responsibility-grounding conditions.

[7] Frankfurt, *Taking Ourselves Seriously*, 22.

[8] Ibid., 23.

primary, role to the intellect in matters of practical normativity.[9] Kierkegaard's position, roughly, is that objective ethical standards exist and can be known by an agent; but, these standards need *subjective appropriation* to become fully comprehended by and thus *motivating* for the agent.[10] This chastened voluntarism has a weak knowledge component locating Kierkegaard between Stump's intellectualism and Frankfurt's strong voluntarism.

Furthermore, Kierkegaard differs from Frankfurt in that he is what Timpe calls a "Wide Source Incompatibilist."[11] As a wide source incompatibilist, Kierkegaard agrees with Stump (a *narrow* source incompatibilist) that morally responsible actions require that the agent be the source of her actions,[12] and that this agential "sourcehood" requirement is incompatible with determinism. His view also accords with Timpe and other wide source incompatibilists that even though the source requirement is the primary freedom-relevant responsibility grounding condition, alternative possibilities are also a necessary part of the "grounding mix" because they sufficiently demonstrate that

---

[9] Though Frankfurt disagrees with Descartes' statement about the *limitlessness* of the will (see Frankfurt's "Concerning the Freedom and Limits of the Will'), he appears to agree with Scotus' insistence "that the will itself, and nothing but the will, is the total cause of its volitions. It is not determined by another, but determines itself *contingenter*, not *inevitabiliter*, to one of the alternatives that are before it." (II Sent., dist. Xxv). Kierkegaard's voluntarism is in some respects similar to Scotus's, but Kierkegaard believes that the agent can intentionally come to see (perceive) the world in new ways, that these ways of seeing can produce new volitions, and that the agent plays an important causal role in both perceiving and willing and therefore is responsible in part due to this causal role. Since perceiving is a capacity of the intellect, Kierkegaard retains an important place for the intellect that Frankfurt's denies.

[10] Davenport, "Entangled Freedom: Ethical Authority, Original Sin, and Choice in Kierkegaard's *Concept of Anxiety*," *Kierkegaardiana* 21 (2001), 141.

[11] Kevin Timpe coins this term in his *Source Incompatibilism and its Alternatives.*

[12] As Johannes Climacus says in the *Postscript*,"…this letting go, even that is surely something; it is, after all, *meine Zuthat* (my contribution)." Quoted in Ferreira, M. Jamie, "Faith and the Kierkegaardian Leap." *Cambridge companion to Kierkegaard*, ed. Alastair Hannay and Gordon Daniel Marino, (Cambridge: Cambridge University Press, 1998), 210.

determinism is false and that the agent, and not some external causal factor,[13] is the ultimate cause of the action.[14] The agent must be the difference-maker when she acts in morally relevant ways and as such her actions cannot be the result of some immanental process where the transitions are logical and necessary, because this immanental process constitutes a responsibility-undermining external force.

Additionally, I argue that Frankfurt's provocative account of volitional necessity or wholeheartedness makes sense only in a libertarian framework such as Kierkegaard's. Something akin to Davenport's "aretaic commitment"[15] is a better term than Frankfurt's own "volitional necessity" for what Frankfurt describes. He makes contradictory claims about volitional necessity. On the one hand he claims that a person may come to care about something so much that "it is impossible for him to forbear from a certain course of action,"[16] which indicates that if a person has certain cares or commitments, it is impossible for him *not* to act on them. On the other hand, he claims that "[w]e can sometimes take steps that inhibit us from loving, or steps that stimulate us to love; more or less effective precautions and therapies may be available, by means of which a person can influence whether love develops or whether it lasts."[17] Statements like this indicate that an agent's loves or cares are revisable and do not *necessitate* that the agent act in a certain way. Rather, an agent can change her cares intentionally but not through a simple

---

[13] The external causal factor can be internal to the agent but outside of the agent's will or ultimate control.

[14] The inclusion of alternative possibilities is what makes a source incompatibilist wide instead of narrow.

[15] Davenport, *Will As Commitment And Resolve,* esp. 472-86.

[16] Frankfurt, "The Importance of What We Care About," 86.

[17] Frankfurt, *Taking Ourselves Seriously*, 41.

act of will. The fact that a love or care can be changed—but not simply or immediately—indicates that when Frankfurt talks about volitional necessity he is really talking about something more akin to deeply engrained character traits—part of the agent's facticity—that drastically limit without eliminating that agent's freedom of will.[18] Character traits are formed over time through repeated actions that stem from the agent and over which the agent must maintain a level of control. This indicates that an historical element is a necessary component of a properly constructed theory of care, and furthermore that caring makes the most sense within a libertarian framework.

Chapter Overview

This chapter is divided into two main sections. In the first section I explicate Kierkegaard's understanding of free will. Specifically I refer to the freedom that he believes is necessary for the cultivation of morally responsible agency, or what he simply calls the "self." I argue that two of Frankfurt's main contributions to the contemporary debates concerning free will and moral responsibility—his hierarchical view of the will and his notion of care—are useful rubrics for interpreting Kierkegaard's notorious theory of the "leap." I argue that we should understand ethical and religious conversions or "leaps" as a function of the agent's *care* rather than as either arbitrary choice or necessary progression. Furthermore, I argue that this way of interpreting the leap defuses the charges that Kierkegaard is an irrationalist,[19] a direct volitionalist,[20] or an unmitigated

[18] Freedom of the will classically defined as alternative possibilities.

[19] See MacIntyre's *After Virtue*, esp. 42, 47.

[20] See Pojman's *The Logic of Subjectivity*, esp. 146.

voluntarist.[21] Kierkegaard's libertarian position is nuanced and modest, leaves an important role for reason, and stresses the limits of the will.

In the final section, I deal with one of the fundamental problems of Frankfurt's position and with semi-compatibilism generally: Frankfurt's lack of an account of moral self-cultivation.[22] This lack is problematic due to Frankfurt's insistence that the highest expression of human free will is wholehearted commitment or love, and these commitments are made up of a cluster of mental states such as desires, emotions, and what Frankfurt at one point calls "irresistible passions."[23] However, Frankfurt does not explain how to cultivate the various mental states that constitute wholeheartedness because he does not believe that agents play a *causal* role in the development of the characteristics that give rise to various mental states. He leaves the reader to wonder if wholeheartedness is simply a matter of moral luck or if moral self-cultivation is even possible. I argue that emotion and desire control is fundamentally important to reach wholeheartedness. Frankfurt's account has no place for desire and emotion control due to his essentially noncognitive view of desires and emotions. Frankfurt ultimately rejects any notion of moral self-cultivation due to his inadequate understanding of character formation and his rejection of a strong control (or alternative possibilities) condition. A reader might come to the conclusion that this problem is inherent in Frankfurt's hierarchical or structuralist view of the will, and that all hierarchical models are incompatible with a robust notion of moral self-cultivation. This conclusion is mistaken.

---

[21] The irrationalist and voluntarist charges are often conflated as in MacIntyre's critique of Kierkegaard.

[22] The connected issue of volitional necessity will also be dealt with in this section.

[23] Frankfurt, "The Importance of What We Care About," 86.

Kierkegaard has a hierarchical view of the will similar to Frankfurt's, yet Kierkegaard also has a robust notion of moral self-cultivation. These two views are brought together in Kierkegaard's *Christian Discourses*, a work designed in part to help the reader alter her emotions and thus to cultivate her character.[24] After exploring Frankfurt's hierarchical view of the will and his subsequent rejection of moral self-cultivation, I argue that Kierkegaard's emotion-therapy is consistent with Frankfurt's hierarchical view of the will and is a plausible explanation for how an agent might cultivate wholehearted loves. Therefore, Kierkegaard provides a more comprehensive understanding of free will and moral responsibility, specifically the freedom-relevant conditions necessary for grounding morally responsibility agency. One of these necessary freedom-relevant conditions is that the agent must be able to participate actively and intentionally in the cultivation of her character. A theory that leaves out some account of this condition fails to provide an adequate account of free will and moral responsibility.

Kierkegaard's Modest Libertarianism

In this section, I cover the first of the three main areas of Kierkegaard's understanding of the freedom-necessary conditions necessary for grounding responsibility and the cultivation of character. Kierkegaard is an incompatibilist with a strong control component, a hierarchical source component which includes a weak knowledge component, and an historical understanding of human agency. Since Kierkegaard is not a contemporary analytic philosopher, and writes texts that cover a

[24] Kierkegaard does not discuss this strategy in the discourses themselves, because they are, after all, discourses (similar to sermons), and not philosophical treatises. Kierkegaard is trying to engender a certain existential response in his reader, not explain how this response is supposed to happen.

wide variety of philosophical and non-philosophical topics, I first deal with recent reconsiderations of his notion of the "leap" to lay the conceptual groundwork for the textual analyses. Then I turn to examples from three texts where these components are interconnected in various ways. *The Concept of Anxiety* explores the nature of sin and argues that sin—or moral responsibility—presupposes alternative possibilities as well as the historical nature of human action. *Either/Or II* demonstrates Kierkegaard's belief that responsibility for selfhood requires the agent to identify with her volitional and motivation states, and thus promotes a hierarchical view. However, *EO II* also stresses the historical nature of responsibility and selfhood. Perhaps most importantly, *EO II* reveals Kierkegaard's commitment that reasons play an important role in major ethical decisions: through the appreciation of reasons, the agent can come to identify with new volitional and motivational states. Finally, I turn to *Christian Discourses*, which once again reveals Kierkegaard's hierarchical understanding of the will. Most importantly, *CD* emphasizes the role that the agent's cognitive faculty plays in free and responsible action. Kierkegaard's analysis of the cognitive nature of desires and emotions connects those mental states with deeper motivations such as cares and concerns, demonstrating that an agent's thoughts shape her desires and emotions. Insofar as some of those thoughts can be freely and intentionally willed, this shows that an important freedom-relevant component of moral responsibility is the cognitive nature of desires and emotions. Take together, the components mined from these texts yield a truly unique and satisfying response to the meta-ethical question at hand, as well as the ways that Kierkegaard's understanding of these issues—though similar in certain respects to

Frankfurt's—can deal with the objections to Frankfurt as well as the objections to Fischer and Stump.

Kierkegaard's 'Leap' Reconsidered[25]

Kierkegaard's strong control component is understood best though his concept of the "leap." He distinguishes between *quantitative* and *qualitative* movements or transitions (qualitative transitions are often referred to simply as leaps). Simply put, a *quantitative* transition is a moral or religious decision that results from a cumulative or necessary process, whereas a *qualitative* transition or 'leap' is a movement that happens as an act of freedom that can never be explained *fully* by reference to prior causal states. In the words of Climacus, "it is, after all, *meine Zuthat* (my contribution)"[26] which points to Kierkegaard's source condition. Kierkegaard understands the leap[27] as a choice that the individual makes that is neither necessary (due to responsibility-undermining external forces) nor arbitrary.

Kierkegaard developed his "theory of the leap"[28] in response to two main ideas. The first is the Hegelian understanding of transition and movement. Hegel posited that the historical progression of Spirit is *necessary*. Referring to Hegel, the pseudonym Vigilius Haufniensis says that "[e]very movement, if for the moment one wishes to use

---

[25] The following interpretation of Kierkegaard's theory of the leap and his understanding of morally responsible action is not the traditional view, but rather has emerged in recent scholarship.

[26] Kierkegaard, *Philosophical Fragments and Johannes Climacus*, edited and translated by Howard V. Hong and Edna H. Hong (Princeton, N.J.: Princeton University Press, 1983), 43.

[27] From here on, unless otherwise noted "leap" will always refer to qualitative leap.

[28] *JP I* 110. Quoted in Ferreira, "Faith and the Kierkegaardian Leap," in *Cambridge Companion to Kierkegaard*, edited by Alastair Hannay and Gordon Daniel Marino (Cambridge: Cambridge University Press, 1998), 207.

this expression, is an immanent movement, which in a profound sense is no movement at all."[29] Kierkegaard's critique here is that if a movement must happen—either logically or due to causal antecedents—then it is not a movement in the ethical sense. An ethical movement "is itself a transcendence that has no place in logic."[30] For movement to be ethical, it must involve elements that transcend logical or causal necessity: "If ethics has no other transcendence, it is essentially logic."[31] Kierkegaard believes that the defining mark of human agency is the ability to perform acts of both good and evil freely. For an act to be considered under ethical categories (under the description good or evil) —and for the agent to be considered morally responsible—the agent must have a robust freedom that transcends necessity. Therefore Kierkegaard believes that ethical action requires a notion of free movement or transition that is not found in Hegel. This notion requires that alternative possibilities are available to the agent either at the moment of the ethically laden choice or during some prior choice that has contributed to the agent's entrenched characteristics which now produce the ethical choice.[32]

Kierkegaard also formulates his notion of the leap in response to the traditional (Western) Christian notion of original sin,[33] an idea with fundamental similarities to

---

[29] *The Concept of Anxiety*, p. 13. I am following Michele Kosch's lead here and taking *The Concept of Anxiety* and *The Sickness unto Death* as essentially Kierkegaard's own voice that were penned as pseudonyms for small reasons and with minor changes. See Kosch's *Freedom and Reason*, 10-13.

[30] *Ibid.*

[31] *Ibid.*

[32] This points to the existence of a notion of "tracing" in Kierkegaard where either there are alternative possibilities available at the moment of choice or we can trace back in the agent's history to a point where those alternative possibilities were available and played a causal role in shaping the characteristics that now produce the choice.

[33] James Giles states in his excellent article "Kierkegaard's Leap: Anxiety and Freedom" that sin is a complex concept made up of several simpler concepts, acting in a way that is ethically wrong, acting is a way that violates a religious or theistic norm, and acting freely when violating those ethical and religious

Hegel's notion of necessary transitions. Kierkegaard argues that the traditional understanding of original sin is that the entire human race is morally responsible for an action that only one human being performed. According to his interpretation of this traditional view, every member of the human race is sinful (suffers the consequences of the act of sin) even though only one member of the race committed the act of sinning that brought the consequences. Kierkegaard argues that this view is mistaken, and that a correct explanation of sin (or morally responsible action) must be able to explain simultaneously how every single person, including Adam, becomes sinful.

Kierkegaard's theory of the leap is a part of his larger commitment to what Michele Kosch calls his "double incompatibilism."[34] Kierkegaard is committed to the traditional libertarian claim that the freedom necessary for moral responsibility is incompatible with causal determinism (the first incompatibilism).[35] However, he also is committed to the claim that the freedom necessary for moral responsibility is incompatible with the ethics of autonomy as self-determination, i.e., the idea that once the agent accepts herself as a self-determined chooser (the first incompatibilism) she necessarily will know what to choose and therefore can be responsible only for good (or morally praiseworthy) choices.[36] Kierkegaard's understanding of human agency is that agents can make choices that are both free from sufficient causal antecedents and can

---

norms. (see Giles, 69-70). He never says as much, but I assume that he thinks that Kierkegaard's concept of sin includes all three components. I think that this view is mistaken, and that all that need happen for sin to occur on Kierkegaard's view is the first and third elements. An agent can sin simply by failing to fulfill his ethical obligation to act freely and responsibility, and this is always a free act for Kierkegaard because the essence of human agency is freedom, and this essence can never be fully eradicated.

[34] Kosch, *Freedom and Reason*, 139-78.

[35] See Kosch, *Freedom and Reason*, 150.

[36] Ibid., 172.

choose both good and evil freely and intentionally. Any theory that denies human agency these abilities will result in an inadequate understanding of the freedom relevant conditions necessary for grounding morally responsible agency.

Much of the best recent scholarship seeking to understand the nature of Kierkegaard's concept of the leap responds to a small but important section of MacIntyre's *After Virtue*. MacIntyre argues that Kierkegaard writes *Either/Or* in several pseudonyms so that he can distance himself from the text and "present the reader with an ultimate choice, himself unable to commend one alternative rather than another…"[37] The choice that Kierkegaard presents is between the aesthetic and ethical ways of life, and MacIntyre thinks that if someone confronts this choice without a prior commitment or inclination, she "can be offered no *reason* for preferring one [way of life] to the other."[38] MacIntyre goes on the say that

> if a given reason offers support for the ethical way of life—to live in that way will serve the demands of duty *or* to live in that way will be to accept moral perfection as a goal and so give a certain kind of meaning to one's actions—the person who has not yet embraced either the ethical or the aesthetic **still has to choose whether or not to treat this reason as having any force**. If it already has force for him, he has already chosen the ethical; which *ex hypothesi* he has not. And so it is also with reasons supportive of the aesthetic.

In other words, MacIntyre understands Kierkegaard as presenting his reader with a radical, criterionless choice.[39] Davenport argues that MacIntyre puts forth a dilemma: either rational evaluation *determines* the will's choice or practical reason is

---

[37] MacIntyre, Alasdair C., *After Virtue: A Study in Moral Theory*. Notre Dame, Ind: University of Notre Dame Press (1984), 40.

[38] Ibid. [Bolded type mine]

[39] Evans, C. Stephen, "Where There's a Will There's a Way: Kierkegaard's Theory of Action." *Kierkegaard on Faith and the Self Collected Essays*. Provost series. Waco: Baylor University Press (2006), 312.

motivationally inert for the agent.[40] Furthermore, MacIntyre seems to argue that if one holds the position that for a decision to be a free and responsible one, practical reasons cannot determine a decision, then this decision will be irrational and arbitrary.[41] Kierkegaard does agree that if reasons are motivating to the degree that those reasons *determine* the decision, then the agent is not acting freely. MacIntyre appears to assume that the only other option is that the decision is arbitrary. He views Kierkegaard's position as the choice between various ways of life that has no basis in reason but only in the will, so in the end the reader must simply choose. However, Kierkegaard does not think that the leap or the choice between various ways of life is arbitrary, but that "the objective significance of moral considerations…*grounds* the ultimate choice…without *determining* this choice."[42]

MacIntyre clarifies his position in response to criticisms of his interpretation of Kierkegaard by stating that the Judge does offer A good reasons "to move from the aesthetic to the ethical and not merely reasons-from-the-standpoint-of-the-ethical."[43] Furthermore, MacIntyre does think that A can *understand* those reasons. However, MacIntyre maintains that on Kierkegaard's account "[o]ne has to have already chosen

---

[40] Davenport, John. "Kierkegaard, Anxiety, and the Will," p.165. Evans makes a very similar point in "Where There's a Will There's a Way," 313.

[41] In other words, one must either be an intellectualist (like Stump for example) or a strong voluntarist like Frankfurt. In am following recent scholarship in locating Kierkegaard in the middle of these two positions as a weak voluntarist.

[42] Davenport, John. "Kierkegaard, Anxiety, and the Will," 163.

[43] MacIntyre, "Once More on Kierkegaard," *Kierkegaard After MacIntyre*, 44. MacIntyre's criticism of Kierkegaard and the response by Davenport and others illuminates the fundamental issues surrounding Kierkegaard's understanding of free will and moral responsibility. MacIntyre is more concerned about the source component and its ramifications for moral responsibility.

oneself as an ethical subject" in order to *appreciate* those reasons.[44] MacIntyre concludes that

> although Kierkegaard does indeed understand human lives as having a telos, and although Kierkegaard does indeed believe that subsequently individuals may come to recognize that there were good reasons for them to move towards that telos out of the aesthetic and into the ethical, *at the time that they did so move theirs was not a progress directed or even guided by reason*, but rather a set of psychological developments. Their story, as I have interpreted it, is a story of the *passions* rather than of *reasons*.[45]

MacIntyre disagrees with Kierkegaard's conception of the will and the role that reason plays in the volitional lives of agents. As Davenport states, Kierkegaard holds that "a choice to posit new ends for the self can be performed *on the basis of* or *on the grounds of* recognized practical reasons, without these having antecedently operated as motives or as objects of appetitive states in the agent."[46] Needless to say, either MacIntyre does not understand this middle ground between intellectualism and strong voluntarism or he finds Kierkegaard's position untenable.

The dilemma that MacIntyre presents—the decision to adopt one way of life over another is either determined by reasons or it is arbitrary—can be understood in two basic ways. On the one hand it is the dilemma between intellectualism and strong voluntarism. On the other hand it can indicate that Kierkegaard must either be what John Robertson calls a "robust moral internalist" or a moral externalist. Robertson defines a robust moral internalist as one who "…holds that moral considerations are intrinsically motivating, guaranteed by their content to affect the motivation of anyone who is rational

---

[44] Ibid.

[45] Ibid., 345. My emphasis.

[46] Davenport, "Kierkegaard, Anxiety, and the Will," 166.

and aware of them." Furthermore, "reflection on what one has reason to do is independent of, and gives direction and content to, one's motivations if one is rational."[47] In this case, Judge William simply has to give A good enough reasons in support of the ethical life and so long as A is a rational person, those reasons will be motivating and he will become an ethical person. Clearly Kierkegaard does not think that reasons alone—even understood and appreciated reasons—will necessarily motivate. Despite the fact that the Judge does present reasons and arguments for the superiority of the ethical life, we are never told whether 'A' makes an ethical leap. MacIntyre is right to see that Kierkegaard is pushing his reader to choose between two ways of life. It would be impossible to reconcile *Either/Or* with this kind of internalism.

If Kierkegaard is clearly rejecting a robust moral internalism, then perhaps he accepts a form or moral externalism where moral reasons (considerations that give normative or justifying reasons) have no conceptual tie to rational motivation.[48] While Kierkegaard accepts a form of moral externalism, he denies the strong version. Davenport interprets Kierkegaard as opposing this view and believes that Kierkegaard's leap should be understood as both "libertarian and yet dispositionally directed."[49] Davenport rejects MacIntyre's dilemma that either reasons *determine* a choice or the choice is *arbitrary* in the sense that practical reason cannot provide any motivation.[50] He argues that Kierkegaard's understanding of the leap is that through a significant moral

---

[47] Robertson, John, "Internalism, Practical Reason, and Motivation," in *Varieties of Practical Reasoning*, ed. Millgram, Elijah. Cambridge, Mass: MIT Press (2001), 130.

[48] Ibid., 132.

[49] Davenport, "Kierkegaard, Anxiety, and the Will," 163.

[50] Ibid. Evans responds to MacIntyre in a similar way in his "Kierkegaard's Theory of Action."

choice, the self gives "full subjective force" to the reasons that it could appreciate *objectively* prior to that choice.[51] In other words, a person's history matters as the self cannot uproot itself completely from prior choices; yet, the self retains the ability to choose new and different ends in part by coming to understand the legitimacy of an alternate life-view.[52] In this way, the agent is given *potential* motivation for the adoption of new ideals, but the agent must actualize this potential by means of the agent's decision. This decision subjectively appropriates the objective moral consideration into the agent's motivational structure. This allows Davenport to locate the "leap" in the space between what Evans terms "intellectualism" (MacIntyre's rationalism) and what I have labeled strong voluntarism (or libertarianism).

Case 1—The Concept of Anxiety

The Concept of Anxiety presents the best articulation of Kierkegaard's strong control responsibility grounding component. In this text we see Kierkegaard's belief that to be morally responsible, an agent must be the source of her actions, and that to be the source of her actions she must be able to avoid performing that act. Kierkegaard thinks an unavoidable act cannot at the same time be a morally responsible act. Furthermore, not only does *CA* reveal Kierkegaard's commitment to *PAP*, it also shows that free and responsible agency has a historical aspect. Kierkegaard believes that free and responsible actions affect the agent and shape her character. Finally, in *CA* we see indications of

[51] Ibid.,166.

[52] This is begging the question as to how this is possible, and I think that Ferreira's interpretation helps. It should be noted that coming to understand a new life-view as legitimate does not automatically mean that this life-view will be adopted, but that this is often a helpful step on the way to the adoption of a new life-view.

Kierkegaard's belief that higher-order mental states are a fundamental aspect of both human agency in general and free and responsible action in particular.

In *The Concept of Anxiety* Kierkegaard examines various traditional Christian ways of understanding the idea that through Adam's sin, sinfulness was brought into the world and subsequently affected all of humanity. He argues that the traditional answer provides one of two explanations: the first explains how the first human being sinned; the second explains why every other human being is sinful. Kierkegaard argues that there is no explanation that makes coherent sense of both of these premises, and that a third kind of explanation is needed. He discusses the difference between these two modes of explanation by asking "[d]oes the concept of hereditary sin differ from the concept of the first sin in such a way that the particular individual participates in inherited sin only through his relation to Adam and not through his primitive relation to sin?"[53] In other words, if the concept of hereditary sin (the second explanation) means that subsequent human beings are sinful simply in virtue of their relationship to Adam, then they are not sinful as a result of their own sinful acts but because of an ontological qualification of the human substance as a result of Adam's sin.[54] Kierkegaard answers that if human beings are sinful only because Adam sinned, then "Adam is the only one in whom it [hereditary sin] was *not* found, since it came into being through him."[55] This is the second of the aforementioned methods of explaining original sin: every subsequent human being is sinful because Adam's sin changed human nature so that now every human being is

---

[53] *CA*, 26.

[54] *CA*, 230, ft.1. This is Reider Thomte's phrase from the notes to his translation.

[55] *CA*, 26. My emphasis.

sinful simply by virtue of being a member of the human race (with the result that Adam is not a sinner like the rest of the human race). But these explanations are two sides of the same coin. If Adam's sin changed the nature of the human race so that all subsequent humans begin life in a predicament different from Adam's, then "Adam is placed fantastically outside of history." Adam no longer has the same nature as the rest of humanity, and his actions do not hold the same consequences for him that they do for the rest of the human race.[56] This difference is a problem in part because it means that Adam is not affected by his own choices like the rest of the human race. Therefore Kierkegaard concludes that "no explanation that explains Adam but not hereditary sin, or explains hereditary sin but not Adam, is of any help."[57]

Kierkegaard argues that the right kind of explanation simultaneously will make sense of every single human being's sin as well as the sin of Adam: "if I can explain guilt in a subsequent person, I can explain it in Adam as well."[58] This explanatory aspect of human agency is *anxiety*. Although anxiety has a level of explanatory power it is not fully analyzable due to its very nature, and it does not *cause* human beings to sin. Kierkegaard thinks that anxiety is an aspect of human agency that makes it *possible* for the agent to will either good or evil without making either outcome *necessary*. "We have nowhere been guilty of the foolishness that holds that man *must* sin…[instead] sin presupposes itself, just as freedom presupposes itself, and sin cannot be explained by anything antecedent to it, any more than can freedom."[59] Kierkegaard holds that if sin

---

[56] The importance of this point is revealed momentarily.

[57] *CA*, 28.

[58] Ibid., 38.

[59] Ibid., 112.

were necessary then guilt and responsibility would not be possible. His conception of sin presupposes avoidability: to sin is to violate some ethical or religious norm *freely*.[60] Consider this passage from Kierkegaard's journals:

> *That "Hereditary Sin" is "Guilt"* is a real paradox. How paradoxical is best seen as follows. The paradox is formed by a composite of qualitative heterogeneous categories. "Hereditary" is a category of nature. "Guilt" is an ethical category of spirit. How can it ever occur to anyone to put these two together, the understanding says—to say something is hereditary which is by its very concept cannot be hereditary.[61]

If anxiety were a quality of human nature that causes sin, then sin could not be an ethical category. Kierkegaard assumes that anything that happens as a result of a thing's nature is necessary and therefore not something for which the agent is responsible. So the action that acquires guilt is not the result of any ultimate external forces and it is, to some extent, inexplicable. Any explanation of the way that guilt is acquired is problematic because if the free acquisition is *fully* explainable, then it is possible to argue that those explanatory factors determine the decision. Kierkegaard denies that the leap can be reconstructed psychologically because his conception of freedom requires that motivations can help *clarify* action, but cannot *explain* action and most importantly do not determine the will.[62]

Anxiety has several characteristics that illuminate Kierkegaard's broader understanding of free agency. First, sin is impossible without anxiety. Kierkegaard does

---

[60] Giles, "Kierkegaard's Leap: Anxiety and Freedom", p. 70. Again, I do not think that sin necessitates that violation of a divine command, but this certainly would qualify as sin because violating the command is shirking one's freedom.

[61] *JP* II 1530; Pap. X^2 A 481. Cited in Reider Thomte's footnotes to *CA*, 230, ftn. 1.

[62] Kosch, *Freedom and Reason*, 213. This is one of the clearest articulations of Kierkegaard's moral externalism and its relationship to his voluntaristic understanding of free agency.

not think that anxiety is all that is required for an agent to commit a free and responsible action, but the quality of anxiety presupposes that the agent's future is undetermined.[63] If the agent is anxious, no causal factors, internal or external to the agent, natural or moral, can *necessitate* that she perform a certain action. An agent is anxious because she perceives her situation as open-ended, that is, as having more than one option available to her.[64] Furthermore, the anxiety results from the "finitude of one's capacities and future lifespan lending a weight of significance to one's choices that would be absent if they could be made again" and arises from the intersection of an undetermined future and an externally given set of constraints, what Kierkegaard calls 'the prohibition'.[65] Though Kosch is right to say that in *CA* Kierkegaard's notion of the prohibition refers to "that fact that the agent is addressed by some demand or set of demands on his conduct that come from the outside," I think that the basic idea is broader than external normative constraints or divine commands. By prohibition Kierkegaard could just as easily mean the agent's immediate grasp of basic moral categories where some actions are wrong and morally blameworthy and should be avoided. The point is that anxiety does not exist unless the agent has an open future, a historical nature (subject to consequences), and is aware on some level that her actions potentially will bring her moral praise or blame.[66]

---

[63] Ibid., 210.

[64] Of course a story could be told where a person feels anxiety because she perceives the situation as open-ended when in reality it is not. But that is why anxiety is a necessary condition for the possibility of sin. Guilt is the other psychological side. It does in fact happen to people that they feel guilty for actions that they could not have avoided, but Kierkegaard believes that guilt should be *ascribed* to a person only if the action could have been avoided.

[65] Kosch, *Freedom and* Reason, 210.

[66] Another (previously mentioned) aspect of anxiety is that anxiety is an explanation that cannot itself be fully explained. "Psychology comes closest [to explaining the qualitative leap] and explains the last approximation, which is freedom's showing-itself-for-itself in the anxiety of possibility." *CA*, 76-77. Quoted in Giles, "Kierkegaard's Leap," 79. This aspect is important because it shows that Kierkegaard was

Kierkegaard's analysis of original sin and anxiety in *CA* also reveals the necessary historical component of free and responsible agency. Kierkegaard argues that to explain sin, which is in part the violation of an ethical norm, the agent's action must result from the responsibility-undermining external force of *genetic compulsion*. Kierkegaard rejects the traditional Augustinian understanding of original sin, which states that when Adam sinned, all of his future progeny were present in him and therefore committed the sin with him and justifiably suffer the same guilt as Adam.[67] Kierkegaard responds that whatever explains Adam's sin must also explain every other human being's sin and vice versa. The traditional Augustinian explanation fails to do this because the manner in which Adam acquires his guilt is fundamentally different than the manner in which the rest of the human race acquire their guilt. Adam acquires his guilt by sinning, uncaused by any external force. The rest of humanity do not acquire their guilt by sinning, but rather because of their guilt (inherited from Adam) they now sin. Therefore Kierkegaard thinks that the traditional understanding of original sin needs to be modified to give the same explanation for the sin of every single human being, including Adam. However, Kierkegaard does agree with at least two aspects of the traditional understanding of original sin. First, he agrees that the *de facto* universal condition of every human being is sinful (but this is because of each individual's sin, not because of Adam's). More importantly, he agrees with the traditional understanding of original sin that human beings are a race; therefore each person is affected by the actions of other people. Thus

---

aware of a classical objection to libertarian freedom, namely, that what happens at the moment of choice can never be fully explained.

[67] Nassif, Bradley L, "Toward a 'Catholic' Understanding of St. Augustine's View of Original Sin," *Union Seminary Quarterly Review*, 39 no. 4 (1984): 291.

the understanding of original sin as a trait or inclination passed down from person to person has some merit, because it suggests that Adam was different in one important (although not fundamental) sense: when Adam sinned, sinfulness had not yet entered the world. Once Adam sinned, sinfulness entered the world and does affect all subsequent human beings by increasing their *propensity* to sin without *causing* their sin, because a person can become guilty only through her own free (i.e. self-determined in the sense of lacking ultimate external determining causes) action. For this reason Kierkegaard states that "…sinfulness moves in quantitative categories, whereas sin constantly enters by the qualitative leap of the individual."[68] In other words, sinfulness (the guilt that enters into the world as the result of an individual's sin) grows greater and greater as more individuals sin. This in turn increases the propensity of future people to sin because "anxiety will be more reflective in a subsequent individual than in Adam."[69] However, sin can enter only by the "qualitative leap," which means that the movement from innocence to guilt cannot be a necessary transition that is ultimately caused by external forces; sin must be caused by the individual's free action.

For Kierkegaard, the three notions of reflexive higher-order mental states, libertarian free action, and the historical nature of selfhood all spring from his basic understanding of the nature of human agency or what he often calls the 'self'. Kierkegaard states in *The Concept of Anxiety* that a human being "is a synthesis of psyche and body, but he is also a *synthesis of the temporal and the eternal*";[70]

---

[68] *CA* 47.

[69] Ibid., 52.

[70] Ibid., 85. Emphasis in original.

furthermore this synthesis "is constituted and sustained by spirit."[71] By a synthesis of mind and body Kierkegaard means that humans consist of both a physical and a mental element. But the nature of selfhood is not exhausted with the physical and the mental. After all, some animals have various mental capacities—perhaps even basic emotions, thoughts, etc. —to go along with their physical bodies. If human beings were simply the physical and the mental then there would be no fundamental difference between humans and animals and thus no concept of *human agency* and the corresponding concepts of free will and moral responsibility. However, these aspects of human agency, or the self, are not realized to the same degree in every person and are not realized automatically: these aspects must be *achieved*.[72] The self is achieved at the most basic level through the reflective synthesizing activity of self-consciousness that Kierkegaard calls 'spirit'.[73] For Kierkegaard, an act of spirit is an act that is free from responsibility-undermining external forces and has real live options. Kierkegaard describes these conditions as the choice to sin or not to sin, to act in full freedom or to shirk that freedom. Reflexive activity occurs when an agent utilizes her ability for reflexivity and becomes aware of herself as an agent He writes,

> Man is a synthesis of the psychical and the physical; however, a synthesis is unthinkable if the two are not united in a third. The third is spirit. In innocence [prior to a person's first sin] man is not merely animal, for if he were at any moment of his life merely animal, he would never become a man. So spirit is present, but immediate, dreaming.[74]

---

[71] Ibid., 81.

[72] Mele, Peter J., *Thinking through Kierkegaard: Existential Identity in a Pluralistic World*, (Urbana: University of Illinois Press, 2005); 882.

[73] Kosch, *Freedom and Reason*, 200. This is seen most clearly in *The Sickness unto Death.*

[74] *CA*, 43.

A sin is a free choice between two real options, one good and one evil, which are both available to the agent at the moment of choice.[75] Before the agent makes a choice, she is innocent or her basic agential properties of freedom, reflexivity, and historicity are *in potentia*. Spirit is dreaming, present in the agent, but not wakened until the agent makes a free and responsible choice.

The free choice that awakens the latent spiritual element in the agent is also a naturally *historical* act that subjectively appropriates the eternal element of selfhood. "The synthesis of the temporal and the eternal is not another synthesis but is the *expression* of the first synthesis, according to which man is a synthesis of psyche and body that is sustained by spirit. As soon as the spirit is posited, the moment is present…*Only with the moment does history begin*."[76] Kierkegaard defines the moment as "that ambiguity in which time and eternity touch each other, and with this the concept of *temporality* is posited, whereby time constantly intersects eternity and eternity constantly pervades time."[77] Kierkegaard often uses the notion of the eternal to signify enduring and meaningful aspects of human existence, while the temporal often signifies commitments that lack significance and meaning. Temporal events will not stand the test of time. However, the temporal often refers to a mode of existence where the agent refuses to commit to the earnest pursuit of "projects that can give narrative shape and

---

[75] In this way Kierkegaard is following in the footsteps of Scotus, and he follows Scotus in other important ways as well (his voluntarism and his rejection of eudaimonistic ethics). However, I will demonstrate that Kierkegaard's view differs from Stump's interpretation of Scotus in that Kierkegaard does not believe that both options are equally available to the agent at the moment of choice, or that the agent is indifferent to those options. Because the agent has certain basic characteristics—freedom, reflexivity, and historicity—it is impossible for the agent to be equally attracted or indifferent to both actualizing and denying her own freedom.

[76] *CA*, 88-89. My emphasis.

[77] *CA*, 89.

enduring meaning to a human life."[78] This is why Kierkegaard says that to sin is to live "only in the moment as abstracted from the eternal." A person who lives in the moment refuses to engage authentically in ground projects[79] that endure through time and therefore refuses to take responsibility for herself as an agent capable of free and responsible action.

When a person makes a fundamental choice between good and evil, the latent spiritual aspect is awakened and selfhood is pervaded with eternality. In other words, the person has made a choice with lasting consequences that will extend that self into the future. "By this division, attention is drawn immediately to the fact that the future in a certain sense signifies more than the present and the past, because in a certain sense the future is the whole of which the past is a part, and the future can in a certain sense *signify the whole*."[80] Once the agent makes a free choice she becomes aware that her choices have consequences that affect both herself and other agents, and that the self that she will become in the future will result from her free and responsible choices. The realization of this social-historical aspect of agency is not immediately obvious. However, Kierkegaard's rejection of Pelagianism makes it clear that the agent's choices not only affect her own character but also affect other people: "[t]he race has its history, within which sinfulness continues to have its quantitative determinability, but innocence is

---

[78] Davenport, ""Towards an Existential Virtue Ethics: Kierkegaard and MacIntyre" in *Kierkegaard After MacIntyre*, p.265. This description fits the pseudonym A. I put off an extended discussion of Judge William's critique of A for the moment, but it suffices to point out that A refuses to let eternality pervade his existence by refusing to "choose to choose."

[79] This phrase is borrowed from Bernard Williams.

[80] *CA*, 89.

always lost only by the qualitative leap of the individual."[81] Kierkegaard thinks that one of the fundamental truths of original sin is that each individual's actions have a real effect on subsequent individuals. Although every person sins through a free action that is not ultimately *caused* by any other agent's actions, none the less, the anxiety (the condition which makes sin psychologically possible) of each generation is greater than the previous generation's. Kierkegaard stresses this point not primarily over a worry about orthodoxy but rather to emphasize the social-historical nature of human agency. Through the agent's free action her historicity is now actualized and she becomes aware of herself as an historical agent whose actions have consequences and through which she extends her self into the future.

Therefore through this "primordial choice"[82] the agent brings to the fore of her conscious life three fundamental aspects of human agency: the mental and physical aspects, the reflective aspect, and the historical aspect.[83] Though this basic primordial choice actualizes these aspects of human agency so that they are now in an awakened or active state, the agent does not necessarily utilize these aspects fully. The self for Kierkegaard is achieved in several steps or stages. The initial free and responsible choice is merely the first and basic step. The self can still refuse to form strongly evaluative higher-order desires,[84] can refuse to recognize her own historicity and the consequences

---

[81] *CA*, 37. It is not surprising that Kierkegaard offers his own understanding of Pelagianism, which while connected to the historical understanding in certain fundamental respects, is nevertheless tailored to suit his larger argument.

[82] Davenport's term. See "The Meaning of Kierkegaard's Choice between the Esthetic and the Ethical: A Response to MacIntyre."

[83] It should be noted that this choice can be avoided, as Kierkegaard takes pains to argue in the section on paganism. This is also seen in the case of A which is examined in a later section.

[84] As is the case with A, or at least Judge William's perception of A in *Either/Or II.*

that her actions have on her self as well as others,[85] and can refuse to live with the eternal constantly in mind.[86] But to refuse to engage any of these fundamental aspects of agency is to live in despair.[87] Most importantly, to refuse to engage any of these aspects of one's agency is to be subject to moral blame. All human agents (people other than very young children and perhaps those with severe developmental or cognitive impairments) can act in a free and responsible manner; that some refuse to do so does not imply that they are not blameworthy, but that they are not very far along to road to mature selfhood.

Case 2—Judge William's Analysis of A in Either/Or II

As in *The Concept of Anxiety*, Judge William's Analysis of A in *Either/Or* shows the overlapping necessity of higher-order desires, the right causal history of the action, and the qualitative leap in major ethical transitions—or, in slightly different terms, free and responsible choices. Davenport's interpretation of this case using Frankfurt's notions of higher-order volitions and his distinction between a wanton and a person is useful but ultimately flawed. His work serves as a good foil for understanding the ways that Kierkegaard's thought both anticipates and deviates from Frankfurt's thought, as well as the ways that Kierkegaard provides the responsibility-grounding freedom-relevant conditions that Frankfurt fails to provide. Davenport argues that Frankfurt's notion of volitional identification "provides a basis for explaining the meaning of the

---

[85] As is the case of certain forms of despair in *The Sickness unto Death*.

[86] As in Kierkegaard's diagnosis of many of his contemporary "Christians" in *Christian Discourses*.

[87] Needless to say that if the refusal fully to accept and engage any of these aspects of agency results in a form of despair, then the self is always responsible for despair because it is always a freely adopted condition, even if the particular state that the self finds itself in at the moment cannot be overcome through a simple act of will. This forms part of the basis of my Kierkegaardian critique of Frankfurt's hierarchical view of the will and the corresponding notion of moral responsibility.

Kierkegaardian 'choice' to make ethically significant choices." Davenport goes on to explain that

> By *ethical* "choice" between good and evil, or "the act of choosing" that is "essentially a proper and stringent expression of the ethical" (Lowrie, EO II, 170), the judge means a volition which satisfies the condition for strong moral responsibility…[or] *volitional identification* in Frankfurt's sense. To *choose* in this sense, one cannot just 'wantonly' act on whatever preference wins out in the "economy of one's desires." Rather, one must actively associate oneself with some form of deliberate action; the higher-order acts of identification this involves will then constitute an authentic inward self. The primordial choice between the aesthetic and the ethical generally, then, is the choice *either* to be wanton, *or* to become a 'person' in the full Frankfurtian sense.[88]

Davenport goes on to argue that Judge William's notion of immediacy roughly corresponds to Frankfurt's notion of a wanton. Furthermore Judge William's notion of personhood is roughly equivalent to Frankfurt's person according to Davenport.

At stake in this disagreement with Davenport is whether A is a proper candidate for responsibility ascriptions. If A does not exhibit volitional identification, then he is not a "person" or agent according to Frankfurt's definition. If A is not a person then he does not exhibit freedom of will. If A does not exhibit freedom of will then he is not a candidate for responsibility ascriptions. I follow Frankfurt's later distinction between volitional identification (which is required for moral responsibility) and care (which is not required for moral responsibility). Davenport focuses mainly on Frankfurt's early work, which does not make this distinction clear.[89] Although Davenport correctly argues

---

[88] Davenport, "The Meaning of Kierkegaard's Choice," 85. The second quote is from Frankfurt's "Freedom of the Will, 18. I have omitted Davenport's subscripts for clarity. Although Davenport's primary concern is not the freedom-relevant conditions for responsibility-ascriptions but rather to determine whether or not the choice that Judge William urges A to make is irrational and arbitrary as MacIntyre claims, his concerns are very close to mine and thus many of his comments are relevant to this discussion.

[89] Frankfurt makes a number of comments in both "Freedom of Will" and "The Importance of What We Care About" that suggest that caring and volitional identification are essentially the same. Take for instance page 16 of "Freedom of Will" where he states that a wanton "does not care about his will. His

that the Judge articulates a hierarchical structure of the will and volitional identification, the Judge does *not* describe A in wanton terms, and his notion of personality is much richer than Frankfurt's notion of a person. In fact, the Judge's understanding of personality is more akin to Frankfurt's notion of caring (with its narrative or historical aspects) than his rather thin definition of personhood (which only requires volitional identification). Whereas Frankfurt's notion of volitional identification cannot properly distinguish free and responsible action from action that results from responsibility-undermining external causes such as manipulation, the Judge's arguments demonstrate that a much richer understanding of volitional identification is necessary to make this distinction between responsible and manipulated action. The Judge's notion of volitional identification is one that branches outwards as well as upwards.[90] The Judge's analysis of A shows that he considers A morally responsible for his actions because A can identify with his motivational states. To identify with her motivations, the agent must demonstrate leeway or avoidability with regards to her choices, must be subject to the historical consequences of her actions, and must be able to appreciate and respond to reasons (including the ability to form new motivations in response to reasons). Since A

---

desires move him to do certain things, without its being true of him either that he wants to be moved by those desires or that he prefers to be moved by other desires." Furthermore, he says that wantons may have higher-order desires but not volitions. Taken together, these two points make it sound as if the crucial aspect of higher-order volitions is that the person with volitions *cares* about his motivational states whereas the wanton does not. But this view cannot be right, because surely Frankfurt does not intend to say that the wanton is morally responsible (since a wanton is not even a person!), and he later admits that all that moral responsibility requires is that the agent identity with his desires *at that moment.* However, he also makes clear that caring gives the agent narrative identity and orders the agent's preferences. But surely an agent can identify with her motives and want certain motives to be effectual without *caring* about that motive in Frankfurt's sense. Therefore these cannot be the same and I hold them as two distinct categories.

[90] This idea is adapted from Alfred Mele's *Autonomous Agents: From Self-Control to Autonomy*, esp. chapter four. I am arguing that higher-order desires (the upward element) are required for moral responsibility, but so is an historical or narrative element as well as other mental states such as emotions (the outward elements).

demonstrates all of these components, he is a person and a proper candidate for responsibility ascriptions.

The Judge's hierarchical view of the will. Judge William has a hierarchical view of the will at work when he advises young A to "choose to choose." The Judge considers A a proper candidate for responsibility ascriptions, but exhorts him to make an "ethical" choice, or what Davenport calls the "primordial choice" to be a chooser in the strong moral sense. The Judge tells A that "[r]ather than designating the choice between good and evil, my Either/Or designates the choice by which one chooses good and evil or rules them out."[91] The Judge later says to A that the choice he is referring to "is between choosing and not choosing…" and that he wishes to force A "to the point where the necessity of making a choice manifests itself and thereafter to consider existence under ethical qualifications."[92] Clearly this is not just any kind of choice. After all, '[a] person living from the aesthetic posture makes decisions, and their decisions can be very reflective…A…is a virtuoso when it comes to deliberation, and of course he makes choices."[93] But the Judge draws a distinction between two kinds of choices when he tells A that A's "choice is an aesthetic choice, but *an aesthetic choice is no choice*."[94] The first (aesthetic) choice is choice in the common sense understanding of the word; choosing to engage in various activities, and to act on certain motivational states instead of others (higher-order desires). The latter (ethical) choice is the choice to be an active

---

[91] *Either/Or,* 169.

[92] *Ibid.*, 177-78.

[93] Mehl, *Thinking Through Kierkegaard,* 15. For instance, the Judge says to A that 'if deliberating were the task of life, then you would be close to perfection." *EO II,* 165.

[94] *Either/Or*, 166. Emphasis mine.

(moral) agent, an agent who recognizes external moral demands, makes choices and volitional identifications in light of those moral demands, and guides herself along a certain path. The judge urges A to embrace himself as an active agent capable of free and responsible choices (the "ethical" choice), instead of remaining a passive agent who lets his own desires act on him, and lets them become responsibility-undermining external forces (the "aesthetic" choice).

This distinction between active and passive agency requires further elucidation. I am claiming that A simultaneously makes choices and identifies with certain desires and motivational states, and that he is passive with regards to his own agency and ultimately lets his desires and motivational states act on him like responsibility-undermining external forces. These two claims are not contradictory when understood in light of Frankfurt's distinction between higher-order volitions and care. A displays the former while lacking the latter. A exhibits higher-order desires and volitions and therefore is not a wanton in Frankfurt's sense. A wanton may have some second-order desires, yet she will not have second-order volitions. A does not demonstrate Frankfurt's notion of care. A's volitional identification and lack of care appears in the Judge's description of A's contemplation of an either/or, either to become a pastor or an actor. After strenuous deliberation, A "decides" to become a pastor, or more accurately, "reflection with its hundred arms seizes the *idea* of becoming a pastor."[95] Once A "decides", he energetically throws himself into the pursuit of this vocation and begins reading, writing sermons, and talking to pastors. A does everything involved in becoming a pastor except *actually becoming* a pastor. A obviously exhibits first-order desires (he desires to be

[95] Kierkegaard, *Either/Or II*, edited and translated by Howard V. Hong and Edna H. Hong (Princeton, N.J.: Princeton University Press, 1987), 165. Emphasis mine.

pastor, to write sermons, etc.) as well as higher-order desires (he identifies with his desire to become a pastor and to write sermons) and he acts on some of those desires. The main issue is how to make sense of A's mental states in light of the fact that he does not identify *fully* with his desire to become a pastor because he never forms the intention to become a pastor. Rather, A forms the intention to *experience* the vocational life of a pastor without committing to that life.

The Judge's analysis of A's second order desire to become a pastor is similar to an example that Frankfurt uses in his first article on the subject, "Freedom of the Will and the Concept of a Person." Frankfurt imagines a psychiatrist who wants to understand and sympathize with his patient's drug addiction.[96] The doctor wants to know what it is like to crave the drug as if he were an addict. He does not, however, want that desire to be *effectual*. He only desires to desire the drug, to feel the intensity of that desire. He does not actually want to take the drug. The doctor has a second-order desire, but not a second-order volition. This case is similar to and different from the case of A wanting to experience the life of a pastor, because A exhibits higher-order volitions while the doctor does not. A may already have the competing first-order desires to become a pastor and/or an actor, whereas the doctor does not actually have the desire to take the drug.[97] A's engagement in activities of a pastor indicates that he may identify with his desire to become a pastor. He wants to experience what it is like to be a pastor, yet his experiment will not move him to become a pastor without the requisite higher-order desires. On the

[96] Frankfurt, "Freedom of the Will," p. 9 (article).

[97] So in the doctor's case is through reflection attempting to create a desire that he does not have, while A already has competing desires, but is not trying to eradicate either but also not fully identify with either.

other hand, Frankfurt's scenario does not clarify whether the doctor engages in various activities related to drug use in order to help him "feel" the desire to take the drug. Because of the nature of the desire and corresponding activity in question (the desire to be addicted to drugs), the corresponding activities are difficult to imagine. Unlike the doctor, A can identify with various desires closely associated with the basic motivational state of desiring to become a pastor while simultaneously refusing to identify with the primary motivational state.

One important similarity between A and the doctor is that they both refrain from forming a higher-order volition about the primary motivational state in their respective scenarios (to take the drug and become a pastor), and they both refrain from caring about their primary motivational states (which is obvious for the doctor since caring presupposes volitional identification). Their similarities end here. However, they differ in that although they both desire to feel the desires associated with a certain life or "vocation" (pastor and drug addict respectively), the nature of A's primary motivation allows him to display volitional identification with some of this desires.

To complicate the scenario further, A demonstrates care (in Frankfurt's technical sense) through his *avoidance* of a vocation. A deliberates between one vocation and another, tries one on for size by playing the role and identifying with the various related desires. Yet he refrains from identifying with the primary motivational state (the decision to actually become a pastor with the cluster of other intentions and commitments that this initial decision includes). When the point comes that he must decide whether or not to identify with that primary motivational state, he proposes a new "Either/Or" and tries out a new vocation. A moves from one experience to the next without committing because

he does not wish his primary motivational state to be effectual because it is not just any motivational state (like the desire to eat cake or workout or any of a host of desires). A is presented by Judge William as a highly deliberative and reflective person who makes plenty of choices. Surely there are instances where his choices are the result of *reflection on* and *identification with* a first-order desire, and furthermore these are not all merely higher-order desires, but higher-order *volitions*. Recall that a higher-order volition can result from weak evaluation; all A must do is have a desire X, reflect on that desire X, and then form a higher-order desire that desire X be effectual. The desire to eat cake for dessert can be an example of a higher-order volition, as long as the agent goes through this process of identification. A exhibits higher-order volitions, but he ultimately fails to care about his life-choices. He does not guide himself along a certain path, except the path of *non-commitment*. Davenport argues that A tacitly wills "not to have any second-order volitions with respect to …the complex of desires and preferences that presently motivate his actions…"[98] I agree with this except for calling this will *tacit*. Judge William presents A as a person who intentionally refrains from identifying with desires that will give him a concrete binding personality. Intentionally refraining from identifying with personality-binding motivations is a higher-order volition, but not an instance of *caring*. As the paradigm aesthete, A above all else wants to satisfy whatever desire happens to come over him in any particular moment. He wants to refrain from identifying with a motivational state that will have a binding effect on him and will require him to make long-term commitments.

[98] Davenport, "The Meaning of Kierkegaard's Choice," 96.

A's *commitment to avoid commitments* is seen in the either/or between a pastor and an actor. As soon as the intensity of the desire to become a pastor passes, the desire to become a lawyer appears, and A begins anew. This shows that A is deliberate and calculating in the management of his desires. He carefully prevents any commitment-granting motivational state from achieving efficacy in the economy of his desires, because he is convinced that long-term commitments only bring despair and boredom, and he wants to avoid these above all.[99] This commitment to remain open to the possibility of new and interesting desires shows that A clearly is a person according to Frankfurt's definition. Frankfurt states that "when a *person* acts, it is according to the desire he wants or the desire he wants to be without (the case of the unwilling addict), with the *wanton* it is neither."[100] The very desires that A wants to be without are desires that require long-term commitments.

Therefore the question is not whether A exhibits volitional identification or is an apt candidate for praise and blame (which is entailed by *VI*), but whether he exhibits *care*. On the one hand it seems that he does. Frankfurt states that for a desire to represent or be the result of a care, the desire must "endure through an exercise of his (the agent's) own volitional activity rather than by its own inherent momentum" and that "[t]he persistence of his desire must be due to the fact that he is unwilling to give it up." [101] Clearly A's higher-order desires not to identify with commitment-granting motivational states result from his own volitional activity and a desire that he is not

[99] It is important to note that Frankfurt realizes that even a wanton can excel in instrumental deliberation, so deliberation alone does not eliminate 'A' from the category of a wanton.

[100] Frankfurt, "Freedom of the Will," 14.

[101] "On Caring," 16

willing to give up. A has moved beyond vulgar hedonism—the commitment to physical pleasure. He decides that he is better off not satisfying certain desires, namely desires that would require long-term commitments and would rule out satisfying whatever other desires happen to come along. A's only commitment is his refusal to make any long-term commitments. A has *weak* second-order volitions. A's higher-order desires demonstrate *volitional identification* because he does want certain desires to constitute his will (desires that will minimize pain and boredom, but will not develop a character that will constrain his desires); but, his higher-order volitions are *weak* because they do not stem from a larger framework of commitments and change from moment to moment based on what desires will yield the most pleasure (or more accurately prevent pain and boredom). This commitment to refrain from commitments is a commitment to *desire-fulfillment*, which Frankfurt explicitly rejects as a form of caring. A believes that he will derive more satisfaction by allowing his strongest desires to be effectual, and he often *desires* that his strongest desires be effectual (as long as they do not bind him to long-term commitments). Frankfurt rejects the notion that desire fulfillment brings happiness. He believes that a person's having a desire, even a very strong desire, does not mean that the person really *cares* about that desire. Care requires a commitment to the desire that gives narrative identity to the person in a way that the commitment to desire fulfillment cannot.[102]

[102] *Ibid.*, 161. Determining whether or not A cares in Frankfurt's sense is difficult at this point. This is because even in his definition of caring sometimes Frankfurt invokes narrative components and other times he seems to indicate that a person can care about something at one moment and not at the next. I am systematizing Frankfurt's thought by drawing what I see as a necessary distinction between higher-order volitions and caring, and the historical or narrative element seems to me to be the key distinction.

Higher-order desires and historicity. A's weak higher-order volitions demonstrate that he identifies with certain desires while rejecting others. The Judge's analysis of A reveals Kierkegaard's hierarchical understanding of the will, and the need for volitional identification for moral responsibility. The Judge does not call into question whether or not A is morally responsible for his actions or for his character. In fact, the Judge argues that A is cultivating character—although unintentionally—for which A is responsible despite the fact that he does not demonstrate care or responsible agency. Judge William argues that A may well find that these weak second-order desires and volitions solidify and that he develops a stable character, albeit unintentionally. The Judge tells A that he is like a captain who believes that he can turn his ship whatever direction he wants at any moment.[103] However, unlike A,

> "the captain will also be aware that during all this the ship is ploughing ahead with its normal velocity, thus there is but a single moment when it is inconsequential whether he does this or that. So also with a person—if he forgets to take into account the velocity—there eventually comes a moment where it is no longer a matter of an Either/Or, not because he has chosen, but because he has refrained from it…."[104]

The Judge does not believe that A's intentional refraining from identifying with commitment-granting motivational states is successful. He argues that A is developing binding characteristics through intentionally refraining from choosing, as he goes on to tell A that "if one puts off the choice, the personality or the obscure forces within it unconsciously chooses." The Judge argues that all persons already have "an outward character consisting of motivations we tend to act on"[105] and that "the personality is

---

[103] *EO II* 164.

[104] Ibid.

[105] Davenport, "The Meaning Of Kierkegaard's Choice," 102.

already interested in the choice before one chooses."[106] Despite A's effort to keep his personality "bare"[107] or free from higher-order volitions and character traits that incline him towards certain motivational states and actions rather than others, he is a person who has the same agential structure as all other human persons. That is, A possesses both negative and positive freedom, and as a human agent he is naturally inclined towards embracing his freedom by making free and responsible choices. Therefore A actively wills against this positive freedom. But as he wills to remain free from binding choices, he makes choices (some conscious and some not) that develop an even stronger personality with dispositions and motivations to act in certain ways.

A is responsible for his actions because he displays the capacity for higher-order volitions, makes free choices even as he seeks to deny himself free choices, and is subject to the historical consequences of this freely-formed higher-order volitions. But because he intentionally refrains from identifying with commitment forming motivational states, A may wake up one day to find that he *cannot* make any long-term commitments even if he wants to. He might become the kind of person who cannot commit, although he has become this kind of person unintentionally. After all, the inability to commit is a binding trait, and therefore an example of the very thing that A is intentionally avoiding. In this way, a person's weak higher-order desires and volitions might give volitional consistency or stability to his life. Judge William and Frankfurt agree that this kind of stability is "…merely fortuitous and inadvertent. It would not be the result of any deliberate or

---

[106] *EO II* 164. Quoted in Davenport, "The Meaning Of Kierkegaard's Choice," 102.

[107] *EO II* 164. The complete quote is as follows: "If one believes that at some moment a person can keep his personality blank and bare or that in the strictest sense one can halt and discontinue personal life, one certainly is mistaken. Already prior to one's choosing, the personality is interested in the choice, and if one puts off the choice, the personality or the obscure forces within it unconsciously chooses."

guiding intent on our part."[108] A person's capacity for higher-order desires and volitions might remain intact despite the fact that she cares about nothing (or cares only about remaining open to the strongest desire), rendering it necessary that a person have more than merely those states. A person who cares is different from someone like A. The person who cares plays an active, willing role in the *continued existence* of binding higher-order dispositions and attitudes by identifying with those desires.

Contra Davenport's interpretation, the case of A demonstrates that Kierkegaard does not think that strong evaluation is necessary for moral responsibility. A believes that he is subject to necessity, but the Judge argues that A is free to become a strong evaluator, that is, an agent with the ability to evaluate his first-order motives and preferences in light of moral concerns.[109] Furthermore, as a free agent he is subject to the consequences of his actions and as such is cultivating characteristics and personality traits for which he is responsible. Thus the main issue at stake in *EO II* is not whether A is free or morally responsible (since the Judge's arguments assume both), but whether or not he will choose to take responsibility for himself by choosing to play an *active* (as opposed to passive) role in the cultivation of his character.[110] "The conversion to an ethical standpoint is, in the Judge's characterization, equivalent to the acceptance of choice, the taking up of responsibility…the 'choice of oneself' is in the first instance a choice of oneself as agent, not the choice of a set of characteristically ethical values or a set of

---

[108] Frankfurt, "On Caring," 162.

[109] A already has higher-order desires and volitions, but they are weak volitions that are not connected to deeper cares or commitments.

[110] Kosch convincingly argues that there is good reason to believe that for Kierkegaard "the most basic division (of life views) is actually between views of life that embrace passivity and those that embrace the possibility of action." *Freedom and Reason,* 155.

more or less hedonistic ones."[111] The Judge's critique of A presents three related but conceptually distinct issues: the metaphysical components of freedom, the freedom-relevant responsibility-grounding conditions, and the necessary components of responsible agency and character cultivation. These distinctions perhaps show the deepest similarity between Kierkegaard and Frankfurt. Just as Frankfurt distinguishes between responsibility—where all that is required is the right mesh of desires at the moment of action—and caring—where the commitment must be formed and carried out over time—so Kierkegaard distinguishes between moral responsibility and becoming an ethically existing self, that is, a strong evaluator who takes responsibility for his character.

This deep similarity also reveals the deepest difference between Kierkegaard and Frankfurt. In Frankfurt's thought the requirements for moral responsibility are disconnected from the requirements for caring. Whereas his necessary conditions for responsibility ascriptions do not distinguish cases of manipulation properly from cases of free and responsible action (since a deterministic force or manipulator could cause the agent to have the right mesh or desires), his notion of caring comes much closer to making this distinction. It is difficult to imagine a case of someone who cares about something in Frankfurt's sense yet is subject to responsibility-undermining external causes such as manipulation. This difficulty is due to Frankfurt's understanding of caring that emphasizes the narrative or diachronic aspect of caring, the hierarchical aspect of caring, and the fact that a person who cares often refuses to identify with other desires. The rift between Frankfurt and Kierkegaard is here. Whereas Frankfurt sometimes

[111] Kosch, *Freedom and Reason*, 150.

indicates that the free will requires alternative motivational possibilities,[112] he clearly rejects this view when he declares that a manipulator could causally produce an agent's first-order and higher-order desires. Also, as long as the agent has the right mesh of desires and is satisfied with those desires, then she is responsible and exhibits the ascription-granting freedom-relevant conditions. Recently Frankfurt admitted, in correspondence with Robert Kane, that hard compatibilism (or global manipulation such as Skinner's *Walden Two* scenario) "is in fact consistent with…[Frankfurt's] general views about identification and wholeheartedness."[113] Since wholeheartedness or volitional necessity is the pinnacle of freedom and caring on Frankfurt's view, he must believe that caring is also consistent with global manipulation or what Kane calls *covert nonconstraining control* (CNC). As long as an agent is free from constraint, can will what she wants to will and is satisfied with the will that she has, manipulation, such as the behavioral engineering seen in *Walden Two,* is irrelevant on Frankfurt's view. Clearly Frankfurt's time-slice hierarchical responsibility theory does not distinguish properly between free and responsible action and action that results from manipulation; morever, neither does his theory of care and wholeheartedness.

Kierkegaard agrees with Frankfurt that there are differences between the requirements for moral responsibility and the requirements for caring or responsible

---

[112] Recall his early definition of free will in "Freedom of Will" where he states that "[a] person's will is free only if he is free to have the will he wants. This means that, with regard to any of his first-order desires, he is free either to make that desire his will or to make some other first-order desire his will instead. Whatever his will, then, *the will of the person whose will is free could have been otherwise*; he could have done otherwise than to constitute his will as he did" (emphasis mine). This statement clearly indicates that in order to have free will, a person must have access to alternative motivational possibilities. However, even in this early work Frankfurt is clear that AMPs are not necessary for moral responsibility, and more recently is clear that responsibility is fully compatible with covert non-constraining manipulation and therefore most certainly does not require alternative possibilities or any sort.

[113] Kane, *The Significance of Free Will*, 225 n. 13.

agency. However, Kierkegaard's requirements for both moral responsibility and responsible agency properly distinguish free and responsible action from action that results from manipulation. Kierkegaard's theory of moral responsibility as well responsible agency requires a cognitive component, an historical component, and a leeway component, all of which Frankfurt rejects. The most crucial of these components is the leeway component, as "the presence of any alternative possibilities is a sufficient condition for the falsity of causal determinism…"[114] Recall that Frankfurt's initial distinction between first and second-order desires and the concept of volitional identification was that the agent is the source of free and responsible action, whereas she is not the source if a manipulator intervenes. This source condition is sometimes referred to as the ability of a free and responsible agent to determine her path, or possess self-determination. Again, semi-compatibilists and source incompatibilists are mistaken to reject alternative possibilities because "the significance of such alternatives…lies in their being indicators of the self-determination manifested by one's actions, which is necessary for responsibility."[115] Kierkegaard understands that alternative possibilities are necessary because those possibilities indicate the falsity of causal determinism. Furthermore, alternative possibilities demonstrate that the agent was the source of her action, and sourcehood is vital when it comes to the agent's ability to engage in responsible agency. Despite the fact that Kierkegaard's requirements for responsible agency are more stringent than the requirements for moral responsibility, they are not fundamentally different.

---

[114] Timpe, "Source Incompatibilism and its Alternatives," 155.

[115] O'Connor, Timothy, "Freedom With a Human Face," 209.

Kierkegaard's Hierarchical-Libertarian Solution to the Problem of Identification and Moral Self-Cultivation

A Frankfurtian Methodology for Self-Cultivation?

This final section deals with another problem with Frankfurt's (and indeed all soft or semi-compatibilists') understanding of the freedom-relevant conditions necessary for grounding moral responsibility. I examine whether any account for moral self-cultivation is derivable from Frankfurt's semi-compatibilistic understanding of free agency. I argue that Frankfurt's theory ultimately fails because he denies that agents play a causal role in the development of agential characteristics and motivational states, and because he subscribes to a version of reasons internalism coupled with what I call motivational noncognitivism. As shown in my first three chapters, the responsibility-grounding freedom-relevant conditions must be different from the agential conditions seen in cases involving manipulation, coercion, and other responsibility-undermining external causes. But an agent who can act *only* on the basis of prior motivation and thus has no choice with regards to her internal motivational possibilities is not fundamentally different from an agent who is determined by an external force, unless however, she has played a formative role in the cultivation of the cluster of mental states, dispositions, and character traits that produce those motivations. To reiterate, my claim is that either 1) the agent must be free to choose from various motivations or to cultivate new motivations at the moment of identification (which Frankfurt's reasons internalism denies) or 2) the agent must have played a causal role in the cultivation of the motivational states that produce the action (in order for the agent to play a truly causal role then i. causal determinism must be false and ii. the agent must have leeway—point 1—at some juncture in the past).

Therefore to affirm both that 1) an agent can be free and responsible who cannot form new motivations and 2) has played no causal role in the formation of her desires and emotions is another way of denying the necessary freedom-relevant responsibility-grounding components and the freedom necessary for morally responsible agency.

Once again on identification. The pinnacle of human freedom as Frankfurt understands it is the ability to commit wholeheartedly to one's cares. He believes that wholehearted concern about final ends endows our lives with meaning and purpose. Furthermore, he argues that "[t]o be a *person* entails evaluative attitudes (not necessarily based on moral considerations) towards oneself…Instead of responding unreflectively to whatever he happens to feel most strongly, he undertakes to guide his conduct in accordance with what he really cares about."[116] A person is able to guide herself along a certain path, namely a path approved by higher-order evaluative attitudes. Yet it is puzzling and troubling that despite his very promising account of identification and wholeheartedness, he gives no indication of how a person might achieve the state of wholehearted love.[117] He does, however, indicate in several places that intentional self-cultivation is possible. For instance, he states that "[w]e can sometimes take steps that inhibit us from loving, or steps that stimulate us to love; more or less effective

[116] "On the Necessity of Ideals," *Necessity, Volition, and Love*, 113-14.

[117] Some might respond that this should not be puzzling at all, since Frankfurt is a compatibilist and thus does not think that wholehearted cares are necessarily at odds with determinism. He thinks that it is quite conceivable that some causal influence outside the agent's control could make it the case that the agent both has wholehearted cares and wants to have these cares. It would not matter that the agent played no causal role in bringing about this state of affairs; therefore there is no need for a theory of self-cultivation. While I grant this objection, I think that two things can be said in response. First, there is nothing inherent in Frankfurt's hierarchical account of freedom that is incompatible with a sophisticated libertarian account. Second, libertarians might go further and argue that we not only want to care wholeheartedly about things, but to care about the things that we have freely chosen to care about, and that we have played a significant role in coming to care about. For an argument along these lines, see Robert Kane's *The Significance of Free Will*.

precautions and therapies may be available, by means of which a person can influence whether love develops or whether it lasts."[118] A creative interpretation of Freud's famous case known as the "Rat Man" proves useful in evaluating this claim about the revisability of love.[119] The Rat Man both desperately loved and violently hated his father, and came to have split personalities "by repressing his hatred and only acknowledging his love."[120] Rat man was clearly neurotic, constantly undoing his own actions and contradicting his own thoughts. The Rat man is divided against himself, but Frankfurt believes that he should (ideally) be wholehearted. So what would Frankfurt have him do?

> In order for the Rat man to be wholeheartedly on the side of his benign attitudes, it would not have been necessary for him to conceal his hostile feelings from himself (which is what he has done). Nor would he have had to refrain from making a conscious effort to deal with those feelings in whatever ways might be effective and helpful. It would have required only that, in the struggle between his hatred and his love for his father, he himself come to stand decisively against the hatred and behind the love.[121]

Thus Frankfurt certainly does not think that that the Rat man should do what he is doing, namely repress the desires that he does not endorse. In fact, he should "deal with those feelings in whatever ways might be effective and helpful" and "stand decisively against

---

[118] Frankfurt, *Taking Ourselves Seriously,* 41. In another context Frankfurt states that "[i]t is surely open to someone for whom an action is unthinkable to alter his own will in such a way that the action becomes thinkable for him. The fact that a person cannot bring himself to perform an action does not entail that he cannot bring himself to act with the intention of changing that fact." "Rationality and the Unthinkable," 187. It is also important to note that Frankfurt does not think that a person can *directly* will to care about something, because "[l]ove is nonetheless involuntary, in that it is not under the immediate control of the will. We cannot love—or stop loving—merely by deciding to do so." Therefore Frankfurt does propose that a person can influence her cares, but this will not happen through a single, direct act of will. Cf. Kierkegaard on patience: "'to gain one's soul'…immediately turns the mind to a *quiet but unflagging activity*…not of making a conquest, of hunting and seizing something, but of becoming more and more quiet…" "To Preserve One's Soul in Patience," *Eighteen Upbuilding Discourses,* 170-71.

[119] J. David Velleman uses this example is his article "Identification and Identity" and Frankfurt continues with this example in his "Reply to J. David Velleman" in *Contours of Agency*, 91-128.

[120] Ibid, 101.

[121] Frankfurt, "Reply to J. David Velleman*,"* 126.

the hatred and behind the love."[122] In other words, Frankfurt's only advice to the person who has desires or other mental states that he does not approve of is to stand against them so that he can unify his will and thus achieve "wholeheartedness."

Despite Frankfurt's statements appearing to support a theory of moral self-cultivation, there are several reasons why Frankfurt never develops any such theory. In fact, not only does Frankfurt not develop a theory of self-cultivation; two main issues in his moral psychology and philosophy of mind disallow such a theory. The first reason that Frankfurt cannot develop a theory of moral self-cultivation (or morally responsible agency) is that he does not think that agents can causally influence the development of the mental states (attitudes, desires, emotions, etc.) that are the fundamental ingredients of character. This is peculiar considering Frankfurt's understanding of cares and the difference between cares and mere desires. Although Frankfurt believes that cares are constituted by desires, the mere existence of a desire in a person's psychological makeup does not mean that the person cares about that desire or its object. Caring about something means that the person is committed to it and does things to keep the desire or desires that constitute the care active. "When we do care about something, we go beyond wanting it…the caring entails, in other words, a commitment to the desire."[123] Not only must the agent be committed to the desire, but the desire must "endure through an exercise of his own volitional activity rather than by its own inherent momentum."[124] When a person cares about something, she is both *committed to* and *active in* guiding the

---

[122] Ibid.

[123] Ibid., 18.

[124] "On Caring," 160.

course of her desires. Furthermore, since Frankfurt clearly understands desires as the essential aspect of a person's character, these claims support the notion that agents do play an active role in the development of character.

However, despite Frankfurt's claim about the active role that agents play in the cultivation of desires that constitute caring, he explicitly rejects Aristotle's claim that "a person becomes responsible for his character insofar as he shapes it by voluntary choices that cause him to develop habits of discipline or indulgence and hence that make his character what it is."[125] In another place he ascribes to Aristotle the view that "a person acquires responsibility for his own character…by acting in ways that are causally instrumental in bringing about that he has the particular set of dispositions of which his character is constituted.[126] Frankfurt does not think that the main issue in moral responsibility and cultivation has to do with whether or not the agent has been active in developing certain habits, dispositions, and characteristics that now give rise to the desires (and the resulting behavior) in question. It does not matter, according to Frankfurt, whether the agent has been "causally instrumental" in the formation of his character. What matters is whether or not the agent now identifies with the desire in question, and has "taken responsibility" for these characteristics.[127] This has been called a "time-slice" theory of identification by Fischer and others, and this label is apt.[128]

---

[125] *Taking Ourselves Seriously.,* 6.

[126] "Identification and Wholeheartedness," 171.

[127] *Ibid.*

[128] See for instance Fischer's "Responsibility and Manipulation," *The Journal of Ethics*, 162, and Watson's "Reasons and Responsibility," *Ethics*, 111, no. 2 (2001): 384. Watson and Davenport both argue that it is not clear that Frankfurt's early view is a time-slice view, but Watson admits that based on Frankfurt's later comments it is clear that Frankfurt has come to embrace fully a time-slice theory of moral

"What counts is our *current effort* to define and to manage ourselves, and not the story of how we came to be in the situation with which we are now attempting to cope."[129] In other words, the agent is morally responsible for whatever desires she identifies with *at the moment of action* (higher-order desires) and not for the first-order desires that she has at that moment.

The last point, about the agent's responsibility to identify with certain desires and alienate other desires, illuminates a second fundamental reason why Frankfurt cannot articulate the freedom-relevant conditions necessary for morally responsible agency. Whereas Frankfurt thinks that many of an agent's mental states (especially her cares—the things that she is attached to, desires, or cause her joy and sorrow, etc.) are things that the agent simply finds herself with, he believes that the agent is responsible for her (higher-order) *response* to those mental states. Frankfurt believes that agents find themselves with "psychic raw elements that nature and circumstances have provided us…" and are content with some of those elements and not with others. Through the willing acceptance of attitudes, thoughts, feelings, desires, etc., the agent takes responsibility for the status of those mental states and they are transformed into authentic expressions of the agent.[130] Frankfurt believes that one of the chief problems in practical reasoning is helping agents determine which desires should be motivating. This separates him from the implausible position that the mere presence of a desire gives the agent a reason (though it may not be very strong) to act.

---

responsibility. I agree that there is a stark contrast between Frankfurt's early writing on caring (which clearly includes a narrative or historical element) and his mesh-theory of moral responsibility.

[129] Frankfurt, "Taking Ourselves Seriously," 7, emphasis mine.

[130] Frankfurt, *Taking Ourselves Seriously and Getting it Right*, 7-8.

Frankfurt's insistence that the agent is responsible for her higher-order desires yet not for her characteristics reveals several important problems. First, as argued in chapter one, his hierarchical mesh theory alone cannot properly differentiate between responsible action and cases of responsibility-undermining external causes such as manipulation. Furthermore, it reveals an inadequate moral psychology and a thin understanding of the nature of desires and emotions. Frankfurt's belief that desires and emotions are something that an agent simply finds in herself indicates that he holds to a noncognitive understanding of motivational states. By *motivational noncognitivism* I mean the notion that all first-order desires are either purely physiological or result from forces completely outside the agent's control rather than being thought-based in any way.[131] Frankfurt's comment that an agent is not responsible for the desires she finds herself with indicates that he thinks that desires are noncognitive in precisely this way. Perhaps the most challenging problem for Frankfurt's motivational noncognitivism is formulating a theory that can simultaneously account for how an agent can freely form a higher-order desire in response to a first-order desire when the agent has no control over that first-order desire. Frankfurt argues that the agent does not play an active role in the formation of the given

[131] An obvious objection might arise here, namely, that desires are the paradigm example of noncognitive mental states, as desires cannot be true or false and have the opposite mind-world fit of beliefs (which are the paradigm example of cognitive mental states) as well as perceptions (of which emotions are a type), both of which can be false. However, my claim about the cognitive nature of desires is not about their direction of fit (which I am not challenging), or about their propositional truth conditions (which clearly they don't have in the way that beliefs or perceptions do), but rather a claim about the relationship between an agent's desires, reflection, and cares or commitments. I will argue that while it is certainly true that human agents have certain unalterable desires, many other human desires arise from concerns and commitments, and since these concerns and commitments are alterable through reflection and action, so also are many human desires (indirectly) alterable. My usage of cognitivism is different from moral cognitivism, which indicates that moral claims can be true or false, but that desires are thought-based in the sense that cognitive strategies can affect an agent's desires. I believe that this is one of the most important unexplored areas of human freedom directly relevant to moral responsibility, and was ironically one of Frankfurt's most intriguing early claims; what it means to be a human agent or person includes the ability to do what one wants as well as the ability to alter freely the direction of one's will. If the will consists chiefly in desires, then a human agent must be able to change what she desires.

(first-order) desires that are part of the agent's character, but does play an active role in forming desires in *response* to those given desires after reflecting on the first-order desires and their relationship to the agent's deeper cares and concerns. The agent is not responsible for first-order desires that are given and completely out of the agent's control; the agent is responsible for higher-order desires that result from (at least weak) evaluation and are at least to some extent under the agent's control.[132]

The most plausible way to understand this formation of higher-order desires is by postulating that the agent thinks about which of her desires she desires more and can form a desire based on those *thoughts* (and be responsible for the at least partially *thought-determined higher-order desire*). However, this is problematic since the desire that the agent is evaluating is a desire for which she is not responsible according to Frankfurt because agents cannot play a role in the formation of the characteristics which in turn give rise to desires (or the desires themselves which result from nature or circumstance and are thus a force that is not under the agent's control). However, both of these assertions cannot be true. Either an agent plays a role in the formation of desires by forming thoughts that in turn shape desires and cultivate characteristics, or the agent cannot evaluate and form higher-order desires based on thoughts.

As I said in chapter three, Frankfurt's way of avoiding this blatant contradiction is to modify the second premise. Frankfurt believes that the process of higher-order evaluation of first-order desires (and other mental states or "attitudes" as Frankfurt calls

[132] Recall that the only freedom-relevant condition that Frankfurt affirms is the source component, and that the basis for his source component is the agent's ability freely to endorse or identify with a first-order desire.

them[133]) is purely value-neutral, and is a matter of weighing the competing desires in question and determining which desire is stronger and is therefore more desirable (in this instance of weak evaluation strength is equated with desirability). Thus the role of the intellect is purely *calculative*.[134] In Charles Taylor's terms, Frankfurt employs a notion of weak evaluation. All that is necessary for an agent to endorse a desire and form a higher-order desire is that the agent desire to X, reflect on the desire to X, and then decide to endorse (form a higher-order desire about) desire X. The mere first-order desire can be taken by the agent as a sufficient reason to endorse and act upon that desire; no further *moral* reflection on (or strong evaluation of) the desire is necessary. But this raises a further issue articulated by Stump and others and sufficiently dealt with in chapter three. If a higher-order desire is merely a mirror or reflection of a first-order desire, then nothing inherent in the nature of higher-order desires is responsibility-conferring. Responsibility-conferring higher-order desires must result from reasons-based evaluation that connects the desire to the person in a fundamental way.[135] Once this category of reasons-based evaluation is introduced, the agent's reasoning ability is no

---

[133] "On the Necessity of Ideals," 113.

[134] Frankfurt often refers to the *instrumental* nature of reason, by which I think he means the *calculative* nature of reason. Part of the role of reason on Frankfurt's view is to weight or calculate the desirability of desires; another part is to decide how to achieve the agent's final ends. This latter function appears to be why he sometimes refers to the instrumental nature of reason, by I think that he is confusing instrumental ends with instrumental reason.

[135] I argue that higher-order desires need not exist in an larger *intellectualist* framework like Stump's in order to ground moral responsibility, yet they do have to have a weak *knowledge* component (where the agent is responsive to reasons) and that this component can (and does) exist in Kierkegaard's voluntaristic account of moral responsibility due in large part to his cognitive understanding of motivational states. Thus I am calling Kierkegaard's theory of evaluation reasons-based because I do not think that he believes that an agent has to engage in strong evaluation to be morally responsible, but that the agent must have the ability to form thought-based motivational states. An agent who refuses to utilize this ability can still be morally responsible on Kierkegaard's account.

longer *merely* calculative, but causally contributes to the formation of responsibility-conferring higher-order desires.

The introduction of reasons-based evaluation and the knowledge component of higher-order desires allows for the possibility of (and in fact demands) a cognitive understanding of desires and emotions, which in turn allows for a way of articulating the agent's role in the cultivation of character via the cultivation of motivational states and their constitutive emotions and desires. Furthermore, this provides the sorely needed connection between the historical, hierarchical, and libertarian responsibility-grounding components. If a necessary component of responsibility-conferring desires and emotions is that they are thought-based, then it is possible that agents play a role in shaping their character-constituting desires and emotions and they play a role in forming higher-order responses to those basic emotions and desires.

The first problem for a Frankfurtian theory of moral self-cultivation—his rejection of Aristotle's causal theory of character—does not square well with some of Frankfurt's own statements about identification. Frankfurt is rejecting the broadly Aristotelian notion that an agent's present desires are connected to character traits which the agent has played a causally instrumental role in developing. Recall from chapter one that Aristotle is concerned with actions being up to the agent, and cites three conditions for responsibility, what I call the source, control, and knowledge components respectively. He articulates a strong control component by stating that it must be within the agent's power to perform or refrain from performing, and a source component by stressing that the agent have within herself "the 'origin' (*arche*) of the action."[136] This

[136] Sorabji, *Necessity, Cause, and Blame: Perspectives on Aristotle's Theory*, (Ithaca, N.Y.: Cornell University Press, 1980), 234. Quoted in Kane, *The Significance of Free Will*, 33.

second reason is crucial, because as Robert Kane puts it, "Aristotle is worried about whether agents are responsible for the characters and motives that are the sources (*archai*) of their actions."[137] Kane believes that in order to impute responsibility to agents for actions that are "volitionally necessary" in Frankfurt's sense, the agent must have made real live choices along the way that contributed to the formation of character from which the action now arises. In other words, whether or not the agent identifies with a desire at the moment that it arises is not the sole issue in ascribing responsibility (though it is certainly important). Instead, Kane follows Aristotle in arguing that agents can contribute to the formation of their own character, and that is why agents are responsible for both the character from which psychic states arise and for how the agent responds to those states.

Finally, in addition to his rejection of a causal theory of character development and a commitment to a form of reasons internalism and motivation noncognitivism, Frankfurt does make clear how agents can deal with the psychic elements that they do not wish to have. He suggests that "[w]hat counts is our current effort to define and to manage ourselves…" and that we can "decisively rid ourselves of any responsibility for their continuation by renouncing them and struggling conscientiously to prevent them from affecting our conduct."[138] Renunciation and struggle may well be the proper response to these elements, but how exactly does Frankfurt think that the struggle is going to make any difference? After all, a person can struggle all she wants against desires, emotions, and attitudes that she wishes she did not have, but it will not matter if

---

[137] Kane, *The Significance of Free Will*, 34. Kane calls this ultimate responsibility, and he argues that it is a key aspect of libertarian freedom, which Kierkegaard clearly affirms.

[138] Ibid.

she does not have any control over her character traits that give rise to those psychic elements because her actions cannot causally contribute to the formation of different character traits.[139] An agent's struggle can succeed in preventing these elements from affecting her future behavior only if she can by that struggle begin to reshape her character so that she will have different desires, emotions, etc. Frankfurt's solution is partially right (and he and Kierkegaard are largely in agreement here); a person must struggle against the elements that she wishes not to have and seek to renounce them. However, his moral psychology will not allow for this solution because his "time-slice" understanding of identification, his subsequent rejection of Aristotle's causal notion of character development, and his acceptance of a reasons internalism coupled with a motivation noncognitivism.

Kierkegaard's motivational cognitivism and moral self-cultivation. Kierkegaard differs from Frankfurt in his insistence that emotions and desires have a cognitive element and thus an agent can play a role in shaping her emotions and desires and thus her character. Furthermore, this cognitive articulation of desires and emotions is placed within the necessary larger libertarian framework that includes both hierarchical and historical elements. Kierkegaard has a notion similar to Frankfurt's "care", what Kierkegaard often refers to as *passion*. The Danish word for passion—*lidenskab*—has a dual function in Kierkegaard's writings. It can refer to "the kind of state that we usually call emotion—a response to particular features (as the subject sees it) of the subject's

[139] More precisely, the agent is not responsible for the character traits that she finds herself with, and has weak control over the responsibility-conferring higher-order desires. What Frankfurt is interested in is a certain structure of will necessary for moral responsibility - the time-slice mesh of first and second-order desires - and not how that structure came to be. My objection here is to the compatibility of that claim with the claim that agents can intentionally alter their loves. I see no way to reconcile these two claims without radically altering one or both.

world." It can also refer to "the concern on which such responses are contingent."[140] My claim is that this second use of passion as concern is very similar to Frankfurt's notion of care. However, Frankfurt sees no connection between care and emotions: emotions happen to a person, but that person can at least potentially alter what she cares about. Kierkegaard understands that "…emotion is one of the chief ways that passion is manifested in human beings; emotions are symptoms and fruitions of passions…"[141] and that "emotions can be expressions of character traits: emotions are based on concerns, some concerns are passions, and passions are character traits, ongoing master concerns that deeply characterize a person."[142] When Kierkegaard thinks about his readers, he does not consider their actions as important as he does their emotions, because emotions are the prime revealers of *concern* (or Frankfurt's care). Kierkegaard generates a strategy for cultivating different cares (or strengthening weak cares or even eradicating cares) because he thinks that passions (both in the sense of cares and emotions) are in some sense thought-determined, at least in the most basic sense that a person must have some conception of the object that interests her to be passionate about that object.[143] In his critique of a Pastor Adler's claim to have received a direct revelation from God, Kierkegaard responds that "not every outpouring of religious emotion is a Christian outpouring…[but rather] emotion that is Christian is controlled by conceptual

---

[140] Roberts, Robert C. "Existence, Emotion, and Virtue: Classical Themes in Kierkegaard," *The Cambridge Companion to Kierkegaard*, ed. by Alastair Hannay and Gordon D. Marino, (Cambridge: Cambridge University Press, 1998), 187.

[141] Roberts, Robert C. "Passion and Reflection" in *International Kierkegaard Commentary: Two Ages*, 90.

[142] Roberts, Robert C. *Spiritual Emotions: A Psychology of Christian Emotions*, 20.

[143] Roberts, "Passion and Reflection," 91.

definitions…"[144] In other words, Kierkegaard's understanding of the emotions is that they "always involve some assessment of one's situation…(and) this assessment will always involve reflection."[145]

Kierkegaard's cognitive understanding of emotions and desires—or what I call his motivational cognitivism—helps solve Frankfurt's hierarchical problem. Frankfurt developed his hierarchical notion of the will as a compatibilist-friendly source component that he believed could ground moral responsibility without reference to alternative possibilities. However, in part due to his motivational noncognitivism and a decidedly time-slice understanding of moral responsibility, there is no responsibility-grounding element present in a higher-order desire because there is nothing about a higher-order desire that connects it to the self as source in a way that distinguishes the higher-order desires from potentially determined first-order desires over which the agent had no control. As Frankfurt admits, "A manipulator may succeed, through his interventions, in providing a person not merely with particular feelings and thoughts but with a new character. That person is then morally responsible for the choices and the conduct to which having this character leads…"[146] Frankfurt can make this statement because he believes that basic mental states (and apparently character as well) are things that agents simply find themselves with, not things that agents play an active role in cultivating. Therefore a manipulator can provide the agent with various mental states as well as a character but the agent is still responsible for how she responds to that character, and her

---

[144] Kierkegaard, Søren, *The Book on Adler*, trans. by Howard V. Hong, and Edna H. Hong, (Princeton, N.J.: Princeton University Press, 1998), 113.

[145] Roberts, "Passion and Reflection," 90.

[146] Frankfurt, "Response to John Martin Fischer," 27.

higher-order response grounds her choice (thus making her responsible) because she is the source of that response in a way that she is not responsible for her basic mental states and character. The problem is that there is nothing that differentiates the agent's responsible higher-order mental state from a mental state caused by external force (such as a manipulator). Thus without a cognitive (or knowledge) component, the higher-order desire does not provide the necessary responsibility grounding source component.

Case 3: Kierkegaard's Motivational Cognitivism and Hierarchical Structure of the Will as seen in Christian Discourses

Kierkegaard solves Frankfurt's hierarchical problem not by connecting the will directly to the intellect as Stump and Fischer try to, but by showing the relationship between the agent's intellect and her emotions and desires. According to Kierkegaard, emotions are thought-based and can be altered through reflection. Furthermore, "…the person at his or her core is not just manifested in emotion, but is somehow actualized or made present in a special way in emotion, just as he is actualized or made present in his action."[147] Kierkegaard understands that there is a fundamental connection between a person's emotions and her concerns or cares. An agent's volitions are her own because they stem from a larger framework of cares and concerns, and these concerns are manifested by (and in turn shaped by) the agent's emotions and desires. Finally, because emotions and desires are to some extent thought-based, they are to some degree under the command of the will.[148] A person is responsible for volitions that stem from her cares as

[147] Roberts, "Existence, Emotion, and Virtue: Classical Themes in Kierkegaard," 198.

[148] Cf. Roberts, "Existence, Emotion, and Virtue: Classical Themes in Kierkegaard," 193.

well as her emotions and desires because all of these motivational states are to some extent thought-based and therefore subject to the agent's thoughts.

In this section I will show that something akin to Frankfurt's hierarchical view of freedom is at work in Kierkegaard's *Christian Discourses*. Whereas Frankfurt's commitments rule out an articulation of moral self-cultivation, Kierkegaard's rich moral psychology, along with his libertarian understanding of the will, allows him to show the role that second-order desires and emotions play in the cultivation of character. In his introduction to part one of the *Christian Discourses*, "The Cares of the Pagans," Kierkegaard makes some observations that set the tone for what he is trying to accomplish. He tells us that he is going to compare and contrast the cares (*bekymring* - anxiety or worry) of the bird and the lily in Matthew 6:25-33 with the cares of the pagan. In the passage, Jesus exhorts his disciples not to be anxious about tomorrow, about what they are going to eat or wear, but to look at the birds of the air and the lilies of the field, and to trust that the same God that provides for them will also provide for his followers. Kierkegaard believes that in this passage we have three kinds of care or passion represented; first, the care of the bird and the lily; second, the care of the pagan who is anxious about tomorrow; and third, the care of the follower of Christ who is not anxious about tomorrow but is in some way like the bird and the lily because they do not worry about tomorrow either.

Kierkegaard's methodology is one of reflection and conceptual analysis. Thinking about the birds of the air and comparing them to pagans and Christians, we can see that the birds desire food, but are not worried about whether or not they will find any food, at least in part because they are not concerned about long-term security and do not form

beliefs about their need for food. The pagans desire food and security generally, and this basic concern is coupled with the belief that there may not be enough food for tomorrow, which in turn generates the worry about not having enough to eat. Christians also desire food and security generally, but this basic concern is coupled with the belief that food is a gift provided by God, and that God will provide. Kierkegaard holds that although the Christian has the same basic concern, her distinct belief will generate different emotions such as gratitude over God's provision. This methodological note indicates how Kierkegaard is thinking about cares in these discourses. First, people's cares can tell us something much deeper about them, namely, whether they are pagan or Christian in their character. Kierkegaard is using "care" primarily as an emotion, and the word may be better translated anxiety or worry. This is made clear by the context of the particular discourses, in which Kierkegaard tells us for example that the pagan gets up every day and worries about what he is going to eat, whereas the Christian wakes up and praises God for what he has. This indicates that there is a connection between a person's emotions and her concerns.

The discourse on "The Care of Poverty" discusses the way in which the bird does not have this care of anxiety over her provision. Kierkegaard draws a distinction between an internal and external condition, between the agent's circumstance and how the agent views or is affected by those circumstances. Kierkegaard says that "if one is to judge according to its (the bird's) external condition one must call it poor, and yet it is not poor…"[149] How is it that the bird can look poor and yet not be poor? "It meets the

[149] *Christian Discourses; The Crisis and a Crisis in the Life of an Actress*, edited and translated with introduction and notes by Howard V. Hong and Edna H. Hong (Princeton, N.J. : Princeton University Press, 1997), 14.

condition of poverty, but does not have the *care* of poverty."[150] The condition is external, but not external as in something that the bird manifests, such as an action or attitude. The condition is the bird's *circumstances.* The bird does not have a storehouse or a savings account, and thus the bird is economically poor. But poverty in the way that Kierkegaard is using the term consists in far more than one's circumstances. Poverty is an *external state* (circumstances) with a corresponding internal state; a *self-understanding* (beliefs, commitments, etc.) that yields or gives rise to specific *cares* or emotions, in this case anxiety over what one will eat day to day. In other words, "[w]hile the poor and rich Christian know about their poverty or wealth, they are not impressed with their condition."[151] Kierkegaard believes that we can discover a person's character by looking at his cares or emotions, and that these emotions reveal that the person in question has a certain mindset that includes beliefs and commitments.

Kierkegaard continues this theme in the discourse entitled "The Care of Abundance." He states that the care of abundance or wealth is actually an ironic care, since many people who find themselves in the external condition of abundance would actually be free from care.[152] However, the fact is that abundance almost always occasions its own anxiety, the anxiety over acquiring/maintaining/increasing one's abundance.[153] Again Kierkegaard argues that the external condition alone is not

---

[150] Ibid, 14 (my emphasis).

[151] Roberts, "Existence, Emotion, and Virtue," 194.

[152] It is made clear in this context that care would be better understood as worry or anxiety.

[153] Perhaps it is more accurate to say that possessing an abundance of money creates the desire to increase or maintain the abundance of money along with the corresponding fear or anxiety that one will not be able to increase or maintain the abundance of money. Most or all of these cares seem to be rooted in a desire for self-sufficiency, and that desire creates certain specific goals for allowing self-sufficiency, all of which are fleeting because no one can be truly self-sufficient according to a Christian understanding.

sufficient to produce the corresponding emotion, because the bird is free from the care of abundance yet is in a sense rich because the bird has the entire world at its disposal. For example, the bird does not have the care of abundance because the bird does not know that it possesses anything. Therefore the way to have the external condition of richness without having the corresponding care of abundance is to be *ignorant* of one's abundance. This is easy for the bird, just as it was easy for it to be ignorant of its poverty. But for a human being it is very difficult to be ignorant of one's circumstances, and it does not happen naturally; this is an *acquired* ignorance. In order to become like the bird, the Christian must change his mindset, and this takes time and effort.

Kierkegaard acknowledges that changing one's mindset is very difficult and will not happen overnight. In a very important passage, he tells us that "to become ignorant in this way can take a *long time*, and it is a difficult task before he (the Christian) succeeds, *little by little*, and before he finally succeeds in really becoming ignorant of what he knows, and then in *remaining ignorant*, in *continuing* to be that, so that he does not sink back again, trapped in the snare of knowledge."[154] This slow process will require strategies elucidated in the discourses, all of them thought or "seeing-as" strategies that can help a person slowly acquire this ignorance. One strategy is very similar to what the Desert Fathers referred to as "The Remembrance of Death."[155] "That is how ignorant he is of this earthly wealth—he became and remains ignorant by having become aware of something else totally different (because by becoming aware of something else one

[154] Ibid, 26. My emphasis.

[155] John Climacus in his *Ladder of Divine Ascent* names the Remembrance of death as the sixth step on the ladder. Here is one of several relevant passages: "Someone has said that you cannot pass a day devoutly unless you think of it as your last. Even the Greeks have said some such thing, because they describe philosophy as a meditation on death."

becomes ignorant of what one once knew), namely, that *he could die this very night*, that *wealth essentially cannot be possessed*, and it is entrusted property, that *he himself is a traveler*—that is how ignorant the rich Christian is of his earthly wealth, yes, just like an absent-minded person."[156] I take it that these three strategies are essentially the same. If the Christian is daily reminding herself that she is mortal and could die at any moment, then she will be drilling into her mind the fact that her wealth is not her own, that she does not and cannot possess it because it does not have any permanence. And the fact that she is mortal and could die at any moment also reminds her that she is not of this world, but is simply a traveler passing through.

These "seeing-as" strategies are important aspects of Kierkegaard's responsibility-grounding modest libertarian components. He agrees with Frankfurt that a person's loves (or concerns) cannot be changed through a single heroic act of will. However, there are ways that the agent can come to see or perceive her situation differently through *intentional mental actions*. By engaging in cognitive and linguistic practices a person can reshape her emotional responses to the world. The person can also come to experience different desires. When a person cares about something, she will necessarily have desires either directly about or related to that thing. For instance, if Amy cares about her garden, she will want the garden to flourish, she will not want it to rain so much that her garden is flooded, etc. Some of those desires will be motivations to action. But if Amy decides that her concern for her garden is causing her to ignore things she considers more important, she can think about the reasons why her volunteer work is more important than her garden, and she can decide to spend more time volunteering and

[156] Ibid, 31. My emphasis.

less time gardening. As she does, she may find that her desire to garden becomes less intense and more infrequent and that her enthusiasm for her volunteer work deepens. In this way thoughts can influence desires as well as emotions. [157]

To show how Kierkegaard's *Discourses* aids the reader in moral self-cultivation, it is necessary to develop an account of the role of second-order desires in the development of character. This is turn will contribute to a richer understanding of Kierkegaard's libertarian conception of freedom. One of the main cognitive strategies found in Kierkegaard *Discourses* is the strategy I call *attentiveness to emotion*. The discourses are written to a person who intends (on some level) to be a practicing Christian, and by analyzing the kinds of emotions and thoughts that a pagan and Christian should have in a given circumstance, Kierkegaard can reveal to a person that she is not existing as a Christian because she does not *feel* or *think* as a Christian. In Frankfurtian language, this reader has competing first-order desires; one the one hand, she has a first-order desire to live like a Christian, and may even have this second-order desire (she is after all reading the discourse, and thus *may* be displaying the desire to reinforce her Christian commitment). However, her pagan emotions reveal that she has a competing first-order desire for money, security, etc., coupled with the belief that these provisions with not be provided, and by showing the reader that the beliefs and emotions related to her basic concerns are actually those of a pagan, Kierkegaard reveals to his reader that her second-order volitions (the effectual desires) are actually pagan. As a result,

[157] Kierkegaard also assumes that one cannot be simultaneously aware and unaware of one's wealth (at least in a certain sense), thus the one thought somehow pushes out the other. Here we find another interesting analogy to emotions. Though it is quite possible to have more than one emotion at the same time, it is more unusual to have sustained emotions that are in tension. Again, St. Climacus seems to agree with Kierkegaard: "The man who wants to be constantly reminded of death and of God's judgment and who at the same time gives in to material cares and distractions, is like someone trying at the same time to swim and clap his hands."

Kierkegaard increases the listener's knowledge of her emotions as well as her understanding of the way in which her emotions are connected to her thoughts and character (at least in the most reflective listener), but also reinforces (or in some instances may actually create) the desire in the reader to think and exist as a Christian. Even in the instances where this desire already exists in the reader, she now becomes aware that the desire is not effectual.

Furthermore, the perceptive hearer may well feel a kind of meta-anxiety over the fact that her anxieties are pagan. This meta-anxiety might in turn strengthen (or bring to light) her desire to live more like a Christian. But surely if the hearer is already a Christian striving (in some sense) to live a Christian life, our perceptive hearer's anxiety would cause her to search her soul, upon which she would find that she in fact has conflicting desires. She would discover that the desire that she wants to be effectual, namely, the desire to live as a Christian, is in fact one of several competing desires and is *not* her *effectual* desire. This in turn could create in the hearer a second-order desire that her first-order desire to live like a Christian be her will. So hearing the discourse has engendered both a second-order emotion towards her pagan emotions as well as a second-order desire to have her desire to live as a Christian be effectual.

Kierkegaard is causing the hearer to reflect on her emotions and desires and through that reflection to create a second-order desire. This may only frustrate the reader over her emotions and desires. That is why it is important that Kierkegaard is also clarifying concepts for the Christian as well as giving reflective strategies that the hearer can employ to help her achieve the desires and emotions that reflect the Christian character that she wishes to have. When he clarifies concepts for the hearer, Kierkegaard

is giving his hearer the categories in which to reframe her situation, and thus perhaps experience the right kinds of emotions. But he is also giving his hearer practices that she can employ to construe her life in Christian terms, such as the remembrance of death and paying attention to what she says (praising and thanking God for what he has given her instead of despairing over what she lacks). These practices are important because simply creating a second-order emotion or desire in the hearer will not automatically result in an effectual first-order desire.

Finally, I think that this reflective process can result in a kind of *feedback loop*. Because the hearer is now aware of how emotions are based on concerns, and the hearer has been equipped with a second-order desire that she desire eternal things, she will now experience second-order desires and emotions in response to future desires and emotions. Now, when she experiences an emotion that she wants to experience, like gratitude for God's provision, upon reflection this emotion might very well generate another emotion of joy or gratitude that she is loving eternal things; but when she experiences an emotion that she does not desire (such as anxiety over economic issues), this emotion, coupled with the new construal that she has, will generate an emotion of anger or guilt or disappointment that may motivate her anew to think and act as a Christian would, and to employ the disciplines that will help enable this. When she feels the desire to acquire more money, upon reflection she may have a second-order desire that this desire not be her will (and thus that it be eradicated), but instead desire that her desire to be satisfied with what God provides be effectual.

Conclusion

The preceding account demonstrates that Kierkegaard provides the necessary conceptual tools to construct an adequate account of the freedom-relevant conditions necessary for grounding moral responsibility. In the process three unique contributions are discernable. First, it has been clearly demonstrated that the unique contribution of semicompatibilism to debates surrounding free will and moral responsibility is the distinction that semicompatibilists make between cases where an agent's action results from external forces such as manipulation, compulsion, phobias, etc., and cases where the agent demonstrates the freedom necessary for responsibility ascriptions. When framed in this way, the contributions and limitations of semicompatibilism are revealed. While semicompatibilism helps to isolate the freedom-relevant components of morally responsible action, it fails to demonstrate that those components are compatible with determinism, which is the claim it openly champions. This leads to the second major contribution of this project. The components that are necessary to properly distinguish actions resulting from manipulation and free and responsible actions only make sense within a broadly libertarian framework, and are found in Kierkegaard's writings. Turning to Kierkegaard was necessary because his view is unique and not fully compatible with any contemporary view. This leads to the third major contribution. I have systematically interpreted Kierkegaard's understanding of free will and moral responsibility and located his view in contemporary analytic categories. This has helped both to clarify Kierkegaard's writings and to provide a unique and compelling answer to a contemporary metaphysical and metaethical question.

Due to the limitations of this project, there are several areas left for future exploration. First, due to Kierkegaard's vast authorship one chapter is not sufficient to give a definitive account of his view of free will and moral responsibility. There are several texts - in particular the Climacus writings and *The Sickness unto Death* - which I did not deal with that should be included in a definitive account. I believe my conclusions would be both backed by such a project and deepened and expanded. Second, due to the transcendental nature of this project, there is much potential for fruitful dialogue between my conclusions and contemporary scientific theories - particularly findings in neuroscience and evolutionary biology - to see if these freedom-relevant components are compatible with recent scientific discoveries.

BIBLIOGRAPHY

Berofsky, Bernard. 2003. "Classical Compatibilism: Not Dead Yet." In *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities.* Edited by David Widerker and Michael McKenna. Burlington, VT: Ashgate.

Bratman, Michael E., John Martin Fischer, and Mark Ravizza. 2000. "Fischer and Ravizza on Moral Responsibility and History." *Philosophical and Phenomenological Research.* 61 (2).

Bratman, Michael. 2002. "Hierarchy, Circularity, and Double Reduction." In *Contours of Agency: Essays on Themes from Harry Frankfurt*. Edited by Sarah Buss and Lee Overton. Cambridge, Mass: MIT Press, 65-85.

———. 2006. "A Thoughtful and Reasonable Stability: A Comment on Harry Frankfurt's 2004 Tanner Lectures" in *Taking Ourselves Seriously & Getting It Right*. Edited by Debra Satz. Stanford, Calif: Stanford University Press, 77-90.

Burgess, Andrew J. 2001. "Patience and Expectancy in Kierkegaard's Upbuilding Discourses 1843-44." Edited by Niels Jørgen Cappelørn, Hermann Deuser, and Jon Stewart. *Kierkegaard Studies: Yearbook 2000.* Berlin: Walter de Gruyter, 205-222.

Buss, Sarah, and Lee Overton. *Contours of Agency: Essays on Themes from Harry Frankfurt*. Cambridge, Mass: MIT Press, 2002.

Campbell, David M. A. 2000. "Kierkegaard, Freedom, and Self-Interpretation." In *Kierkegaard and Freedom*. Edited by James G. Giles. Houndmills, Basingstoke, Hampshire: Palgrave.

Campbell, Joseph Keim, Michael O'Rourke, and David Shier. *Freedom and Determinism*. Cambridge, Mass: MIT Press, 2004.

Clarke, Randolph. 2003. *Libertarian Accounts of Free Will*. New York: Oxford University Press.

Cooper, David Edward. 1990. *Existentialism: A Reconstruction.* Oxford, UK: Blackwell.

Davenport, John J. 2000. "'Entangled Freedom': Ethical Authority, Original Sin, and Choice in Kierkegaard's Concept of Anxiety." *Kierkegaardiana* 21, 131-151.

———. 2001. “Kierkegaard, Anxiety, and the Will.” *Kierkegaard Studies. Yearbook.* Berlin: W. de Gruyter, 158-181.

———. 2001. “The Meaning of Kierkegaard’s Choice between the Esthetic and the Ethical: A Response to MacIntyre.” *Kierkegaard after MacIntyre: Essays on Freedom, Narrative, and Virtue*. Edited by Davenport and Anthony Rudd. Chicago: Open Court, 75-112.

———. 2001. "Towards an Existential Virtue Ethics: Kierkegaard and MacIntyre.” *Kierkegaard after MacIntyre: Essays on Freedom, Narrative, and Virtue*. Edited by Davenport and Anthony Rudd. Chicago: Open Court, 265-323.

———. 2002. "Fischer and Ravizza on Moral Sanity and Weakness of Will." *The Journal of Ethics.* 6, no. 3: 235-259.

———. 2007. *Will as Commitment and Resolve: An Existential Account of Creativity, Love, Virtue, and Happiness.* New York: Fordham University Press.

Ekstrom, Laura Waddell. 1998. "Protecting Incompatibilist Freedom." *American Philosophical Quarterly.* 35 (3): 281-291.

Evans, C. Stephen. 1983. *Kierkegaard's "Fragments" and "Postscript": The Religious Philosophy of Johannes Climacus.* Atlantic Highlands, N.J.: Humanities Press.

———. 2006. “Kierkegaard’s View of the Unconscious.” *Kierkegaard on Faith and the Self Collected Essays.* Provost series. Waco: Baylor University Press, 277-300.

———. 2006. “Where There’s a Will There’s a Way: Kierkegaard’s Theory of Action.” *Kierkegaard on Faith and the Self Collected Essays.* Provost series. Waco: Baylor University Press, 313-30

Ferreira, M. Jamie. 1991. *Transforming Vision: Imagination and Will in Kierkegaardian Faith.* Oxford: Clarendon Press.

———. 1998. "Faith and the Kierkegaardian Leap." *Cambridge companion to Kierkegaard*. Edited by Alastair Hannay and Gordon Daniel Marino. Cambridge: Cambridge University Press.

Fischer, John Martin, and Mark Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.

Fischer, John Martin. 1994. *The Metaphysics of Free Will*. Oxford, UK: Blackwell.

———. 1999. "Recent Work on Moral Responsibility." *Ethics.* 110 (1): 93-139.

———. “Frankfurt-Style Compatibilism.” In *Contours of Agency: Essays on Themes from Harry Frankfurt*. Edited by Buss, Sarah, and Lee Overton, 1-26. Cambridge, Mass: MIT Press, 2002.

———. “Responsibility and Agent Causation.” 2006. *My Way: Essays on Moral Responsibility*. New York: Oxford University Press.

———. “Responsibility and Alternate Possibilities.” 2006. *My Way: Essays on Moral Responsibility*. New York: Oxford University Press.

———. “Responsibility and Self-Expression.” 2006. *My Way: Essays on Moral Responsibility*. New York: Oxford University Press.

———. “Responsiveness and Moral Responsibility.” 2006. *My Way: Essays on Moral Responsibility*. New York: Oxford University Press.

———. 2007. *Four Views on Free Will*. Malden, MA: Blackwell Pub.

———. 2007. “Compatibilism,” in *Four Views on Free Will*. Great debates in Philosophy. Malden, MA: Blackwell, 44-84.

———. 2010. "Precis of My Way: Essays on Moral Responsibility." *Philosophy and Phenomenological Research.* 80 (1): 229-241.

Flanagan, Owen, 1996. *Self-Expression: Mind, Morals, and the Meaning of Life.* New York: Oxford University Press.

Frankfurt, Harry G. 1988. *The Importance of What We Care About: Philosophical Essays.* Cambridge: Cambridge University Press.

———. 1999. *Necessity, Volition, and Love.* Cambridge, U.K.: Cambridge University Press.

———. 2004. *The Reasons of Love.* Princeton, N.J.: Princeton University Press.

Frankfurt, Harry G., and Debra Satz. 2006. *Taking Ourselves Seriously & Getting It Right*. Stanford, Calif: Stanford University Press.

Grant, W. Matthews. 2010. “Can Libertarians Hold that Our Free Acts are Caused by God?” *Faith and Philosophy 27:1*: 22-37.

Giles, James. 2000. "Kierkegaard's Leap: Anxiety and Freedom." *Kierkegaard and Freedom.* Edited by James Giles. Houndmills, Basingstoke, Hampshire: Palgrave, 69-92.

Glannon, Walter. 1999. "Responsibility and Control: Fischer's and Ravizza's Theory of Moral Responsibility." *Law and Philosophy.* 18 (2): 187-213.

Harned, David Baily. 1997. *Patience: How We Wait upon the World.* Cambridge, Mass: Cowley Publications.

Irvine, William Braxton. 2006. *On Desire: Why We Want what We Want.* Oxford: Oxford University Press.

Jackson, Timothy P. 1998. "Arminian Edification: Kierkegaard on Grace and Free Will." *The Cambridge Companion to Kierkegaard*. Edited by Alastair Hannay and Gordon Daniel Marino. Cambridge: Cambridge University Press.

Kane, Robert. 1996. *The Significance of Free Will.* New York: Oxford University Press.

———. 2004. "Agency, Responsibility, and Indeterminism: Reflections on Libertarian Theories of Free Will." In *Freedom and Determinism.* Edited by Campbell, Joseph Keim, Michael O'Rourke, and David Shier. Topics in contemporary philosophy. Cambridge, Mass: MIT Press.

———. 2005. *A Contemporary Introduction to Free Will*. New York: Oxford University Press.

Kierkegaard, Søren. 1941. *Kierkegaard's Concluding Unscientific Postscript.* Trans David F. Swenson, and Walter Lowrie. Princeton: Princeton University Press.

———. 1956. *Purity of Heart is to Will One Thing: Spiritual Preparation for the Office of Confession.* Translated by Douglas Van Steere. New York: Harper.

———. 1967-78. *Søren Kierkegaard's Journals and Papers,* 7 vols. Edited and translated by Howard V. Hong and Edna H. Hong, assisted by Gregor Malantschuk; with index, vol. 7, by Nathaniel Hong and Charles Barker. Bloomington: Indiana Press.

———. 1978. *Two Ages: The Age of Revolution and the Present Age: A Literary Review.* Edited and translated by Howard V. Hong and Edna H. Hong. Princeton, N.J.: Princeton University Press.

———. 1980. *The Concept of Anxiety: A Simple Psychologically Orientating Deliberation on the Dogmatic Issue of Hereditary Sin*. Edited and translated by Reidar Thomte. Princeton University Press: Princeton, N.J.

———. 1980. *The Sickness Unto Death.* Edited and translated by Howard V. Hong and Edna H. Hong. Princeton, N.J.: Princeton University Press.

———. 1983. *Philosophical Fragments* and *Johannes Climacus.* Edited and translated by Howard V. Hong and Edna H. Hong. Princeton, N.J.: Princeton University Press.

———. 1987. *Either/Or II.* Edited and translated by Howard V. Hong and Edna H. Hong. Princeton, N.J.: Princeton University Press.

———. 1990. *Eighteen Upbuilding Discourses.* Edited and translated by Howard V. Hong and Edna H. Hong. Princeton, N.J.: Princeton University Press.

———. 1992. *Concluding Unscientific Postscript to Philosophical Fragments*, 2 vols. Edited and translated by Howard V. Hong and Edna H. Hong. Princeton, N.J.: Princeton University Press.

Kosch, Michelle. 2000. "Freedom and the Immanence." *Kierkegaard and Freedom*. Edited by James G. Giles. Houndmills, Basingstoke, Hampshire: Palgrave, 121-41.

———. 2006. *Freedom and Reason in Kant, Schelling, and Kierkegaard.* Oxford: Clarendon Press.

MacIntyre, Alasdair C. 1981. *After Virtue: A Study in Moral Theory*. Notre Dame, Ind: University of Notre Dame Press.

———. 2001. "Once More on Kierkegaard." In *Kierkegaard After MacIntyre: Essays on Freedom, Narrative, and Virtue*. Edited by Davenport, John J., Anthony Rudd, 339-56. Chicago: Open Court.

McKenna, Michael. 2001. "Source Incompatibilism, Ultimacy, and the Transfer of Non-Responsibility." *American Philosophical Quarterly.* 38 (1): 37-51.

———. 2001. "Review of Responsibility and Control: A Theory of Moral Responsibility." *The Journal of Philosophy.* 98 (2): 93-100.

———. 2008. "A Hard-line Reply to Pereboom's Four-Case Manipulation Argument." *Philosophy and Phenomenological Research.* 77 (1): 142-159.

Mehl, Peter J. 2005. *Thinking through Kierkegaard: Existential Identity in a Pluralistic World*. Urbana: University of Illinois Press.

Mele, Alfred R. 1995. *Autonomous Agents: From Self-Control to Autonomy*. New York: Oxford University Press.

———. “Soft Libertarianism and Flickers of Freedom.” In *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*. Edited by Widerker, David, and Michael McKenna. 2003. Burlington, VT: Ashgate.

Nassif, Bradley L. “*Toward a 'Catholic' Understanding of St. Augustine's View of Original Sin*.” Union Seminary Quarterly Review 39:4 (1984) 287-299.

O’Connor, Timothy. “The Agent as Cause.” *Free Will.* Blackwell Readings in Philosophy, 3. Edited by Robert Kane. Malden, MA: Blackwell Publishers, 196-205.

———. 2005. "Freedom with a Human Face." *Midwest Studies in Philosophy*, 29: 207–27.

———. 2005. “Free Will,” in *Free Will*, *Vol.I. Concepts and Challenges.* Edited by John M. Fischer. Abingdon: Routledge Press, 7-20.

Pereboom, Derk. 1995. "Determinism al Dente." *Nous.* 29 (1): 21-45.

———. 2003. “Source Incompatibilism and Alternative Possibilities.” In *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*. Edited by Widerker, David, and Michael McKenna, 185-200. Burlington, VT: Ashgate.

———. 2005. "Defending Hard Incompatibilism." *Midwest Studies In Philosophy.* 29 (1): 228-247.

———. 2007. "Hard Incompatibilism." In *Four Views on Free Will*, by John Martin Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas, 85–125. Oxford: Blackwell.

Pillar, Christian. 2008. “Morality’s Place: Kierkegaard and Frankfurt.” *Revista Porteguesa de Filosofia* 64: 1207-1219.

Pink, Thomas. 2004. *Free Will: A Very Short Introduction.* Very Short Introductions. Oxford: Oxford University Press.

Pojman, Louis P. 1984. *The Logic of Subjectivity: Kierkegaard's Philosophy of Religion.* University: University of Alabama Press.

Roberts, Robert Campbell. 1984. “Passion and Reflection.” *International Kierkegaard Commentary 14: Two Ages.* Edited by Robert L. Perkins. Macon, GA: Mercer Univ. Press.

———. 1995. “Kierkegaard, Wittgenstein, and a Method of ‘Virtue Ethics.’” *Kierkegaard in Post/Modernity.* Studies in Continental thought. Edited by Martin Joseph Matuštík and Merold Westphal. Bloomington: Indiana University Press, 142-66.

———. 1997. "Dialectical Emotions and the Virtue of Faith." *International Kierkegaard Commentary 12: Concluding Unscientific Postscript to Philosophical Fragments*. Edited by Robert L. Perkins. Macon, Ga., USA: Mercer University Press, 73-93.

———. 1998. “Existence, Emotion, and Virtue: Classical Themes in Kierkegaard.” *The Cambridge companion to Kierkegaard*. Edited by Alastair Hannay and Gordon Daniel Marino. Cambridge: Cambridge University Press.

———. 2003. “The Virtue of Hope in Eighteen Upbuilding Discourses.” *International Kierkegaard Commentary: Eighteeen Upbuilding Discourses*. Edited by Robert B. Perkins (Macon: Mercer University Press), 181-203.

———. 2003. *Emotions: An Essay in aid of Moral Psychology.* Cambridge, UK: Cambridge University Press.

Robertson, John. 2001. “Internalism, Practical Reason, and Motivation.” In *Varieties of Practical Reasoning*. Edited by Millgram, Elijah, 127-52. Cambridge, Mass: MIT Press.

Searle, John R., 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.

Sorabji, Richard. 1980. *Necessity, Cause, and Blame: Perspectives on Aristotle's Theory*. Ithaca: Cornell University Press.

Stump, Eleonore. 1988. "Sanctification, Hardening of the Heart, and Frankfurt's Concept of Free Will." *The Journal of Philosophy.* 85 (8): 395-420.

———. 1996. "Persons, Identification, and Freedom." *Philosophical Topics* 24:183-214.

———. 1996. "Libertarian Freedom and the Principle of Alternative Possibilities.” In *Faith, Freedom, and Rationality: Philosophy of Religion Today*. Edited by Jordan, Jeff, and Daniel Howard-Snyder, 73-88. Lanham, Md: Rowman & Littlefield.

———. 1999. "Alternative Possibilities and Moral Responsibility: The Flicker of Freedom." *The Journal of Ethics.* 3 (4): 299-324.

———. 2001. "Augustine on Free Will." In *The Cambridge companion to Augustine*. Edited by Stump, Eleonore, and Norman Kretzmann, 124-147. Cambridge, UK: Cambridge University Press.

———. 2003. "Freedom: Intellect, Action, and Will." In *Aquinas*, 277-305. London: Routledge. http://sites.google.com/site/stumpep/onlinepapers.

Taylor, Charles. 1985. *Human Agency and Language: Philosophical Papers I.* Cambridge: Cambridge University Press.

Timpe, Kevin. 2006. "A Critique of Frankfurt-Libertarianism." *Philosophia.* 34 (2): 189-202.

———. 2007. "Source Incompatibilism and Its Alternatives." *American Philosophical Quarterly.* 44 (2): 143-55.

———. 2008. *Free will: Sourcehood and its Alternatives*. London: Continuum.

Van Inwagen, Peter. 1983. An Essay on Free Will. Oxford: Clarendon Press.

———. 2004. "Van Inwagen on Free Will." In *Freedom and Determinism.* Edited by Campbell, Joseph Keim, Michael O'Rourke, and David Shier. Topics in contemporary philosophy. Cambridge, Mass: MIT Press, 213-230.

Watson, Gary. 1975. "Free Agency." *The Journal of Philosophy.* 72, no. 8: 205-220.

———. 1987. "Free Action and Free Will." *Mind.* 96, no. 382: 145-172.

———. 2002. "Volitional Necessities," in *Contours of Agency: Essays on Themes from Harry Frankfurt*. Edited by Buss, Sarah, and Lee Overton. Cambridge, Mass: MIT Press.

———. 2004. "Soft Libertarianism and Hard Compatibilism." *Agency and Answerability: Selected Essays.* Oxford: Clarendon Press.

———. 2001. "Reasons and Responsibility." *Ethics.* 111 (2): 374-394.

Weston, Michael. 2000. "Kierkegaard and the Literature of Freedom." *Kierkegaard and Freedom*. Edited by James G. Giles. Houndmills, Basingstoke, Hampshire: Palgrave.

Widerker, David, and Michael McKenna. 2003. *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*. Burlington, VT: Ashgate.

Wolf, Susan. 1980. "Asymmetrical Freedom." *The Journal of Philosophy* 77:3, 151-66.

———. 2002. "Sanity and the Metaphysics of Responsibility." *Free Will.* Blackwell Readings in Philosophy, 3. Edited by Robert Kane. Malden, MA: Blackwell Publishers, 145-63.

———. 2002. "The True, the Good, and the Lovable: Frankfurt's Avoidance of Objectivity," in *Contours of Agency: Essays on Themes from Harry Frankfurt*. Edited by Buss, Sarah, and Lee Overton. Cambridge, Mass: MIT Press.