

2003

## The Need for Autonomy

Paul Jude Naquin

*Louisiana State University and Agricultural and Mechanical College*

Follow this and additional works at: [https://digitalcommons.lsu.edu/gradschool\\_theses](https://digitalcommons.lsu.edu/gradschool_theses)



Part of the [Arts and Humanities Commons](#)

---

### Recommended Citation

Naquin, Paul Jude, "The Need for Autonomy" (2003). *LSU Master's Theses*. 3685.  
[https://digitalcommons.lsu.edu/gradschool\\_theses/3685](https://digitalcommons.lsu.edu/gradschool_theses/3685)

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact [gradetd@lsu.edu](mailto:gradetd@lsu.edu).

# THE NEED FOR AUTONOMY

A Thesis

Submitted to the Graduate Faculty of the  
Louisiana State University and  
Agricultural and Mechanical College  
in partial fulfillment of the  
requirements for the degree of  
Master of Arts

in

The Department of Philosophy

by  
Paul Jude Naquin  
B.A., Louisiana Tech University, 2000  
May 2003

## **ACKNOWLEDGEMENTS**

I would like to thank all of the people who were kind enough to read or listen to my arguments and help me clarify the ideas that follow. In particular, I would like to thank my committee members, Husain Sarkar and James Spence; as well as Harry Frankfurt, Jason Megill, Gabe Cate, and William Osborne. Special recognition is due to my thesis director James Taylor for his unending patience and incredible wealth of knowledge, without which I could not have even pursued this project. Finally, I wish to thank Stacy Todd, who has been both a keen critic and the mainstay of my sanity during the writing of this thesis. You all have my deepest gratitude.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	ii
ABSTRACT.....	iv
CHAPTER 1. THE PROJECT.....	1
1.1 Nature of the Beast.....	1
1.2 Amorphous Ideas.....	2
1.3 Theoretical Underpinnings.....	5
CHAPTER 2. THEORETICAL HISTORY.....	8
2.1 Evolution of a Paradigm.....	8
2.2 Variation on a Paradigm.....	21
CHAPTER 3. WHAT IS NEEDED.....	26
3.1 <i>Ab Initio</i> in Depth.....	26
3.2 The Need for a Need.....	31
CHAPTER 4. THE SOLUTION.....	38
4.1 The Need that Satisfies a Need.....	38
4.2 Getting the Job Done.....	42
4.3 Conclusions and Implications.....	46
REFERENCES.....	49
VITA.....	51

## ABSTRACT

Autonomy researchers over the last three decades have largely focused on the hierarchical, content-neutral theories proposed by Harry Frankfurt and, to a lesser degree, Gerald Dworkin. Both of these theories claim that one must have higher-order endorsement of her lower-order desires to be autonomous with respect to the lower-order desires. However, neither of these theories makes the claim that one must be autonomous with respect to the higher-order endorsing desire. This leads to a dilemma known as the *ab initio* problem. Specifically, the problem is that it is not clear how one can become autonomous with respect to one desire by appealing to a desire to which one does not bear an autonomous relation.

In this essay, I attempt to argue that the *ab initio* problem can be solved by modifying the currently content-neutral theories to be instead substantive. In other words, I claim that a hierarchical account of autonomy must appeal to a theoretically-specified (substantive) mental state. I believe that the solution to the *ab initio* problem is to appeal to a need for self-worth as the appropriate mental state. The need for self-worth can be used to explain how any individual identifies with her desires because an individual cannot rationally pursue a goal that she believes will damage her overall worth. Therefore, the need for self-worth explains how a person comes to be autonomous with respect to her desires.

## CHAPTER 1. THE PROJECT

### 1.1 Nature of the Beast

Over the last several decades, the nature and application of autonomy has accounted for much debate in moral philosophy. At the heart of these debates have been Harry Frankfurt's and – to a lesser degree – Gerald Dworkin's hierarchical accounts of autonomy. Most contemporary discussions of autonomy focus on criticizing, developing, or applying one of these two theories (or some version of one of them). This essay will endeavor to provide a solution to a particularly persistent problem with both of these theories.

Both Frankfurt and Dworkin are faced with a two-prong problem, originally outlined by Gary Watson. On one hand, these two theories are confronted with a regress problem, wherein an infinite array of desires is necessary to endorse any basic desire. On the other hand, halting the regress of desires introduces an *ab initio* problem, which addresses the issue of how an endorsement from a non-autonomous desire can cause a person to become autonomous with respect to another desire. My goal will be to address – and hopefully solve – the *ab initio* difficulty by positing that people have a basic evolutionary need for self-worth, and that the satisfaction of this need is a necessary condition for autonomy. Additionally, if I succeed in solving the *ab initio* problem, then the regress will not exist. The ultimate goal of this essay, then, is to eliminate both problems.

I shall begin by briefly explaining the general concept of autonomy, as well as the overall problems with that general concept. This will lead us into a close examination of Frankfurt's and Dworkin's theories, as well as some related ideas that are of particular

interest to our discussion. Once I have outlined the important positions, I shall give a close account of how the regress and *ab initio* problems affect each theory. Then I will argue that humans have a basic psychological need to develop self-worth. Lastly, I shall attempt to show that the satisfaction of this need will allow a hierarchical theory of autonomy to avoid the *ab initio* problem.

## 1.2 Amorphous Ideas

Before anything else, I need to briefly explain the nature of the concept of autonomy. However, there is already an immediate problem: the goal of this paper is to explore the nature of autonomy, so I do not yet possess a working definition for it. Moreover, there does not seem to be much unity in the way that philosophers have used the term. In other words, we have little information to aid us in even forming a coherent set of characteristics that define autonomy. Dworkin notices the difficulty and suggests the possibility that there are simply many conflicting concepts, but he rejects this conclusion.

What is more likely is that there is no single conception of autonomy but that we have one concept and many conceptions of autonomy – to make use of a distinction first introduced by H.L.A. Hart and developed by Rawls. The concept is an abstract notion that specifies in very general terms the role the concept plays. Thus, a certain idea of persons as self-determining is shared by very different philosophical positions. Josiah Royce speaks of a person as a life led according to a plan. Marxists speak of man as the creature who makes himself; existentialists of a being whose being is always in question; Kantians of persons making law for themselves. At a very abstract level, I believe they share the same concept of autonomy.<sup>1</sup>

---

<sup>1</sup> Gerald Dworkin, *The Theory and Practice of Autonomy*, (New York, New York: Cambridge University Press, 1988), 9.

Dworkin's conclusion seems like a reasonable starting point, but it still leaves us missing a clear notion of autonomy. However, lacking a solid definition with which I might evaluate any claims, I must offer any more detailed explanations of autonomy only tentatively. Evaluation will therefore be limited to a more tangential approach. In this essay, theories will be assessed based on their internal consistency, as well as their consistency with common reasonable intuitions about autonomy.

It seems that the easiest way to begin examining autonomy is etymologically. Autonomy comes from two Greek words: **αυτος** (autos), which means 'self,' and **νομος** (nomos), which means 'rule.' Thus, we might start by saying that autonomy refers to the ability to rule oneself, *i.e.*, the ability to control one's own actions and desires. Although this does not give us a specific denotation, we can use this foundational understanding of autonomy as a guide for evaluating cases. Since most of our discussion will revolve around applying our intuitions to examples, this should be an excellent starting point.

Before we try to evaluate theories, we still need to refine our general understanding of autonomy, and we can do this through the use of examples. Even though we do not yet have a clear definition to work with, we can still use our general notions to categorize certain individuals as autonomous or non-autonomous with respect to their desires. We can illustrate these general notions by considering comparison cases. First, we may consider two individuals at a party. Each of these people has had a substantial amount of alcohol, and they are both inebriated. One of our subjects – we can call her Courtney – decided before the party that she wanted to become intoxicated, and she has done so with great success. The other person, Jason, is a recovering alcoholic who has been overwhelmed by his desire to drink, and he simply could not restrain



himself. Even with only our general understanding of autonomy, we can see that Jason is not autonomous with respect to his desire to imbibe alcohol because he is not fully in control of his own desires. In contrast, Courtney is autonomous in regards to her desire to drink (even if she regrets it later).

We can also consider autonomy from the view of manipulation or coercion. Imagine that I have three friends, Stacey, Angela, and Francisco, and all three of my friends like strawberry ice cream. My friends also know that I usually have strawberry ice cream in my house, although I do not have any right now. When Stacey visits my house, she immediately goes to the refrigerator to find some strawberry ice cream. Unfortunately for her, there is no ice cream. However, it seems that she is autonomous in relation to her desire to acquire the ice cream. When Angela comes to my house, I lie and tell her that there is strawberry ice cream inside my refrigerator. Like Stacey, Angela walks to the refrigerator and initiates a futile search for ice cream. Unlike Stacey, it seems as if Angela is not autonomous with respect to her desire because I deceived and manipulated her. When Francisco arrives at my house, I point a gun at him and yell that he should immediately search for ice cream. Wisely, Francisco goes to the refrigerator and hunts for ice cream in vain. In this example, each of my friends has formed a desire to seek out ice cream in my refrigerator, but my friends do not appear to be equally autonomous with respect to those desires. Stacey seems to possess full autonomy because her desire has not been manipulated externally.<sup>2</sup> Angela and Francisco seem to

---

<sup>2</sup> It might be argued that some sort of socio-environmental factors may have influenced Stacey's desire, and she is therefore not autonomous with respect to it. I shall address this issue directly later in the essay, but for this example I am assuming that there are no hidden influences undermining her autonomy.

lack it because I have taken control of their desires, at least in part.<sup>3</sup> In Angela's case, I manipulated her belief set so that she would acquire her desire, and Francisco's desire was the direct result of my violent coercion.

It is possible that someone may disagree with me concerning the degree to which each person in my examples is autonomous. Even if this is the case, it only proves that people do have some intuitions about the nature of autonomy. My task is to provide a consistent framework in which those intuitions, or at least the most important ones, can coexist.

### 1.3 Theoretical Underpinnings

From the examples we have just considered, it is apparent that people have some ideas about the characteristics of autonomy, but we need to solidify those notions into a concise and useful denotation. For our specific definitions, there are two major branches of autonomy theory to which we may appeal. The first is the Kantian account of autonomy, which essentially equates autonomy with acting on reason completely independent of any other influences. While the Kantian idea is still held strongly by many ethicists, the second branch of theory, based on Frankfurt's and Dworkin's hierarchical models, has engendered much more debate in recent years.

Although my main foci here are the hierarchical theories, I would like to digress momentarily to discuss theories of autonomy in the Kantian tradition. There are two

---

<sup>3</sup> One could claim that I have taken control Francisco's actions, but not necessarily his desires. "On Frankfurt's hierarchical analysis of what it is for a person to be autonomous with respect to his desires a person who acts under duress (even under duress of a certain degree of harshness) will be autonomous with respect to that first-order desire that moved him to comply with the threat that he was faced with. However, it is clear from 'Coercion and Moral Responsibility' that Frankfurt does not accept that such a person is thereby fully autonomous with respect to his compliant action." Taylor, James Stacey, "Autonomy, Duress, and Coercion," *Social Philosophy and Policy* (Forthcoming), (June 2003), 16-17. However, we need not be concerned about this worry here, as my goal with this example is merely to show that people do have certain intuitions about the fundamental nature of autonomy.

common interpretations of Kant's theory: moral autonomy and personal autonomy. As Christine Korsgaard explains, Kant's own view (moral autonomy) is that autonomy is inextricably linked to morality.

According to Kant, as each impulse to action presents itself to us, we should subject it to the test of reflection, to see whether it really is a *reason* to act. Since a reason is supposed to be intrinsically normative, we test a motive to see whether it is a reason by determining whether we should allow it to be a *law* to us. And we do that by asking whether the maxim of acting on it can be willed as law. ...So the reflective endorsement test is not merely a way of justifying morality. *It is morality itself.*<sup>4</sup>

Korsgaard shows that the Kantian notion of autonomy is strongly tied to morality. However, one may also filter Kant's theory in order to derive a theory of personal autonomy devoid of direct ties to morality. John Christman defines personal autonomy in the Kantian tradition as "the self imposition of laws of reason. ...It is open to the theorist of autonomy to specify morally *neutral* conditions for the autonomous formation of preferences."<sup>5</sup> The common thread in Kantian theories of autonomy is that a person must act on reason without interference from desires or manipulation.

As I remarked earlier, this discussion of Kantian autonomy is only a digression, and a brief one at that. However, it seems reasonable to assume that someone might ask why I am focusing my attention exclusively on non-Kantian theories of autonomy, especially since Kantian theory has such a rich history. There are two responses to this question. The first answer is given by Wright Neely.

Theories of freedom have been offered according to which a man acts freely only when he acts morally or even only

---

<sup>4</sup> Christine M. Korsgaard, *The Sources of Normativity*, (New York, New York: Cambridge University Press, 1996), 89.

<sup>5</sup> John Christman, "Introduction," *The Inner Citadel*, ed. John Christman (New York, New York: Oxford University Press, 1989), 14.

when he acts out of a sense of duty. Philosophers who have seriously put forward this view, however, typically hold other views which require that a man be acting freely when he acts *immorally*, when he succumbs to temptation; thus the view is usually replaced by one according to which a man acts freely only when he acts either morally or immorally. This dialectic can be seen quite clearly in Kant's moral philosophy, and it can be argued that he never really was able to render the second position consistent with the considerations which led him to the first.<sup>6</sup>

Essentially, Neely's point is that if we want to hold a person responsible for her wrong actions, we need to show that she is capable of performing those actions autonomously. If autonomy is inextricably tied to right actions, as it is on a Kantian model, then there is no way to hold people responsible for their wrongdoings.

The second answer to this question is simply one of metaphysics. On a Kantian metaphysical view, it will be unsurprising that Kantian autonomy makes perfect sense. Nonetheless, on the Humean view, which I espouse, Kant's theory is incoherent. While Kantian autonomy requires strict adherence to reason, David Hume takes the stance that "Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them."<sup>7</sup> Admittedly, one may contend that the Humean metaphysical picture is flawed, but providing counterarguments to that contention would carry us too far from the topic of this paper. (In fact, that argument could be a separate book unto itself.) Moreover, the theorists that I am considering all adopt a Humean stance, and this should be sufficient reason to adopt also the Humean metaphysical model here.

---

<sup>6</sup> Wright Neely, "Freedom and Desire," *The Philosophical Review* 83, No. 1 (1974): 35.

<sup>7</sup> David Hume, *A Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton (New York, New York: Oxford University Press, 2000), 266.

Given that the Kantian model of autonomy is untenable on Hume's metaphysical view, there is only one major set of autonomy theories available: the hierarchical models, which are written from a Humean perspective. Frankfurt and Dworkin posited the two primary hierarchical theories in the 1970s, and these theories have been the mainstay of research in autonomy since then. As these theories will also be the focal point of this paper, I shall proceed with the argument by explaining them in detail.

Unfortunately, before I give such an explanation, I need to digress one more time from the specific topic. As both of the theories at the core of this study are concerned with the interaction of desires, it would be useful to develop a working philosophical definition of the term 'desire.' In order to do this, I will invoke the aid of Professor Neely. "'Desire' in this sense... is meant to apply to everything from fleeting whims to certain kinds of abiding character traits, from primary appetites to self-love, and from any of these to moral convictions. Desires also include (or overlap with) such things as purposes, intention, volitions, and (in some senses) motives."<sup>8</sup> Neely admits that this denotation suffers from some degree of vagueness, but it should suffice for my purposes. With that clear, I can return to the hierarchical theories of autonomy.

---

<sup>8</sup> Neely, 33.

## CHAPTER 2. THEORETICAL HISTORY

### 2.1 Evolution of a Paradigm

Let me begin with Frankfurt. He describes autonomy as a relation between three types of internal desires. A ‘first-order desire’ is simply a desire either to execute or refrain from some action. This is the simplest form of desire, and it seems likely that many sentient creatures are capable of possessing first-order desires. A ‘second-order desire’ either endorses or repudiates a particular first-order desire. Second-order desires are formed when an agent reflects on which desires she wants to approve and which she wants to reject. Frankfurt also distinguishes a special type of second-order desire that is absolutely vital to his theory. A ‘second-order volition’ is a desire, not only to have a first-order desire, but also to have that first-order desire move the agent to action. For Frankfurt, a second-order volition is indicative of a person’s will. “[The will] is the notion of an *effective* desire – one that moves (or will or would move) a person all the way to action.”<sup>9</sup> Though this understanding of the will may seem unusual, it is neither new nor accepted solely by followers of Frankfurt; it originates with Thomas Hobbes. “In *deliberation*, the last appetite, or aversion, immediately adhering to the action, or to the omission thereof, is that we call the WILL; the act, (not the faculty,) of *willing*.”<sup>10</sup> For Frankfurt, then, a person’s will is her effective first-order desire, and she is autonomous to the extent that her will conforms to her second-order volitions.

Let us consider a brief example to clarify the relationship between these types of desires. Imagine that someone is preparing to attack you, and you have two first-order

---

<sup>9</sup> Harry Frankfurt, “Freedom of the Will and the Concept of a Person,” *The Importance of What We Care About*, (New York, New York: Cambridge University Press, 1988), 14.

<sup>10</sup> Thomas Hobbes, *Leviathan*, ed. J. C. A. Gaskin (New York, New York: Oxford University Press, 1996), 40.

desires: (a) to retaliate, and (b) to run away. It seems reasonable to assume that most animals are capable of such basic desires, as are all humans – with the likely exception of people suffering from certain types of brain dysfunction. Now, add to the example the assumption that you are a pacifist, and you have a second-order desire to want to avoid fighting. At the same time, you do not possess a second-order desire to retaliate against your attacker because you are a pacifist. At first glance, it might seem that this is sufficient for an account of autonomy; one should simply act on one's second-order desire.

Nonetheless, Frankfurt foresees two quandaries with a model using only first and second-order desires. The first difficulty is the possibility of 'wantons,' or beings that have first-order desires, and possibly even second-order desires, but do not possess second-order volitions. How could an individual have second-order desires while lacking second-order volitions? A wanton might rationally consider which of his first-order desires are strongest in order to decide which desires to satisfy and how to satisfy them. However, the wanton would not consider whether or not he *should* want to comply with his strongest desires. "What distinguishes the rational wanton from other rational agents is that he is not concerned with the desirability of his desires themselves. ...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."<sup>11</sup> In addition to the problem of wantons, the absence of volitions would leave no way to adjudicate between conflicting second-order desires.

Frankfurt addresses both of these complications by introducing second-order volitions into his theory. In order to see how this works, we can modify our example.

---

<sup>11</sup> Frankfurt, "Freedom of the Will," 17.

You are a pacifist once again, but the assailant is attacking your child rather than you. You have the same first-order desire to avoid conflict, as well as the same second-order desire to have that first-order desire. However in this case, you also have a first-order desire to protect your child, and a corresponding second-order desire to want to have the desire to protect your child. In this scenario, you must choose between conflicting second-order desires. We shall assume that you want the desire to protect your child to move you to action, even if you must engage in violence in order to do so. In choosing this desire, you have formed a second-order volition to protect your child, so you are autonomous with respect to your desire to save the child.

In describing his theory, Frankfurt takes note of a possible regress problem. One might ask if higher-order desires and volitions are necessary to confirm that an individual truly endorses any given second-order desire. The fact is that it often makes sense to say to someone: “Tell me what you want, what you really, *really* want.”<sup>12</sup> If a person must confirm each second-order desire by appealing to a higher-order desire, then of course an infinite regress appears inevitable. Frankfurt believes that “it is possible, however, to terminate such a series of acts without cutting it off arbitrarily. When a person identifies himself *decisively* with one of his first-order desires, this commitment ‘resounds’ throughout the potentially endless array of higher orders.”<sup>13</sup> Thus, according to Frankfurt, autonomy seems to be ultimately dependent upon resolute identification with a first-order desire.

The nature of Frankfurt’s decisive identification can best be seen through an example. Imagine that someone – I will call him Andrew – is in a restaurant. While

---

<sup>12</sup> Spice Girls, “Wannabe,” *Spice*, (Los Angeles, California: Virgin Records, 1996), [Compact Disc]: track 1. I apologize for making this reference, but as James Taylor has pointed out, the description is all too apt.

<sup>13</sup> Frankfurt, “Freedom of the Will,” 21.



looking at the menu, Andrew realizes that he has a desire to eat a rare steak. When Andrew comes to this realization, he also begins to wonder whether he truly wants steak, or if he is merely about to act on a random first-order desire. To solve the problem, Andrew reflects on his desires for various types of food, and he forms a second-order volition endorsing his first-order desire for steak. In other words, Andrew started out merely wanting steak. Now he also wants to want steak, and he wants his desire for steak to move him to act. Regrettably, Andrew finds himself unsatisfied; he begins to wonder whether his second-order volition is genuine. In order to guarantee his second-order volition, he might be inclined to form a third-order volition. Obviously, Andrew can repeat this exercise indefinitely, and a regress begins to threaten. It is at this point that decisive identification becomes important. Frankfurt points out that people do not actually go through such a process of forming legions of higher-order desires and volitions to guarantee their autonomy. Instead, individuals form second-order volitions whose genuineness need not be questioned at all. In the example, Andrew would not look for confirmation for his second-order volition; he would merely form a volition with which he decisively identifies. “The decisiveness of the commitment he has made means that he has decided that no further question about his second-order volition, at any higher order, remains to be asked.”<sup>14</sup> It seems that Andrew can eat his rare steak in peace because he will, at some point, become disinterested in rechecking the genuineness of his desire.

However, a setback appears when we consider the obvious fact that people are often indecisive. Frankfurt’s requirement of wholehearted decisiveness seems to classify people as non-autonomous whenever they confront a truly difficult decision. In fact, the

---

<sup>14</sup> Frankfurt, “Freedom of the Will,” 22.

strong decisiveness that Frankfurt speaks of is not commonly experienced outside of relatively trivial cases of decision making, and it would seem extremely odd to classify a person as lacking in autonomy every time she makes a momentous decision which leaves her with some concerns about her decision. Incidentally, I should make a distinction here between having concerns about the consequences of a decision and concerns about the decision itself. One may be wholehearted even if one has lingering worries about possible ill-fated outcomes of a decision. However, wholeheartedness is not compatible with one's being uncertain about whether one truly identifies with the decision. This would render many people non-autonomous when they make some of the most important decisions in their lives, such as career choices. For instance, my desire to become a professional philosopher is inconsistent with my desires to engage in other careers, such as professional wrestling. (Sadly, unlike in Plato's time, philosophy and wrestling each require too much time for me to pursue both of them simultaneously.) Upon forming my second-order volition to become a philosopher, I must consider the possibility that my volitional endorsement is misplaced, and then it does not seem that my identification is, in fact, wholehearted. If I do not identify decisively with my desire to become a philosopher, then I must choose from two conclusions. The first possibility is that I am simply non-autonomous with respect to my desire to be a professional philosopher. But this conclusion appears unwarranted; the fact that I may still coherently question the genuineness of my desire does not seem to undermine my autonomy with respect to it.

The other option is that the dilemma I have just described is a misunderstanding of Frankfurt based on too harsh a reading of his work. In "Identification and Wholeheartedness," he has an answer to my problem. "If he has made a genuinely

unreserved commitment to the view that adopting the answer is his most reasonable alternative, he can anticipate that *this* view will be endlessly confirmed by accurate reviews of it. The fact that a commitment resounds endlessly *is* simply the fact that the commitment is *decisive*.<sup>15</sup> On this understanding of decisive commitment, my problem dissolves because, presumably, I would decide to become a philosopher every time I consider my career options. However, this definition of decisiveness is dependent upon the subjective mental states of an agent, and that leads directly to two other stumbling blocks.

If a person's autonomy depends solely upon that person's subjective mental states, then it is possible that a person may be manipulated into having the appropriate mental states. Intuitively, manipulation sounds like it will undermine autonomy, but the problem is even more complex than it might first appear. There are two ways in which we might understand manipulation, and I shall take just a moment to differentiate between them here. One can note a difference between agential and nonagential forms of manipulation. Agential manipulation is enacted on someone by another agent(s); it is precisely what most people think of when they hear the word 'manipulation.' Unsurprisingly, nonagential manipulation has no immediate controlling agent(s). Nonagential forms of manipulation might include such things as socialization, socioeconomic standing, or cultural background. In addition, Frankfurt's compatibilism commits him to the view that certain influences on our desires are perfectly licit. In order to deal with this predicament, he needs to offer us an explanation of manipulation that separates acceptable influences from ones that endanger or undermine autonomy.

---

<sup>15</sup> Harry Frankfurt, "Identification and Wholeheartedness," *The Importance of What We Care About*, (New York, New York: Cambridge University Press, 1988), 168.

Aside from the problem of manipulation, we can show that an infinite regress once again threatens Frankfurt's theory. If a person must make a decisive commitment to halt a regress of higher-order desires, then that individual might begin to question her desire to enact the commitment itself. Since it is easily conceivable that someone might ask such a question of herself, Frankfurt has not successfully avoided the infinite regress.

Frankfurt recognizes both of these concerns, and he attempts to address them by modifying his definition of wholeheartedness. To this end, he introduces the notion of 'volitional unity,' the degree to which all of one's effective first-order desires are consistent with one another. This allows him to say that instead of making a decisive commitment, one need only be satisfied with a desire in order to be autonomous with respect to it.<sup>16</sup> He then states that satisfaction "is a matter of simply *having no interest* in making changes."<sup>17</sup> In more specific terms, we might think of satisfaction in terms of competing desires. Initially, a person may have several desires in competition, but at some point, the individual will decide volitionally to endorse one of those desires.<sup>18</sup> If the person is satisfied with her decision, then the competing desires will lose (some of) their motivational power. In this case, the person is satisfied, not only because she volitionally endorses her effective first-order desire, but also because she feels no inclination to reconsider her choice or the reasons that caused her to choose it.

Conceiving of wholeheartedness in this way allows Frankfurt to address the issues of manipulation and regression. Satisfaction helps counteract manipulation by showing

---

<sup>16</sup> Frankfurt is not employing a coherentist view, in which a person achieves autonomy if her desires fit into a coherent framework. He is merely noting the fact that having a consistent desire set allows a person to act on more of her desires, thus giving her the ability to satisfy more of them.

<sup>17</sup> Harry Frankfurt, "The Faintest Passion," *Necessity, Volition, and Love*, (New York, New York: Cambridge University Press, 1999), 105.

<sup>18</sup> It is entirely plausible that an individual might only have one desire in a given situation. If so, then there is no conflict and the person is satisfied by default. However, we can assume that all people will also encounter situations in which they have competing desires.

that some influences on our desires can be allowed as long as they are (generally) consistent with our overall desire sets. This allows a human being to maintain her autonomy even if some of her desires are affected by her culture and upbringing. Moreover, appealing to satisfaction permits Frankfurt to avoid the regress problem altogether because satisfaction does not require any action on the part of the agent. Since satisfaction is simply a state of being disinterested in further investigation of one's desires, one simply *will not* question those desires. Therefore, satisfaction leads to identification with a desire because the satisfied person accepts the desire as her own in a way that is intrinsically beyond question.

Does satisfaction perfect Frankfurt's theory at last? In truth, there are still some issues that need to be considered. First, Frankfurt may have succeeded in showing that some forms of non-agential manipulation may not threaten autonomy, but he has not given us an account of how illicit manipulation does undermine autonomy. Additionally, he has given us no way to adjudicate between licit and illicit manipulation. If some forms of manipulation do not undercut autonomy, then why should others do so? More to the point, what are the vital distinctions between legitimate and illegitimate types of manipulation? The second worry, which I shall only mention at this time, is that Frankfurt still has not explained how his theory can deal with the *ab initio* problem. That is, how can an agent, who is not initially autonomous with respect to any of her desires, become autonomous with respect to some of her desires?

Finally, the third predicament hinges on two possible understandings of the foundational connotations of autonomy. The dilemma is that both understandings of autonomy seem to be problematic for Frankfurt. One of these interpretations is that an

autonomous person directs her own actions. Since Frankfurt's view is that satisfaction with second-order volitions makes a person autonomous with respect to those volitions, an agent can be autonomous with respect to any volition as long as she does not care to question its legitimacy. On this understanding of autonomy, there seems to be no possible way to block out illicit influences on an agent's desires. For example, imagine that I want my friend Dan to play Frisbee golf with me. With my amateur mad scientist kit, I implant a microchip in Dan's brain that gives him a desire to play Frisbee golf. Being a studious follower of Frankfurt, I also bestow upon Dan a second-order volition endorsing his desire to play Frisbee golf, as well as a predisposition to be uninterested in examining his Frisbee golf desires. It seems that my interference in Dan's desires completely usurps his autonomy, yet Frankfurt cannot account for why this is the case if we understand autonomy to be nothing more than self-direction. The other major connotation of autonomy – the one that Frankfurt seems more concerned with – is that an autonomous agent must identify with her desires; that is, she must hold those desires to be an integral part of herself. We can recall Frankfurt's claim that satisfaction with a desire is necessary to ensure that a person identifies wholeheartedly with that desire. But if satisfaction cannot occur until a desire has already been volitionally endorsed, then Frankfurt's view implies that a person can only be autonomous with respect to a desire retroactively. Intuitively, it seems odd to say that a person is not autonomous with respect to a desire when she is volitionally endorsing it, yet she may become autonomous after the fact if she is satisfied with her decision. To put a finer point on this dilemma, let us look at each of the relevant steps in the process. At time  $t_1$ , a person forms a second-order volition endorsing one of her first-order desires. Then, at time  $t_2$ , the person

becomes satisfied with her volition. Thus, the agent is not autonomous in relation to her desire when she volitionally endorses it, but she becomes autonomous when she later achieves satisfaction. As there is no theoretical reason to assume that  $t_1$  and  $t_2$  could not be separated by periods as long as years, this notion of retroactive autonomy conflicts powerfully with our normal intuitions. Frankfurt could attempt to respond to this objection by stating that since satisfaction is not an intentional action, it can appear simultaneously with the second-order volition. Unfortunately, in order to answer the temporal worry, Frankfurt needs to show that satisfaction *always* occurs simultaneously with a person's volition (and this claim seems almost certainly false). Also, if satisfaction does occur simultaneously with all second-order volitions, then apathy would insulate all of an agent's volitions from scrutiny. Of course, if a person cannot examine her volitions, then she cannot distinguish the licit from the illicit ones, even if she does have criteria for making such a classification.

At this point, we can consider Frankfurt's most recent work to see if it sheds light on a solution to our problems. We started by considering his theory in terms of volitionally endorsing desires through rational reflection, and Frankfurt addresses this directly.

What is essential to persons is not, in my view, a capacity to measure the value of their desires or to assess the desirability of their impulses. Rather, it is a capacity to identify themselves with (or refrain from identifying themselves with) their tendencies to be moved in one way or another. These reflective attitudes of identification or of withholding are often based on or grounded in evaluations of desirability. However, they need not be. A person may identify himself with (or withhold himself from) a certain

desire or motivation for reasons that are unrelated to any such assessment, or for no reason at all.<sup>19</sup>

This explanation by Frankfurt illustrates the change in his notion of identification, but now we must ask exactly what he means by identification, especially if someone may identify with a desire without reasons.

It is the notion of identification as acceptance, which I adumbrated first in “The Faintest Passion.” Since I may identify with desires of which I do not approve, identification does not entail endorsement. Since I may identify with desires that I consider to be quite trivial, such as a desire to have some ice cream, identifying does not entail caring. Perhaps it is more or less true, as Watson observes, that ‘what Frankfurt has in mind in cases of volitional struggle is the endeavor to go against what one cares about.’<sup>20</sup> To the extent that this is true, however, it is not because there is some essential linkage between identification and caring. It is just because a person ordinarily has no reason or motive to struggle except in cases where there is something at stake that he cares about.<sup>21</sup>

Since Frankfurt has greatly weakened the connection of identification with reflective endorsement and caring, our only recourse is to focus more closely on autonomy as it is associated with identification as acceptance. Along this line of thought, Frankfurt eventually states that identification is more concerned with confidence than truth. Specifically, having confidence entails that a person believes – although such a belief need not be explicit or even noticeable to the agent – that the things she values are in fact valuable, even if only to herself.<sup>22</sup> For example, Jim has a first-order desire to protect his

---

<sup>19</sup> Harry Frankfurt, “Reply to Gary Watson,” *Contours of Agency*, ed. Sarah Buss and Lee Overton, (Cambridge, Massachusetts: The MIT Press, 2002), 160.

<sup>20</sup> Gary Watson, “Volitional Necessities,” *Contours of Agency*, ed. Sarah Buss and Lee Overton, (Cambridge, Massachusetts: The MIT Press, 2002), 148, quoted in Harry Frankfurt, “Reply to Gary Watson,” *Contours of Agency*.

<sup>21</sup> Frankfurt, “Reply to Gary Watson,” 161.

<sup>22</sup> Harry Frankfurt, “Reply to Barbara Herman,” *Contours of Agency*, ed. Sarah Buss and Lee Overton, (Cambridge, Massachusetts: The MIT Press, 2002), 277.



children from harm along with an endorsing second-order volition. On the acceptance model of identification, Jim does not need to reflectively endorse his volition to protect his children – nor does he need to have done so previously – if he has held this volition unquestioningly since its inception. Even if no one else in the universe thinks that Jim’s children are valuable, Jim’s confidence that he values his children is sufficient for him to identify with his desire to protect them. Since he identifies properly with his desire to protect his children, Jim is autonomous on Frankfurt’s revised view. If this is the case, then a person is autonomous when she chooses her volitions based on what she holds to be valuable, regardless of the opinions of anyone else.<sup>23</sup>

Since we have now traced Frankfurt’s theory from its inception to its current incarnation, I shall briefly outline how the final draft of this model works. An individual starts out with a set of first-order desires, many of which are in conflict with one another. The person then forms second-order desires by dividing her desires into those that she endorses and those that she repudiates.<sup>24</sup> This is (usually) done by means of rationally considering the desirability of the various desires, based on what she considers valuable. Next, the agent will consider all of the first-order desires that are endorsed by second-order desires, and she will decide which of these desires that she wants to move her to action. She makes this choice by wholeheartedly identifying with one of her desires, and

---

<sup>23</sup> To be candid, I am not exactly sure what Frankfurt is trying to say in this last set of essays. By email message, he explained that “in speaking of ‘endorsement’ I didn’t intend to speak of positive evaluation or approval, but of something closer to mere acceptance. Being satisfied with it just means that the person has no inclination to resist or to alter the desire for it. The satisfaction need not be based upon a judgment that the desire (or its object) is valuable. It is more important that the person be wholehearted, which is pretty much the same as being satisfied.” While I am still unclear as to his precise meaning, I am encouraged by the fact that no one else really seems to understand the recent changes to his theory, either.

<sup>24</sup> There are two minor details which might be addressed here. First, a person can change her mind about whether to endorse or repudiate a given desire based on differing circumstances or information. Second, one might ask if an individual might have a desire that she neither endorses nor repudiates. The answer is that this may be possible if the desire is a trivial one, but the issue is unimportant. Even if a person is completely indifferent to a desire, the person will still undoubtedly prefer to pursue a desire that she specifically endorses over one that she merely does not reject.

in doing so, she forms a second-order volition. Finally, if the agent is satisfied with her second-order volition, then she is autonomous with respect to it.

Now, we can examine the most recent version of Frankfurt's theory to see what difficulties remain. The most obvious difficulty is the lack of clarity concerning identification, which is the crux of this theory. As it is unclear how people come to identify with their desires, the problem of manipulation must still be considered a possible threat to this model. Frankfurt is also still faced with the *ab initio* problem, since he has yet to explain how a person initially becomes autonomous with respect to a desire. I propose that we need to answer the latter problem first, as manipulation is irrelevant unless it is clear that someone can initially become autonomous.

## 2.2 Variation on a Paradigm

Not long after Frankfurt first proposed his hierarchical theory of autonomy, Dworkin proposed his own theory. Though much of Dworkin's theory is similar to Frankfurt's, there is one difference of special importance that I would like to examine. For this reason, I shall offer a very brief discussion of Dworkin's theory of autonomy.<sup>25</sup> Like Frankfurt, Dworkin has a hierarchical theory of autonomy, based on the notion of higher-order endorsement of lower-order desires. However, Dworkin adds a component to his theory that Frankfurt's lacks. Frankfurt's view of autonomy is 'ahistorical' in that his theory does not consider the formation of an agent's desires. Frankfurt is merely interested in whether the agent's desires fit into the proper structure. Dworkin, on the other hand, thinks that it is important to consider the method by which a person's desires

---

<sup>25</sup> It is important to note that I am only interested in Dworkin's early theory of autonomy. As time has progressed, he has changed his view significantly. However, his recent view seems to have become unnecessarily complicated and vague, so contemporary autonomy theorists tend to focus exclusively on his early work. I shall do the same.

were formed; hence, his theory is ‘historical.’ It is this difference that interests me, because my aim is to show that the addition of the historical component to a hierarchical theory of autonomy does not solve the *ab initio* problem.

With a general overview in mind, we are now able to consider the details of Dworkin’s notion of autonomy. Dworkin agrees with Frankfurt that an agent can be autonomous with respect to a lower-order desire only if she has an endorsing higher-order desire. However, Dworkin thinks that this is merely a necessary, rather than sufficient, condition for autonomy. He writes that his theory “may be characterized, in desperate brevity, by the formula autonomy = authenticity + independence.”<sup>26</sup> As Dworkin defines them, authenticity is the hierarchical aspect, and independence is the historical.

Authenticity, at least in this context, is similar to Frankfurt’s notion of identification. In order for a person to authentically hold a desire, she must approve and endorse her motives for the desire. Alternatively, someone can be autonomous with respect to a desire even when she is not aware of her motives as long as she would have approved of those motives under conditions of full information. To illuminate this notion of authenticity, we can return to the previous example of pacifism and analyze it in Dworkin’s terms. In this situation, you are a pacifist under attack by an assailant. You have a desire to fight the attacker and another desire to flee. You then evaluate your desires based on the motives that you endorse, and you form higher-order desires. Your authentic higher-order desire endorses the lower-order desire to escape, because you identify with the desire to avoid violence. Authenticity need not be absolute, though; this

---

<sup>26</sup> Gerald Dworkin, “Autonomy and Behavior Control,” *Biomedical Ethics: 3<sup>rd</sup> Edition*, ed. Thomas A. Mappes and Jane S. Zembaty, (New York, New York: McGraw-Hill Book Company, 1991) 275.

is evident if we change the scenario just as we did when considering Frankfurt's view. Once again, the assailant is attacking your child instead of you, and the only way to prevent the attack is to assault the aggressor. In this case, you may approve of your desire to protect your child more fully than your desire to shun fighting, so your higher-order parent-desires will be more authentic than your higher-order pacifist-desires. Therefore, authenticity is dependent on the degree to which someone approves of her desire.

Even so, Dworkin does not believe that authenticity alone is sufficient for autonomy; independence is also necessary. He claims that there are two types of independence: procedural and substantive. Procedural independence generally refers to the requirement that an agent's desires be formed through an appropriate process; whereas, substantive independence consists in being free from commitments.

First, let us look at procedural independence. Dworkin explains what this should mean in *The Theory and Practice of Autonomy*.

Spelling out the conditions of procedural independence involves distinguishing those ways of influencing people's reflective and critical faculties which subvert them from those which promote and improve them. It involves distinguishing those influences such as hypnotic suggestion, manipulation, coercive persuasion, subliminal influence, and so forth, and doing so in a non ad hoc fashion.<sup>27</sup>

Although Dworkin does not explicitly state how procedural independence can be judged, it is obvious that his goal is to address the problem of agential manipulation, or intentional control by an outside agent. We have strong intuitions that a person cannot be autonomous with respect to desires that were instilled or cultivated in the person by

---

<sup>27</sup> Dworkin, *Theory and Practice*, 18.

outside forces, and Dworkin proposes procedural independence in order to explain our intuitions. We can see how this works with a simple example. Imagine that Quentin is invited by his friends to go camping in the forest. Quentin's mother does not want him to go, so she tells him that there are evil ghosts in the forest who will attack him. Owing to his mother's manipulation, Quentin forms a higher-order desire that repudiates his desire to join his friends. Quentin lacks procedural independence because his mother has hindered his ability to rationally evaluate his desires. If Quentin had formed his higher-order desire based on some reasonable assessment of the dangers of camping – for instance, if he were allergic to tent canvas – then his desire would be procedurally independent.

Substantive independence, on the other hand, consists in being free from all external influences or commitments. It appears that Dworkin's original intention for proposing this condition is to eliminate the influences of harmful nonagential manipulation. Marilyn Friedman claims that "clearly our highest level principles are as much the products of socialization as any of our lowest level motivations."<sup>28</sup> Undoubtedly, Dworkin is attempting to address such charges that hierarchical theories of autonomy cannot account for the influence of society. To see how substantive independence is supposed to work, we can return to the example of Quentin. This time, his mother does not tell him that there are ghosts in the forest; however, Quentin has been raised in a society that views camping as extremely unmanly. Quentin still forms the higher-order desire to avoid camping, but now his reason is that he is afraid of humiliation if he ignores his prescribed gender role. Although no one has directly

---

<sup>28</sup> Marilyn Friedman, "Autonomy and the Split-Level Self," *The Southern Journal of Philosophy* 24, (1986): 25.

manipulated him, Quentin lacks autonomy because he has a commitment to social norms that interferes with his rational evaluation.

Although substantive independence does succeed in preventing nonagential manipulation, Dworkin recognizes a serious drawback. “The compassionate or loyal or moral man is one whose actions are to some extent determined by the needs and predicaments of others. He is not independent or self-determining.”<sup>29</sup> Dworkin’s concern is that substantive independence may be impossibly demanding, as well as counterintuitive. Of course, without substantive independence he cannot address the problem posed by Friedman. Essentially, the issue is how to distinguish between licit and illicit higher-order desires, so Dworkin must confront the *ab initio* problem just like Frankfurt.

---

<sup>29</sup> Dworkin, “Autonomy and Behavior Control,” 277.

## CHAPTER 3. WHAT IS NEEDED

### 3.1 *Ab Initio* in Depth

I have now explained both of the foundational hierarchical theories of autonomy. By examining these theories, it can be seen that any hierarchical account of autonomy will have to deal with three interrelated problems. To restate briefly, these are: the regress problem, the *ab initio* problem, and the problem of manipulation. I will quickly review how these problems fit together so that I can explain my own project.

To begin, recall that we are dealing with hierarchical theories. Any theory which explains autonomy in terms of higher-order desires endorsing lower-order desires must explain how to avoid an infinite regress of higher-order endorsement. In order to avoid the regress, a theorist needs to specify the conditions under which an agent can develop autonomy from a set of desires that the agent is not autonomous with respect to. The question of how an individual initially becomes autonomous in relation to a desire is the *ab initio* problem. Finally, a theory of autonomy must account for manipulation; that is, it should be able to make a clear distinction between licit and illicit (or licitly- and illicitly-formed) desires. It is important to see how these three issues fit together, because now my own endeavor should make sense. I am attempting to address the *ab initio* problem because it *needs* to be addressed first. If this dilemma can be resolved, then the regress problem will be immaterial. Essentially, the regress will end when an agent fulfills the initial requirements for acquiring autonomy in relation to a given desire. Moreover, since the problem of manipulation is about whether or not autonomy can be counterfeited in some way, manipulation should not be a primary concern until after a

theorist has attempted to provide a solution to the *ab initio* problem. This is because one cannot create a facsimile of autonomy if there is no original explanation.

Now that I have explained why the *ab initio* problem needs to be addressed, let us examine exactly how it affects Frankfurt's and Dworkin's theories. The explanation of the problem can most easily be seen in the example of a 'willing slave.' In this example, imagine that a person has been raised in slavery throughout her life. Eventually, she is freed from her slavery. Nevertheless, she believes that slavery is her only and, more importantly, her proper destiny. Although this notion of a human being may seem impossible to most of the people in modern first-world countries, Sigurdur Kristinsson points to a relevant article from the New York Times about a young woman named Fatma Mint Mamadou, who is a dishearteningly accurate illustration of a willing slave.

"God created me to be a slave," she says in her quiet voice, "just as he created a camel to be a camel." ...Raised to obey – to haul water, herd animals, cook, sweep and reproduce – Fatma shows no will to resist either history or that stark landscape [of Mauritania, North Africa]. She will be a slave forever, in her own mind. In fact, that is the only place she remains a slave, for she is no longer technically indentured to anyone.<sup>30</sup>

While this might give the impression of merely a dispiriting digression from the topic, the fact is that the very existence of someone like Fatma is a threat to content-neutral autonomy. Kristinsson neatly explains how this is the case.

Neutral accounts suggest, roughly, that if someone consistently endorses (or could endorse) a desire to follow orders without question or condition, then he may be autonomous in acting on that desire. Intuitively, however, an autonomous action must involve the agent's own choice of ends at a more concrete level than that of pursuing

---

<sup>30</sup> Elinor Burkett, "God Created Me To Be a Slave," *The New York Times*, 12 October 1997, [On-line]: <http://query.nytimes.com/gst/abstract.html?res=F10E17FD34590C718DDDA90994DF494D81>. Note: The URL listed here will only allow access to the article abstract. The full article can be accessed for a fee.



whatever ends some authority commands. At the very least, his commitment to particular ends must be sustained by his reason-states in general, i.e. his relevant beliefs and desires.<sup>31</sup> This relationship between the agent's ends and his relevant beliefs and desires seems broken if he is unconditionally committed to pursue whatever ends the authority commands regardless of how he might feel about them.<sup>32</sup>

Kristinsson's point is that Fatma does not seem to be in any way autonomous with respect to her desire to remain in servitude. However, she also appears to satisfy all of the conditions for autonomy on Frankfurt's view.

On Frankfurt's account, an autonomous agent needs to be satisfied with her second-order volition, which endorses a first-order desire. Fatma does have a first-order desire to be a slave, and she has a second-order volition endorsing that desire. Finally, her assertion that God made her a slave seems indicative of satisfaction with her volition. It looks as if Fatma does identify with her slave-desires on Frankfurt's view. Additionally, Frankfurt has already stated that social pressures do not inhibit autonomy – at least, not in any vital way – so there can be no appeal to social manipulation as an explanation for why Fatma might actually lack autonomy.<sup>33</sup>

We can now consider Fatma's case in light of Dworkin's theory. Since Fatma is (or could be) both aware and approving of her second-order desires endorsing her first-order slave-desires, she meets the qualifications of authenticity. The real question is

---

<sup>31</sup> In a recent manuscript, Bernard Berofski describes 'contingencies,' or desires that simply exist without explanation. "Often, there is no (completely) objective justification for those commitments that express the deepest values of an agent – religion, profession, personal ideals, lifestyle, personal relationships. We have preferences and are prepared to grant that the distinct preferences of others are no less securely grounded." Bernard Berofski, "Autonomy Without Free Will," (unpublished manuscript). These contingencies would not be responsive to one's reason-states, but they do not seem to be autonomy-undermining. In order to explain this, I would have to define the difference between licit and illicit manipulation, and that project would take too much time to be pursued here.

<sup>32</sup> Sigurdur Kristinsson, "The Limits of Neutrality: Toward a Weakly Substantive Account of Autonomy," *Canadian Journal of Philosophy* 30 (2000): 257-286.

<sup>33</sup> Frankfurt, "Freedom of the Will," 23-25.

whether she can meet the requirements of procedural and substantive independence. In terms of procedural independence, we can ask if Fatma has acquired her slave-desires through an appropriate process. Recall that procedural independence is determined by “the distinction... between those modes of evaluation that interfere with the rationality of higher-order reflection and those that do not.”<sup>34</sup> It is obvious that Fatma’s ability to reason is not inhibited by her desires, as she seems perfectly capable of making rational assessments. She simply does not come to the same conclusions about the nature of her existence that most of the inhabitants of first-world countries would. Fatma does seem to have procedural independence; however, she may not have substantive independence (a lack of external commitments). On substantive independence, one might say that Fatma is simply committed to her master, and therefore she fails to fulfill this condition. However, a closer look at Fatma reveals that she is not particularly committed to any one person; she is merely committed to the life of slavery, regardless of the owner. “She knows that she was *abd*, a slave, just as she knows that her mother was inherited by Sidi M’Hamed Ould Hamadi when the old master died. But that state is no more remarkable to her than the rich brown of her skin.”<sup>35</sup> One might respond that Fatma still fails to achieve substantive independence because her desires are still being formed on the basis of external commitments to the lifestyle of slavery. Apparently, Fatma does not possess substantive independence, so Dworkin can account for the intuition that she is not autonomous with respect to her slave-desires. Surprisingly, accounting for the lack of autonomy in this situation still leaves a serious difficulty. We have seen that Fatma’s substantive commitment is to the lifestyle that is her cultural heritage – horrible as that

---

<sup>34</sup> Dworkin, *Theory and Practice*, 161.

<sup>35</sup> Burkett, paragraph 5.

legacy may be. If Dworkin denies that an autonomous individual can have this sort of commitment, then he is forced to claim that people whose desires are influenced by a commitment to their social background cannot be (fully) autonomous. This understanding of substantive independence is unrealistically demanding, as no one can fully (or perhaps even substantially) escape the influence of society. It is for this reason that Dworkin eventually abandons the notion of substantive independence.

The conception of autonomy that insists upon substantive independence is not one that has a claim to our respect as an ideal. It is a conception which violates one of the constraints I imposed [previously] – that of being consistent with other important values we hold. In particular it makes autonomy inconsistent with loyalty, objectivity, commitment, benevolence, and love.<sup>36</sup>

Even if Dworkin does give up substantive independence, the case of Fatma is problematic for him. Specifically, if Dworkin allows that commitments to one's culture do not undermine autonomy, then Fatma has fulfilled all of the requirements for autonomy even though it is quite apparent that she is not autonomous in relation to her slave-desires.

Both Frankfurt and Dworkin are in a quandary concerning the willing slave problem. If they claim that autonomy exists solely in virtue of an agent's subjective mental states, then it seems as though one can always be manipulated into meeting all of the necessary mental state requirements, but such manipulation intuitively undermines autonomy. In contrast, if autonomy is made dependent upon some objective fact about a person, then it seems that her freedom is dependent on something entirely outside of her control. What is needed, then, is a test for identification with a desire that neither appeals to outside endorsement, nor allows for the possibility of manipulation. In other words, we need a solution to the *ab initio* problem.

---

<sup>36</sup> Dworkin, *Theory and Practice*, 21.

### 3.2 The Need for a Need

In the past few years, several philosophers have suggested that the *ab initio* problem can be solved by positing a substantive theory of autonomy. At this point in the discussion, I need to distinguish between ‘content-neutral’ and ‘substantive’ autonomy theories. Both Frankfurt’s and Dworkin’s views are entirely content-neutral because, on either of these views, an agent’s desires, beliefs, or other mental states are not required to have any specific property in order for the agent to be autonomous with respect to a given desire. On Frankfurt’s view, a person’s desires need only fit into the proper hierarchical structure, and Dworkin only requires that a person form her higher-order desires according to a proper procedure. This is in direct opposition to substantive theories, which claim that an agent does need to possess a substantive mental state. For the purposes of this discussion, a substantive mental state is a mental state that has some property which is fixed by the theory of autonomy. In other words, a proponent of substantive autonomy believes that, in addition to other criteria, an agent must have a mental state that has some constant property in order for the agent to be autonomous with respect to a desire. We can see the difference between content-neutral and substantive theories in an over-simplified example. (For the purpose of this example, I will concentrate solely on the relations between beliefs and desires. I am ignoring any concerns about satisfaction or independence for the sake of focusing on the basic difference between neutral and substantive theories.)

Hierarchical Content-neutral Autonomy: Person *P* is autonomous with respect to desire *D* at time *T* only if *P* has a higher-order desire *H* that endorses *D* at *T*.

Hierarchical Substantive Autonomy: Person *P* is autonomous with respect to desire *D* at time *T* only if *P* has a higher-order

desire  $H$  that endorses  $D$  at  $T$ , **and**  $P$  has substantive mental state  $M$ , which does not repudiate  $H$  at  $T$ .

The essential difference is that a substantive theory of autonomy demands that an agent must have a substantive mental state which acts as a gatekeeper for higher-order desires. This does not mean that an agent cannot have a desire that is repudiated by the substantive mental state, but it does mean that the agent cannot be autonomous with respect to such a desire.

If a theorist intends to offer a substantive account of autonomy, then she must specify exactly which mental states will be required for judging higher-order desires. Some theorists that believe autonomy requires a substantive basis have focused on the notion of self-respect.<sup>37</sup> The main idea is that a person can only be autonomous in relation to desires that are oriented toward, or at least not repudiated by, self-respect. It behooves me to explain exactly what is meant by 'self-respect' in this context.

Those who have self-respect are often said to have a "sense of their own worth." While often moving and doubtless important, this idea is also puzzling. ...Though the topic remains a difficult one, one conjecture emerges from our examples: at least part of a sense of one's own worth is having, and living by, personal standards or ideals that one sees, whether objective or not, as an important part of oneself.<sup>38</sup>

While Thomas Hill, Jr. seems wary of committing himself fully to a firm definition of self-respect, it seems reasonable to accept this as a working definition, since it does appear to capture our normal intuitions about self-respect. Unfortunately, Hill's definition illustrates exactly why self-respect cannot be the basis of autonomy. The basic

---

<sup>37</sup> One good example is Thomas Hill, Jr., who takes a Kantian approach that focuses on respect for oneself and others. For his full discussion, read his book *Autonomy and Self-Respect*.

<sup>38</sup> Thomas Hill, Jr., *Autonomy and Self-Respect* (New York, New York: Cambridge University Press, 1991), 23.

problem is that there are two ways in which one could develop a sense of his own value in the way that Hill describes, and both of these methods undermine autonomy.

One manner in which a person may derive her self-respect is from the approbation of others. At first glance, this method of acquiring self-respect might seem plausible because the disapprobation of others almost always impacts a person's self-respect negatively. Conversely, we might think that the approval of other people would increase our self-respect. However, this leads to the original problem that Aristotle mentioned concerning honor. "Men of culture and *action* seek a life of honor; for the end of political life is almost this. But this good... is thought to depend on those who bestow rather than on those who receive honor."<sup>39</sup> If self-respect is developed merely through the support of other people, then it cannot be of any use in determining whether an individual agent identifies with a given higher-order desire.

Obviously, then, self-respect must be grounded in something other than the approval of outsiders. In order to see how this can be the case, let us reexamine our definition of self-respect, outlined above. At the very least, self-respect must contain the recognition that one possesses certain principles or values and lives by them.<sup>40</sup> However, we are now faced with the task of explaining exactly what the principles and values are that a self-respecting person lives by. Since the person must live by the values in question, those values have to be both motivating and action-guiding. In the context of this discussion, we can understand 'values' to be desires that a person identifies with; moreover, the person can acquire self-respect by acting in accordance with her most

---

<sup>39</sup> Aristotle, *Nicomachean Ethics*, trans. Hippocrates G. Apostle, (Grinnell, Iowa: The Peripatetic Press, 1984), 4.

<sup>40</sup> One might object that an agent could pride herself on avoiding the 'trap' of principles and values, but the agent only seems to be accepting deviance as a value instead of more ordinary ones.

closely-held values. Now we can see that an appeal to self-respect as a judge of higher-order desires appears to be circular. Self-respect cannot determine whether or not an agent is able to become autonomous with respect to a given desire because she can only gain self-respect by acting in accordance with desires to which she *already* bears an autonomous relation. It looks as though relying on self-respect in this way actually prevents anyone from attaining autonomy at the outset.

One might respond to this argument by claiming that a person can gain self-respect by satisfying desires that she holds non-autonomously. For example, a child might develop self-respect by living in accordance with values that she has internalized through continuous interaction with her family. Later, when the child develops sufficient self-respect, she will be equipped to evaluate her desires and identify with them. However, troubles arise when we try to apply this understanding of self-respect to our willing slave case. Fatma originally acquired her values, such as the desire to be a good slave, from her family and friends. It is reasonable to assume that she gained self-respect when she acted properly in accord with her values. We can also assume that she eventually developed – or could have developed – the ability to determine whether her higher-order desires conflicted with her desire to gain self-respect by being a good slave. In this case, we can see that Fatma’s self-respect not only fails to weed out illicit desires, it probably eliminates any vestiges of (intuitively) legitimate desires.

It is important to note that the underlying problem with self-respect is not the fact that values are genetically or environmentally conditioned. Rather, the trouble is that self-respect is dependent upon the very desires that it is supposed to adjudicate. Fortunately, recognition of this difficulty with self-respect does give a hint as to what

specific kind of mental state is needed in a sufficient substantive account of autonomy. To elaborate, the mental state needed for a substantive account of autonomy will have to conform to several conditions. First, the substantive mental state must be able to motivate a person to alter her desires. Second, the mental state must also be objective in the sense that it must be immune to the influences of socialization and manipulation. Third, all humans must necessarily possess the substantive mental state. Finally, a satisfactory mental state must be capable of distinguishing between licit and illicit higher-order desires.

At first glance, one might think the combined conditions would lead to a *reductio* on the possibility of a substantive notion of autonomy. In spite of this, I shall attempt to use the aforementioned conditions to find an appropriate mental state. Since the mental state in question must be effective in motivating an agent to modify her desires, the mental state must itself be a desire of some sort.<sup>41</sup> To meet the second qualification of immunity from external influences, the search for a serviceable mental state must be confined to a subset of desires known as ‘needs.’ For this discussion, I shall stipulate that a need is a certain type of desire wherein the frustration of the need will have a serious adverse affect on the person who possesses it. It should be immediately apparent that needs are not completely immune to external influences, but they are more resistant than most of the subsets of desires, such as whims or wants. I admit that I have not yet fulfilled the requirement for imperviousness to outside influence; however, I will return to this condition momentarily. Now that the search is restricted to needs, I must find a type of need that is possessed by every person in existence. To fill this requirement, I

---

<sup>41</sup> Some philosophers will undoubtedly take issue with the claim that desires, not reasons, motivate people to action. However as I stated earlier, I cannot fully argue for the Humean metaphysical view in the space and time available to me.



shall appeal to Abraham Maslow, a psychologist who attempted to discover the ‘basic needs’ of human beings. “Some values are common to all (healthy) mankind.... What I have called basic needs are probably common to all mankind and are therefore shared values.”<sup>42</sup> In strictly biological terms, we can conceive of basic needs such as food, shelter, and sex. (I should briefly remark that, given my definition of ‘need’ and ‘basic need,’ any proposition that is true for all desires is true of all needs, and any proposition that is true for all needs is true for all basic needs.) We can find these needs in all healthy human beings and, more importantly, it is obvious that people have these needs in order to insure their survival. (Admittedly, the need for sex does not directly affect one’s own survival, but it remains a basic need because it exists in all people for the survival of the species.) If we turn our attention briefly back to the second criterion, we can see that we now have a set of mental states that cannot be eradicated or significantly altered by environmental influences. If an agent’s basic needs were somehow destroyed by her environment, then the quality of her life would quickly diminish, and she would most likely die. If she lost her need for food or shelter, then she would be unmotivated to provide the most basic sustenance for herself. Moreover, while one might argue that the loss of a need for sex might not necessarily affect any individual life adversely, the widespread loss of that need would be threatening to the survival of humankind on the whole. From an evolutionary standpoint, it makes perfect sense to claim that all healthy humans will possess the basic needs that I have listed, including the need for sex. The last condition – which is the tricky one – for a substantive mental state is that it must be able to distinguish between legitimate and illegitimate higher-order desires, but needs for

---

<sup>42</sup> Abraham Maslow, “Psychological Data and Value Theory,” *New Knowledge in Human Values*, ed. Abraham Maslow, (New York, New York: Harper & Brothers, 1959), 122. For Maslow’s original discussion of psychological basic needs, read his book *Motivation and Personality*.

food, shelter, or sex certainly do not have even a remote possibility of performing such a function. Since this is the case, I shall try to posit the existence of an additional basic need: a need for self-worth.

## CHAPTER 4. THE SOLUTION

### 4.1 The Need that Satisfies a Need

By self-worth, I am referring to an unembellished combination of the words ‘self’ and ‘worth.’ Essentially, self-worth is a person’s belief that her life has some merit or usefulness. This is different from ‘self-respect,’ which has connotations of pride attached to the meaning. A person can have self-worth even if she does not take pride in her life. For example, consider a man who is ashamed of his lifestyle and his job. However, he has a young daughter that he cares for very much. The man may believe that his life has merit because he is caring for his daughter, but he can still be ashamed of his life.

Now that the definition of self-worth is clear, I can evaluate my hypothetical need by using the convenient set of criteria outlined above. In order to satisfy the first three conditions, I need to show that the human longing for self-worth is due to a basic need. To make the case, I shall assume that a basic need for self-worth does not exist, and I will show that the lack of such a need would endanger humanity on an evolutionary scale. What exactly does it mean to say that self-worth is not a basic need? Essentially, this implies that humans, as a species, would be able to survive the process of natural selection without every one of them – or almost every one – satisfying the need. It is important to remember that this discussion focuses on humans *as they actually exist in the real world*. I am not attempting to explain the needs of some ideal creature similar to humans (such as a being that possesses perfect rationality or lacks desires), or humans in unrealistically hypothetical circumstances (such as a single person who is the only creature in the world).

My first task is to establish how a basic need for self-respect might protect survival, and then I can consider whether humanity as we know it could have endured without it. Typically, we do not think that animals seek self-worth.<sup>43</sup> Likewise, people generally do not think that animals are not capable of critical self-reflection. In other words, an animal is not able to evaluate its own actions, character, and life. These general ideas indicate that self-worth might only be necessary for the survival of self-reflective creatures. Let us assume that during evolution some primitive humans become capable of critically reflecting on their own existence, but none of them has any desire for self-worth. Without a desire to improve their own worth, the primitives have no reason to reflect on their overall characteristics or lives, and would have no motivation to change if they did. Even if they do use their abilities to self-reflect, the information they gain will be of interest to them only insofar as it helps them satisfy their physical desires. Since the people in this example have no wish to improve themselves, they would have no motivation to put forth much effort toward any enterprise that does not have an immediately foreseeable reward. Academic and artistic enterprises require a great deal of effort for very little tangible or immediate gain, so such our primitives would not pursue them. For someone who is not motivated by a desire for self-worth, there is no reason to attempt any action that does not appear to directly lead to the satisfaction of some physical desire. For our hypothetical primitives, it seems more likely that the ability to self-reflect would atrophy and vanish without the inherent desire to use the ability. Thus we can see that self-reflective creatures without any desire for self-worth might survive the process of evolution, but they would not become *Homo sapiens*.

---

<sup>43</sup> This does not mean that animals do not act in ways that would add to a human's self-worth. However, these actions can be explained by instinct or training.

So far, I have only shown that a small number of people must want to improve themselves; it is still possible that only a scarce few are motivated by self-worth. By no means does this imply that self-worth is a basic need. To achieve the results I require, I shall revisit the previous example. Imagine that a group of primitive humans become capable of critical self-reflection, and one of them possesses a desire for self-worth. If she attempts to fulfill this desire to gain merit, she has a better-than-normal chance of improving her own and her tribe's chances of survival. This could happen in any number of ways, such as building shelters, making weapons, or making fire. Someone who improves the tribe's life would be highly valued and others would undoubtedly emulate this individual. If just a few people in one tribe acquired this desire, then that tribe would have a great adaptive advantage over other tribes, and the 'self-worth tribe' would flourish. For this reason, we can conclude that, in self-reflective creatures, a desire for self-worth is a great evolutionary advantage on both an individual and social scale. This advantage makes it likely that the desire would become widespread, and that it would quickly become a basic need in order to survive.

While I have shown that there appears to be an evolutionary explanation for the existence of a desire for self-worth, I have not shown that there any reason for modern people to have a need for self-worth because society has, for the most part, eliminated the threat of death by natural selection. In addition, some people may claim that the survival of the human race is not a sufficient justification for compliance with a desire. (This is why it is so important that sex is extremely pleasurable.) Why, then, can we not discard the basic need for self-worth, or at least relegate it back to the status of a mere optional desire?

The first answer to this question is that a basic need for self-worth continues to be an adaptive advantage for each *individual* human being. Obviously, a person who helps other people will be valued by those other people, and since most people conceive of their self-worth as inclusive of their usefulness to others, individuals who have self-worth are considered more valuable to other people. It stands to reason that someone who has a persistent need for self-worth is more likely to work diligently for it than someone who lacks this need. This is why a basic need for self-worth remains practically useful and adaptive. As this argument is strictly prudential, many philosophers might find it unsatisfying; however, there is a stronger reason to accept the need for self-worth as basic.

I have already argued that a desire for self-worth is necessary for a person to exercise her ability to critically self-reflect. As critically self-reflective creatures, it is in our nature to question the purpose for our existence in the universe. More to the point, each person eventually considers her own individual reason for existing. Owing to the desire for self-worth, no one is likely to be satisfied with the conclusion that the entire purpose of her existence is to assist in the futile task of attempting to perpetuate the species until entropy eradicates all life in the universe. The more general point that I am indicating here is that if the worth of someone's existence has nothing to do with that person's strongly-held values, then she cannot have any control over the worth of her life. This idea is intuitively unsatisfying, and it conflicts with our common notions about what it means for someone to have worth. However, there are two other answers that might satisfy the desire. One of these possible answers is that there is no preordained, teleological purpose to human life, but each person is capable of giving her life worth by

acting in accordance with her values. The other possibility is that a deity imbues each human life with a purpose that reflects the deity's values, and each person can give her life worth by living in accord with the values imparted by the deity. Taking into account the fact that people who accept the second possibility will take the deity's values to be their own, it seems that the same answer actually applies in both cases. A critically self-reflective human being can give her life worth by living in accordance with her values.

Notice that giving one's life worth is an ongoing process, and a person must periodically reflect on her actions and values in order to be certain that she is (a) identifying with the values that she believes are most important, and (b) correctly living in accordance with her values. Now it is clear that a person continually needs to need self-worth in order to have a fulfilling life. One might object that it looks as though I am once again grounding the basic need for self-worth in survival, but this would be a misunderstanding of my claim. I am arguing that people, owing to their human nature, cannot survive without having a greater purpose than mere survival. This necessity for a greater purpose is the need for self-worth, and it is basic because people require a greater purpose than survival throughout the entirety of their lifetimes.

#### 4.2 Getting the Job Done

I have established that people have a basic need for self-worth, but to prove that it can serve a role in a substantive theory of autonomy, I must still demonstrate that it can adjudicate between legitimate and illegitimate higher-order desires. To test whether my theory works, I shall attempt to address three cases: the willing slave, the dutiful soldier, and the humble cleric. My aim is to show that the need for self-worth will repudiate the

higher-order desires of people who are intuitively non-autonomous, but it will allow the higher-order desires of individuals that we take to be autonomous.

We can begin with the willing slave, Fatma. When we last evaluated Fatma's case on Frankfurt's theory, she qualified as autonomous with respect to her slave-desires, even though she intuitively seems to lack autonomy. For Dworkin, Fatma provided an illustration of the problems with his notion of substantive independence. However, in light of a need for self-worth, her case makes much more sense. Fatma's insistence on subsuming her own desires in the will of another person seems to inherently undermine her ability to pursue self-worth. Therefore, her higher-order volitions that endorse the slave-desires appear to be incompatible with her need for self-worth, so we can now classify Fatma as non-autonomous.

But perhaps I am too quick to accept my conclusion. What if Fatma truly believes that being a slave is the only way to enhance her self-worth? To clarify, let us consider what is necessary for any person to act on a desire. An agent cannot act on a given desire unless (a) the desire is present, and (b) the agent believes that the desire has at least a remote possibility of being satisfied. A man who is lost in the desert may desire water, but he will not search for it if he believes that there is absolutely no water to be had, regardless of whether accessible water actually exists. In Fatma's case, she may possess the need for self-worth, but she believes that her worth is limited exclusively to her usefulness to a master. However, I would like to scrutinize the belief that a person's worth may be entirely dependent on another person. We can recall that the need for self-worth is persistent, that is, it can never be permanently satisfied. Even if someone were to spend every moment of her life increasing her self-worth, she would still possess the



need to increase it further. A person could theoretically increase her self-worth for an infinite period of time because there is no conceptual limit to the amount of worth that an individual may accrue. However, at any specific time each person will have a certain finite amount of actual self-worth. If Fatma's beliefs indicate that the possibility of her self-worth is completely dependent on her usefulness to a master, then the conceptual limit of her self-worth is the degree to which he can actually exploit her. In addition, her slave-desires prevent her from reevaluating her beliefs and desires in order to expand the possibilities of finding self-worth. In short, she has reduced a limitless possibility to an extremely limited one. The desire/belief sets that constitute her slave-desires now interfere with her attainment of self-worth by denying her the opportunity to pursue it. As she cannot satisfy her need for self-worth while she retains her slave-desires, she cannot be autonomous with respect to those desires. Aside from demonstrating that Fatma does not have autonomy, I have also shown why no person can autonomously choose to be a slave to another person.

If my theory shows that a willing slave cannot possess autonomy, then how does it work in a similar case wherein the agent does intuitively seem to be autonomous? For example, imagine that a man – I shall call him John – is a 'dutiful soldier.' John's soldier-desires include the desire to follow the orders of his commanding officer(s), even if those orders mean John's death. Nevertheless, our intuitions indicate that John is autonomous in relation to his soldier-desires. The key is that John's self-worth is not solely dependent on his commander(s) or, by extension, his country's actual worth. The desire to follow orders only exists because John has a general desire to acquire self-worth by defending his country. John's overall soldier-desires do not exclude him from

deciding at some later point to pursue self-worth by another method. John is autonomous with respect to his soldier-desires because these desires are not inconsistent with satisfying his need for self-worth, and he can adjust or discard his soldier-desires later if necessary.<sup>44</sup> The distinction between Fatma the willing slave and John the dutiful soldier is that the willing slave holds desires that cannot be changed if (and when) they endanger her ability to acquire self-worth, while the dutiful soldier is able to change his desires if they interfere with his pursuit of self-worth.

The case of the dutiful soldier might make us think that autonomy is dependent on the ability to evaluate and change one's own desires, but another example will illustrate that this is not the case. We can think of Kendal, the humble cleric, who joins a monastery in order to minimize his personal worth so that he can devote his life to his deity. Kendal repudiates all of his desires, except for the desire to do God's will. Although Kendal feels the same way about God as Fatma feels about her master, our general intuition is that Kendal is autonomous with respect to his cleric-desires. (In fact, we might even admire his dedication to his religion.) For my theory to succeed, I must show that there is some essential difference between Fatma and Kendal, and that this difference is discernible in view of a need for self-worth. Fatma's lack of autonomy stemmed from the fact that her commitment to her master severely limits her ability to acquire self-worth, so it seems reasonable to see if a commitment to God limits Kendal's ability to acquire worth. Fatma has problems because her self-worth is limited to her finite usefulness to her master, and nothing else. Therefore, she cannot satisfy her need for self-worth at any time that her master does not need her for something. Like Fatma,

---

<sup>44</sup> There might be practical difficulties if John changes his desires under certain circumstances. For example, there might be steep penalties if John deserts the army during a war. However, practical concerns will only affect whether or not John can satisfy his desires, not whether or not he can change them.

Kendal's self-worth is dependent on his usefulness to God. But unlike Fatma's master, God (at least according to the western tradition) is capable of using Kendal's talents in an infinite number of ways, so the commitment to God does not put any conceptual limits on Kendal's ability to pursue self-worth.

The fact that there are no limits on Kendal's possible self-worth does not address the more basic concern with this case. Even if it is possible for Kendal to pursue all the self-worth he can hold, his actual desire is to abase himself before God. The specific problem is Kendal's desire appears to be directed toward reducing his self-worth. Furthermore, because Kendal's cleric-desires, like Fatma's slave-desires, prevent him from later changing his own desires, Kendal cannot be autonomous for the same reason as John the dutiful soldier. The solution to this dilemma can be found by clarifying Kendal's exact beliefs and desires. Kendal believes that he has a body and a soul, and his soul is where his true self resides. It is true that Kendal repudiates his physical desires, but he does so because he wants to increase the worthiness of his soul. Now it is apparent that our humble cleric's desires do satisfy his need for self-respect, and his autonomy remains intact.

I have now shown that humans have a basic need for self-worth, and that this need can explain our intuitions about our three test cases: willing slave, dutiful soldier, and humble cleric. Given this, it seems reasonable to conclude that the need for self-worth does satisfy all of the criteria for a substantive mental state.

#### 4.3 Conclusions and Implications

Before I consider the implications of my theory, allow me to briefly summarize my argument. Current hierarchical theories of autonomy cannot explain willing slave

cases, but the addition of a substantive requirement can explain such cases. An appropriate substantive mental state must satisfy four criteria: (a) it must be able to motivate people to modify their desires, (b) it must be immune to socialization and manipulation, (c) it must be universal across humanity, and (d) it must be able to correctly categorize higher-order desires as legitimate or illegitimate. In order to satisfy the first three conditions, the mental state must be a basic need, or a desire that must be satisfied by all (or at least the vast majority of) people for the survival of humanity. Finally, I have argued that a need for self-worth fulfills the fourth criterion; hence, a basic need for self-worth is an appropriate mental state for a substantive hierarchical theory of autonomy. By supplying the necessary substantive mental state, I believe that I have solved the *ab initio* problem by explaining how an individual can initially become autonomous with respect to a given desire.

With the argument fresh in mind, I will conclude by mentioning some of the implications of my theory. The most obvious result is that the regress problem should dissolve, because the hierarchical theories of autonomy now have a foundation. Second, my conclusions indicate that human beings – and, in fact, any critically self-reflective creatures – may not be able to survive without autonomy. This undoubtedly has interesting implications for the philosophy of the mind, as well as moral psychology. Also, it seems to me that more research is needed to explore whether or not the basic need for self-worth can solve the problem of manipulation. If it can, then Frankfurt's theory of autonomy, with this substantive addition, could be an adequate explanation for all of autonomy. If self-worth cannot protect an agent from manipulation, then we need to explore other options in depth, such as Dworkin's procedural independence

requirement. The last, and perhaps most interesting implication of this argument, is that it may give us an indication of the difference between creatures that are fully sapient and creatures that are merely sentient, though they may be instrumentally reasons-responsive. If so, we might have reasons to think that a very few great apes are sapient, and more than a few humans are not.

## REFERENCES

- Aristotle. *Nichomachean Ethics*, translated by Hippocrates G. Apostle. Grinnell, Iowa: The Peripatetic Press, 1984.
- Burkett, Elinor. "God Created Me To Be a Slave." *The New York Times*, 12 October 1997, [On-line]:  
<http://query.nytimes.com/gst/abstract.html?res=F10E17FD34590C718DDDA90994DF494D81>.
- Christman, John. "Introduction." *The Inner Citadel*, edited by John Christman, 3-23. New York, New York: Oxford University Press, 1989.
- Dworkin, Gerald. "Autonomy and Behavior Control." *Biomedical Ethics: 3<sup>rd</sup> Edition*, edited by Thomas A. Mappes and Jane S. Zembaty. New York, New York: McGraw-Hill Book Company, 1991.
- Dworkin, Gerald. *The Theory and Practice of Autonomy*. New York, New York: Cambridge University Press, 1988.
- Frankfurt, Harry. "The Faintest Passion." *Necessity, Volition, and Love*. New York, New York: Cambridge University Press, 1999.
- Frankfurt, Harry. "Freedom of the Will and the Concept of a Person." *The Importance of What We Care About*. New York, New York: Cambridge University Press, 1988.
- Frankfurt, Harry. "Identification and Wholeheartedness." *The Importance of What We Care About*. New York, New York: Cambridge University Press, 1988.
- Frankfurt, Harry. "Reply to Barbara Herman." *Contours of Agency*, edited by Sarah Buss and Lee Overton. Cambridge, Massachusetts: The MIT Press, 2002.
- Frankfurt, Harry. "Reply to Gary Watson." *Contours of Agency*, edited by Sarah Buss and Lee Overton. Cambridge, Massachusetts: The MIT Press, 2002.
- Friedman, Marilyn. "Autonomy and the Split-Level Self." *The Southern Journal of Philosophy* 24, (1986): 19-35.
- Hill, Jr., Thomas. *Autonomy and Self-Respect*. New York, New York: Cambridge University Press, 1991.
- Hobbes, Thomas. *Leviathan*, edited by J. C. A. Gaskin. New York, New York: Oxford University Press, 1996.

Hume, David. *A Treatise of Human Nature*. New York, New York: Cambridge University Press, 2000.

Korsgaard, Christine. *The Sources of Normativity*. New York, New York: Cambridge University Press, 1996.

Kristinsson, Sigurdur. "The Limits of Neutrality: Toward a Weakly Substantive Account of Autonomy." *Canadian Journal of Philosophy* 30, (2000): 257-286.

Maslow, Abraham. "Psychological Data and Value Theory." *New Knowledge in Human Values*, edited by Abraham Maslow. New York, New York: Harper & Brothers, 1959.

Neely, Wright. "Freedom and Desire." *The Philosophical Review* 83, (1974): 32-54.

Spice Girls: Victoria Addams, Melanie Brown, Emma Bunton, Melanie Chisholm, & Geri Halliwell. "Wannabe." *Spice*. [Compact Disc]. Los Angeles, California: Virgin Records, 1996.

Taylor, James Stacey. "Autonomy, Duress, and Coercion." *Social Philosophy and Policy* (Forthcoming), (June 2003): 16-17.

Watson, Gary. "Volitional Necessities," *Contours of Agency*, ed. Sarah Buss and Lee Overton. Cambridge, Massachusetts: The MIT Press, 2002.

## VITA

Paul Jude Naquin received his baccalaureate degree in psychology from Louisiana Tech University in 2000. He went on to Louisiana State University, where he received a master of arts degree in philosophy in 2003. His primary interests are theoretical and applied ethics, especially in the area of bioethics. He is also interested in philosophy of religion and logic. He is the author of “Theism’s Pyrrhic Victory” in *The Southern Journal of Philosophy*, and a coauthor of “Overcoming Inefficient Reading Skills” in the *Journal of Educational Psychology*.