

ABSTRACT

Title of Dissertation: IDENTIFICATION AND AUTONOMY: A
MEDITATION ON THE PHILOSOPHY OF HARRY
FRANKFURT

Teresa Marie Chandler, Doctor of Philosophy, 2004

Dissertation directed by: Emeritus Professor Raymond Martin,
Department of Philosophy,
University of Maryland at College Park &
Professor, Union College, Schenectady, NY

Harry Frankfurt offers an account of freedom as “autonomy” in which identification plays a central role. Identification is supposed to be a process or psychic configuration by which attitudes – desires in particular – become or count in some sense as one’s own. When we are propelled to action by desires with which we “identify,” we’re acting autonomously. On Frankfurt’s view, whenever we act, a desire is involved, but sometimes the desire that moves us is one by which we do not want to be moved, a desire from which we are alienated – with which we fail to identify. Paradigm examples are addictions and compulsions. When moved by one of these “alien” desires, we lack autonomy.

Frankfurt’s account of autonomy, then, rests on a basic distinction. Of the desires that move us to action – “effective desires” – only some will be desires with which we identify. The main claim of this dissertation is that Frankfurt needs to

maintain this distinction, but in the end, doesn't. There are two basic problems. First, as Frankfurt develops his conception of identification, it shifts, and as it shifts, it becomes broader, so much so that it no longer marks the narrow internality. Second, neither of Frankfurt's alternatives – wholeheartedness, caring – clearly functions to mark out a narrow internality, either.

In the case of caring, Frankfurt gives an account that's dispositional – that is made out in terms of effective desire, so that in the end caring is not clearly distinguishable from having an effective desire – from simply being moved to action. In the case of wholeheartedness, Frankfurt introduces the concept as a way of understanding identification and therefore as a criterion of narrow internality, but he defines wholeheartedness in a way that presupposes a criterion of narrow internality.

Given the shifting conception of identification and the problems with wholeheartedness and caring, we are unable to distinguish between effective desires and desires that are truly an agent's own, and therefore, are left with an account of autonomy that remains unclear.

IDENTIFICATION AND AUTONOMY:
A MEDITATION ON THE PHILOSOPHY OF HARRY FRANKFURT

by

Teresa Marie Chandler

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2004

Advisory Committee:

Emeritus Professor Raymond Martin, Chair
Professor John Caughey, Dean's Representative
Professor Judith Lichtenberg
Professor Christopher Morris
Research Scholar David Wasserman

©Copyright by
Teresa Marie Chandler
2004

TABLE OF CONTENTS

I. INTRODUCTION	1
What is ‘identification’?	9
II. FRANKFURT’S EVOLVING VIEW	16
Frankfurt’s project	16
Identification as higher-order desire	21
Problems with identification as higher-order desire	24
Identification as decision	27
Problems with identification as decision	31
Identification as wholeheartedness	33
Problems with identification as wholeheartedness	36
III. EXPLICIT TO IMPLICIT: IDENTIFICATION AND CARING	38
Wholeheartedness and the unwilling addict	43
Caring.....	54
Problems with caring	61
Explicit to implicit – some conclusions	67
IV. TWO STANDARDS OF INTERNALITY	71
Two kinds of coherence	71
Two kinds of internality.....	74
Wholeheartedness and identification	78
V. WHAT MIGHT WORK	82
Why satisfaction is not passive	88
Choosing self	90
Degenerate desires	93
Authority	99
VI. CONCLUSION.....	100

I. INTRODUCTION

As Americans, we value freedom above all else. Our personal dignity and worth is bound up with our entitlement to its exercise; its protection is central to our way of life; our wars are fought under its banner. We are also, or so we like to think, pragmatic, “results-oriented,” and naturalistically rather than metaphysically inclined. Arguably this is true even of the pious among us. This dual commitment, to freedom on the one hand, and a largely naturalistic outlook on the other, creates difficulties when we attempt to articulate for ourselves what this freedom to which we are so fiercely committed is, when it is genuinely threatened, and how best to protect it.

An easy place to start is by understanding freedom as the liberty to move about, speak, and do as we please. This conception of freedom is often called freedom of action and it is defined in negative terms – as the absence of obstacles to getting (pursuing) what we want. Although we may relish this sort of freedom ourselves, we are less comfortable with our neighbors having it. We think that their liberty should stop at our property line – if not a little further back. Certainly, it should not block our view. Perhaps because we want to restrict our neighbors’ freedom, we are willing to accept restriction of our own. In fact, it is not even clear that we fully trust ourselves with this sort of freedom. Legend and myth are full of the woe that befell those who wished for the wrong things and were empowered to get them. In any case, the freedom under whose banner we fight is not the freedom to do entirely as we please, but the freedom to pursue certain valued ends in a manner consistent with others pursuing those valued ends as well.

But freedom of action is not the end of our concern with freedom. When we reflect on how it is that we come to want what we want, on whether we truly want ourselves to want it, or whether we can even help wanting it, we are asking questions not about freedom of action, but about freedom of will. When we apply the practical, technical framework that has been so useful for predicting and controlling events in the natural world to understanding ourselves we quickly generate a basis for doubting whether our wills are free. From this standpoint, it looks as if what we want and what we choose must be the consequence of causal chains leading back to determining factors that are external to us and over which we ultimately have no control. If these determining factors do not originate with us, although we might be free in the sense that we can do as we choose, it seems hard to see how we could have chosen otherwise than we did. If we couldn't have chosen otherwise, it's hard to see how our wills themselves can be understood as free.

There is a response to this worry, which goes as follows. What matters for freedom is that if circumstances had been (in some relevant way) different, you would have chosen differently. Although technically it's never the case that you could have chosen otherwise than you did, you had the capacity to choose otherwise in the sense that you would have chosen otherwise if (relevant) circumstances had been different – for example, if your concerns had been different. Furthermore, this response continues, the notion of the will *itself* as a causally independent, self-determining force is an illusion. The best explanation of what happens when someone acts makes reference not to the “will,” but simply to beliefs and desires.

This “belief-desire” theory of action is best exemplified in the work of Hobbes, though it is often also referred to as “Humean” by way of contrasting it to the principal rival view, put forth by Kant. This “belief-desire” view of action posits two basic kinds of mental states: beliefs and desires. Beliefs are about what is; they are capable of being true or false. Desires are about how we’d like the world to be. One intuitive way to characterize this distinction is as a matter of “direction of fit” between mind and world. Belief is about mind “fitting” the world; desire is about world “fitting” the mind.¹

Whenever anyone acts, on this view, we can explain his or her behavior in terms of a motivating reason constituted by a belief-desire pair. For example, Sue went to See’s to buy some mint truffles because she desired some chocolate mints and she believed that she could get some by going to See’s.² Acting out of respect for the moral law, on this view, would also be interpreted in terms of a motivating reason constituted by a belief-desire pair – a belief about what the law commands and a desire to honor the law.

Since, on this view, no one acts intentionally without a motivating reason, constituted by a belief-desire pair, it follows that no one acts freely or morally without such a pair, either. No one acts freely or morally without such a pair because all moral or free action is intentional and all intentional action is explicable in terms of belief-desire pairs. Such explicability is what it is to have acted intentionally,

¹ G. E. M. Anscombe, *Intention* (London: Basil Blackwell, 1957; reprint, Cambridge: Harvard University Press, 2000).

² For this view to be plausible, (1) “desire” must be understood formally, as Davidsonian “pro-attitudes,” rather than phenomenally, as attractions and passions; (2) desire need not be narrowly self-regarding.

which is to have *acted* at all. Consequently, there is no such thing as a will, independent of naturally occurring beliefs and desires. If beliefs and desires are the natural consequences of causal chains whose origin precedes even our own existence, so be it. What matters is freedom from coercion or external obstacles – being free to do as we please (including, if we so desire, to honor the moral law). In other words, on this view, freedom of action really is the best we can hope for. This is a view that's generally attributed to Hobbes, Hume, and Adam Smith, and it is the dominant view in contemporary social science.³

What's wrong with this response? One complaint about it is that it gives in too easily and gives up too much. Freedom of action is not freedom of will. Freedom of will might seem to require a will that is undetermined, and some have thought that the news that the universe is not fully deterministic creates the space for such will. However, if we want to defend a notion of freedom of will, it is not enough to show that our choices and motives are not somehow externally determined. In the first place, it is hard to make sense of how an entirely *undetermined* choice or act is free, rather than simply random or arbitrary. Random decisions and behavior would be unintelligible, and not therefore free in any recognizably human sense. Just because something is undetermined, that is, does not mean that it is free in any sense that we would value as freedom in persons.

A second reason for thinking that freedom of the will itself cannot be understood simply as the will's being undetermined is related to a further criticism of the Hobbesian/Humean response above. Even if we understand freedom of the will

³ Although contemporary belief-desire theorists acknowledge that decisions and other "intention-like states" may not be reducible to beliefs and desires. See Michael E. Bratman, *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge, UK: University of Cambridge, 1999).

as freedom to pursue what we truly want, there is a perfectly intelligible sense in which addictions and compulsions can rob us of this freedom. So one reason this response is not sufficient is that it doesn't deal with the phenomenon of internal obstacles – addictions and compulsions. If metaphysical freedom is constituted by (something like) our power of decision being an uncaused cause, either it can coexist with addictions and compulsions or it can't. If it does coexist with them, it is hard to see what such freedom amounts to. If it cannot coexist with addiction and compulsion, then it is hard to see how it is a metaphysical attribute of persons.

Finally, there is the paradoxical phenomenon of freedom of the will seeming to be best exemplified in cases in which the will is most tightly constrained. Harry Frankfurt refers to these cases as cases of “volitional necessity.” Martin Luther’s “here I stand, I can do no other,” as he refused to recant writings for which he'd been excommunicated, is the paradigm example, introduced into the contemporary discussion by Daniel Dennett,⁴ but connections between freedom and necessity can be seen at least as far back as the Stoics.

Many thinkers have been led to conclude that an adequate account of freedom of the will itself will not be a matter of a person's intentions and actions being *undetermined*, but being determined in the right way – or, on a non causal construal, developing in accord with the right sort of pattern. Kant held that free will is a presupposition of practical reason, that we understand the will as a causal power, and that *a causal power must be determined by law*, in this case, the law it gives to itself, which is the categorical imperative. Kant's account is an account of self

⁴ Daniel Dennett, *Elbow Room* (Cambridge, MA: The MIT Press, 1984).

determination, which he construes as determination by reason. Alternatively a free will could be one that is determined by the good, for instance, as dictated by God.⁵ Or a free will could be a will aligned with the laws of nature.⁶

Any account of freedom that understands it as a person's realized intentions being determined in the right way (or in accord with the right sort of pattern) requires distinctions among the possible ways it could be determined (or patterns according to which it could be intelligible as free). Such an account needs not only a way of telling whether a person is the source of his own actions by distinguishing "internal" from "external" determinates of willing, broadly understood,⁷ e.g., distinguishing my sitting down from my being pushed down, but also, when obstacles to freedom are in some sense internal, as with addictive craving, a way of distinguishing motives or

⁵ John Locke: "That God has given a rule whereby men should govern themselves, I think there is nobody so brutish as to deny. He has a right to do it...and he has power to enforce it by rewards and punishments of infinite weight and duration in another life...This is the only true touchstone of *moral rectitude*; and by comparing them to this law, it is that men judge of the most considerable *moral good or evil* of their actions: that is, whether as *duties or sins*, they are like to procure them happiness or misery from the hands of the ALMIGHTY." Book II, Chapter XXVIII, 8. John Locke, *An Essay Concerning Human Understanding*, new abridged ed. (London: Everyman, [1690] 1993).

"...the care of ourselves, that we mistake not imaginary for real happiness, is the necessary foundation of our *liberty*." Book II, Chapter XXI, 51. Locke, *An Essay Concerning Human Understanding*.

"...there is a case wherein man is at liberty in respect of *willing*, and that is the choosing of a remote good as an end to be pursued..." Book II, Chapter XXI, 56. Locke, *An Essay Concerning Human Understanding*.

⁶ Epictetus: Epictetus understands freedom as "learning to desire each thing exactly as it happens" and to learn freedom is to learn to "keep our wills in harmony with what happens" Epictetus asserts that there is a "law governing hypotheses – that we must accept what the hypothesis or premises demands. But much more important is the following law of life – that we must do what nature demands." Book I, xii 12-19, xxvi. 1-2. Epictetus, *Epictetus: The Discourses as Reported by Arrian, the Manual, and the Fragments*, ed. G. P. Goold, trans. W. A. Oldfather, 2 vols., vol. 1, *The Loeb Classical Library* (Cambridge, MA: Harvard University Press; reprint, 1989).

⁷ The causal account's "Where's the Agent" problem. This is the self that determinism threatens. See Yaffe for Locke's consciousness solution. Gideon Yaffe, *Liberty Worth the Name: Locke on Free Agency* (Princeton, NJ: Princeton University Press, 2000).

modes of willing that are authorized or authentic from those that are unauthorized or inauthentic.⁸

There are a number of proposals for how we ought to make these distinctions. The first distinction, that between my sitting down and being pushed down, is also described as a distinction between voluntary and involuntary movement or as a distinction between being active and being passive. The standard way to make this distinction has long been in terms of movement caused by will, motive, choice, or decision, versus movement resulting from any other cause. Hobbes, Locke, and Hume all explicitly endorse such a distinction, and this distinction serves at least as a starting point for virtually every account of human action.

The second distinction is less often made and enjoys less unanimity. The idea of an “authentic” decision, choice, or motive, in particular, emerges only after Kant. Kant himself drew a distinction between the will’s being self-determined, that is, autonomous, determined according to its own law, the categorical imperative, on the one hand, and the will’s being heteronomous, or determined by anything other than its own law – for example by desire or inclination. Kant lays groundwork for this distinction in a discussion of actions exhibiting moral merit.⁹ A merchant may refrain from overcharging a naïve customer because the merchant is merely inclined to do so or because it is what the moral law requires. In both cases, he acts voluntarily and in

⁸ Frankfurt says “A person acts autonomously only when his volitions derive from the essential character of his will.” Harry G. Frankfurt, *Necessity, Volition, and Love* (Cambridge: Cambridge University Press, 1999), p. 132. Kant distinguishes between heteronomous and autonomous willing, where heteronomy refers to cases in which an agent is ruled by his desires, but they are internal in the sense that they are *his* desires.

⁹ Immanuel Kant (1724-1804), *Ethical Philosophy: The Complete Texts of Grounding for the Metaphysics of Morals and Metaphysical Principles of Virtue*, trans. James Ellington (Indianapolis: Hackett, 1983). *Grounding*, section 1, 397.

accordance with the moral law, but only in the second, when acting from the motive of duty, does the merchant act autonomously. The motive of duty, then, for Kant, is what I have termed an “authorized or authentic” determinate of willing.

What I shall explore in this dissertation is whether and how a conception of “identification,” that is, roughly the process by which we take attitudes as our own, might be used to draw a distinction between willing that is “authorized or authentic,” which I will call “true,” and willing that is not. Identification is likely susceptible to as many interpretations as the distinction it’s supposed to be clarifying. In that sense, simply wheeling it in and letting it sit there won’t be much help. Instead, I shall explore some proposals about how to clarify this notion.

Contemporary discussion of this topic has taken off from Harry Frankfurt’s introduction of “identification” in his influential 1971 article, “Freedom of the Will and the Concept of a Person.” Frankfurt’s account of autonomy is among the best known contemporary accounts in the Hobbesian/Humean tradition. Philosophers working in various theoretical traditions have used his views as a starting point or foil for their own views, and “identification” has emerged as one point of contention.

I believe, that, ultimately, thinking about the psychological processes and other elements involved in “identification” will enable us to see agency in terms of integration rather than authority, which in turn may help us to better understand internal conflict and the development, breakdown, and recovery of agency. In this dissertation, however, my primary aim will be to get clear about the role that identification plays in Frankfurt’s account of autonomy as a means of investigating

the challenges in developing an account of autonomy and as a way of seeing what further distinctions may need to be drawn.

WHAT IS 'IDENTIFICATION'?

In the course of discussing his own conception of "surrogate self-identification" Raymond Martin refers to the "larger complex of identificatory processes that collectively figure so importantly in constituting selves." He goes on to say "For the most part philosophers have ignored these identificatory processes. This neglect is puzzling since mention of them often surfaces near the heart of an important philosophical thesis."¹⁰ The thesis he refers to is Frankfurt's thesis that identification provides a basis for distinguishing free action from free willing. Martin is right that not a lot of progress has been made in clarifying what this important phenomenon amounts to, and he is right that various conceptions of it are "near the heart of" what I'd amend as important philosophical theses.¹¹

¹⁰ Raymond Martin, *Self Concern: An Experiential Approach to What Matters in Survival* (Cambridge, UK: Cambridge University Press, 1998), p. 109.

¹¹ People are beginning to get around to it, but I think there is still much work to do. Part of the difficulty is that this work is somewhat interdisciplinary. Identification is also a phenomenon or set of phenomena that is being investigated empirically. However, I think there are useful conceptual distinctions to make in philosophical work that is already widely discussed, and that is what I will focus on here.

In his book, Martin focuses on neurophysiological pathology. He mentions Oliver Sachs, *An Anthropologist on Mars* New York, Knopf, 1995. In conversation, he also mentioned Antonio Damasio, "A Neurobiology for Consciousness," and Thomas Metzinger, "The Subjectivity of Subjective Experience: A Representationalist Analysis of the First-Person Perspective," both in Thomas Metzinger, *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, ed. Thomas Metzinger (Cambridge, MA: The MIT Press, 2000). They, along with philosopher/neuroscientists Steven Quartz and neuroscientist Terrence Sejnowski, are among many who have interesting discussions of this set of phenomena. Stephen R. Quartz and Terrence J. Sejnowski, *Liars, Lovers, and Heros: What the New Brain Science Reveals About How We Become Who We Are* (New York: HarperCollins, 2002). There seems to be a large research effort looking into the brain's "executive function" (see Richard Saltus "Lack Direction? Evaluate Your Brain's CEO," *the New York Times*, Aug 26, 2003 D1, as example of recent popular review) and I'm sure there's much there that might end up being illuminating.

Because Frankfurt's version of identification has been so influential – as David Velleman points out, “Frankfurt is responsible for bringing the term ‘identification’ into widespread use among contemporary philosophers and for shaping their intuitions about it”¹² – it is Frankfurt's version of identification, and his attempts to use it to draw distinctions among various modes of willing, that I will examine here.

In a series of articles, beginning with “Freedom of Will and the Concept of a Person,” Frankfurt elaborates an account of autonomy in which he uses his conception of identification as a basis for saying which desires are truly an agent's own – expressive of his will – and which (e.g. obsessions and compulsions) are alien – or obstacles to his will.

Now a person is active with respect to his own desires when he identifies himself with them, and he is active with respect to what he does when what he does is the outcome of his identification of himself with the desire that moves him in doing it...the notion of identification is admittedly a bit mystifying... however, it grasps something quite fundamental in our inner lives, and it merits a central role in the phenomenology and philosophy of human mentality¹³

Frankfurt defines autonomy as a person's being “active” rather than “passive” with regard to his “motives”¹⁴ and he intuits that this can be made out in terms of identification. Identification, in turn, enables us to distinguish motives that are, roughly, “internal” from those that are “external” to the agent. While Frankfurt offers this account of autonomy as (among other things) a clarification the freedom required

¹² J. David Velleman, "Identification and Identity," in *Contours of Agency*, ed. Sarah Buss and Lee Overton (2002), p. 105.

¹³ Harry G. Frankfurt, "Three Concepts of Free Action," in *The Importance of What We Care About* (1988), p. 54.

¹⁴ Harry G. Frankfurt, *The Reasons of Love* (Princeton: Princeton University Press, 2004). Quotation at p. 20, n. 5.

for moral responsibility, he asserts, “I do not think it is mainly for the sake of moral responsibility that we care as much as we do about being free.” Later he claims, “It is identification that indispensably constitutes the source and the ground of reasons. The creation of reasons is, indeed, close to being the whole meaning of it.”¹⁵

How exactly to understand identification and what one might need to say about it, though, have proven to be elusive challenges. Philosophers have taken issue with Frankfurt’s account of autonomy and one of the points of contention is his concept of identification and the work it’s supposed to be doing. No one, to my knowledge, denies that identification is a phenomenon central to our experience and understanding of agency and related phenomena such as responsibility and practical reason. Gary Watson characterizes identification as one of two major problems for compatibilist theories of free will (the other being causation), but he notes that “difficulties with the concept of identification are everyone’s problem,” not just a problem for compatibilists.¹⁶ As Sarah Buss and Lee Overton, editors of a recent volume of essays on Frankfurt’s work, point out, “Even incompatibilists concede that being free from determination by events in the past does not *suffice* for being morally responsible for one’s actions.”¹⁷ Identification, whether as cause, consequence, or corollary, is certainly involved. You might think that Kantians possess a theoretical apparatus that enables them to bypass identification, explanatorily at least, but Barbara Herman points out that, like Frankfurt, many contemporary Kantians,

¹⁵ Harry G. Frankfurt, “Reply to Richard Moran,” in *Contours of Agency*, ed. Sarah Buss and Lee Overton (2002), p. 209.

¹⁶ Gary Watson, “Free Action and Free Will,” *Mind* (1987): p. 151.

¹⁷ Sarah Buss and Lee Overton, eds., *Contours of Agency: Essays on Themes from Harry Frankfurt* (Cambridge: The MIT Press, 2002), p. xiv.

Christine Korsgaard in particular, eschew Kantian metaphysics, which leaves them having to “bootstrap their way from essentially Humean materials, plus a more complex story of the needs and identity conditions of agency to the autonomous will...”¹⁸ In fact, in her account of normativity, Korsgaard leans heavily on a conception of identification, claiming that, “the reflective structure of human consciousness requires that you identify yourself with some law or principle which will govern your choices. It requires you to be a law unto yourself. And that is the source of normativity.”¹⁹ Michael Bratman credits Frankfurt’s work on identification with having broken a “tendency in the philosophy of action to limit attention to two main types of theories” and having “mapped some of the ‘contours’” of a “middle ground” between Kantian and Humean models of action.²⁰ I think that one reason people are being drawn to figure out just what identification is, is that it seems to hold the key to this middle ground.²¹

¹⁸ Barbara Herman, "Bootstrapping," in *Contours of Agency*, ed. Sarah Buss and Lee Overton (2002).

¹⁹ Christine M. Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), p. 103-04.

²⁰ Michael E. Bratman, "Hierarchy, Circularity, and Double Reduction," in *Contours of Agency*, ed. Sarah Buss and Lee Overton (2002), p. 79.

²¹ Watson kicked off a critical discussion by suggesting that identification could play the role in a theory of agency that Frankfurt wished only if identification were itself understood in terms of some evaluative capacity not itself reducible to desiring. Gary Watson, "Free Agency," in *Free Will*, ed. Gary Watson (1982), 96-110. Geoffrey Sayre-McCord and Michael Smith are currently working on a paper in which they flip this. They agree that evaluation is necessary for autonomy, but they attempt to show that identification is essential for an account of evaluation, not vice versa. Michael Bratman and David Velleman both joined in the debate, taking a number of passes at characterizing identification and its role in theory of action. Bratman, who has a forthcoming article, "Desires of one's Own," thinks intention is what is missing from belief-desire accounts. (See Bratman, *Faces of Intention: Selected Essays on Intention and Agency*.) He tries to assimilate identification into intention formation and has thus far favored an account centered on "deciding to treat a desire as a reason." (Ibid., See also Bratman, "Hierarchy, Circularity, and Double Reduction.", Michael E. Bratman, "Identification, Decision, and Treating as a Reason," in *Faces of Intention: Selected Essays on Intention and Agency* (1999).) Velleman, in a forthcoming book, *Self to Self*, in which he distinguishes three senses of identification corresponding to three conceptions of self, all of which may be in play, but not distinguished, in Frankfurt's work. The account of identification Velleman thinks is pertinent

All philosophies of action include reference to beliefs and desires. For some philosophers (and many social scientists) beliefs and desires suffice to explain what happens when someone acts. Others, Kantians in particular, question whether desires are also necessary. Many philosophers, however, both Kantian and Humean, have thought that beliefs and desires are not sufficient. These philosophers debate over what more might be needed. Frankfurt supplements beliefs and desires with identification.

Identification is a process or psychic configuration by which attitudes, desires in particular, become or count in some sense as one's own. The basic idea is that when we are propelled to action by desires with which we "identify," we're acting autonomously. A creature like us, a human, has all sorts of desires. Whenever we act, on Frankfurt's view, a desire is involved. But sometimes the desire that moves us is one by which we don't want to be moved, a desire from which we are alienated, with which we fail to identify. Paradigm examples are addictions and compulsions. When we're moved by one of these "alien" desires, we are not autonomous. Frankfurt's account of autonomy, then, rests on a basic distinction. Of the desires that move us to action – "effective desires" – only some will be desires with which we identify.

So, we have "effective desires" that in some broad sense belong to us and we have effective desires with which we identify, which I call here "true" desires, which

to agency is similar to one offered also by Joseph Raz. (Joseph Raz, "We Are Ourselves: The Active and the Passive," in *Engaging Reason: On the Theory of Value and Action* (Oxford, 2000).) Daniel Dennett discusses the issue in *Freedom Evolves*, endorsing an early take of Velleman's on what Frankfurt perhaps ought to have said. Bratman, Velleman, T. M. Scanlon, and Richard Moran, who, like Velleman, distinguishes among kinds of internality, all take on the issue of identification in a recent collection of essays about Frankfurt's work, *Contours of Agency*.

belong to us in a narrower sense. The main claim of this dissertation is that Frankfurt needs to maintain this distinction, but in the end, doesn't. There are two basic problems.

The first problem is that as he develops his conception of identification, it shifts, and as it shifts, it becomes broader, so much so that it no longer marks the narrow internality. For example, with the broader conception of identification we can no longer distinguish the unwilling addict's desire to take the drug and his desire to refrain in terms strictly of identification, because in the broader sense, the unwilling addict will usually identify with both desires.

The second problem is that none of Frankfurt's alternatives clearly functions to mark out a narrow internality, either. Frankfurt develops a couple of closely related concepts – wholeheartedness and caring – that look as if they might do the work identification originally is supposed to do, mark certain desires as narrowly internal. I argue that neither of these concepts does the job. In the case of caring, Frankfurt gives an account that's dispositional, that's basically made out in terms of effective desire, so that in the end caring is not clearly distinguishable from having an effective desire – from simply being moved to action.

The case of wholeheartedness is complicated. One immediate problem is that the concept of wholeheartedness is introduced as a way of understanding identification and therefore as a criterion of narrow internality, but it's also defined in a way that presupposes a narrow criterion of internality that's left unspecified. So it looks as if there may be two narrower internalities, one of which is not explained, and that identification is mixed up with both.

Given the shifting conception of identification, and given the structure of these other concepts, where that leaves us is unable to clearly distinguish between effective desires and desires that are in some sense truly an agent's own, and therefore, with no account of autonomy – no account that works only with a single simple conception of identification, that is.

In spite of this, I wonder whether Frankfurt might still be able to capture some of our intuitions about volitional struggle without the identification that was supposed to define the narrow internality. I look at some problems – passivity and “degenerate desires” – that Frankfurt's general approach may be saddled with, and review responses Frankfurt does or could make.

II. FRANKFURT'S EVOLVING VIEW

FRANKFURT'S PROJECT

“*Consciousness*” vs. *reflexiveness*. In his early and best known article on autonomy, “Freedom of the Will and the Concept of a Person,” Frankfurt starts out trying to get at what is special about us. He is responding to an article by Peter Strawson in which Strawson assumes that “consciousness” is sufficient, to mark us as persons.²²

Frankfurt rejects this “consciousness” approach, which he later, more accurately, characterizes as “mentality.” His basic point is that mentality, the mere having of mental states, beliefs and desires, is not sufficient for personhood. Many entities might count as having beliefs and desires yet still not count as persons – some animals, for instance. While Frankfurt doesn’t go so far as to ask what it is “like” to be us, his initial rejection of this “consciousness” or mentality criterion asks us to introspect and notice that we not only have beliefs and desires – attitudes – but that we also have attitudes about those attitudes. He’s right. We do understand and recognize ourselves as having certain attitudes, and as we bring these before us – or as they emerge in consciousness – we find new attitudes forming in response. This is the “reflexive” capacity.²³

²² P. F. Strawson, *Individuals* (London: Methuen, 1959), pp. 101-02.

²³ Compare to Daniel Dennett, “Conditions of Personhood,” in *The Identities of Persons*, ed. Amélie Rorty (Berkeley, CA: University of California Press, 1976). Compare also to Carol Rovane, *The Bounds of Agency* (Princeton: Princeton University Press, 1998). Contrast with Korsgaard, *The Sources of Normativity*. Contrast also with Rawls (John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971).) and others who emphasize a gap between self-awareness and action which must be closed by reason. Also, Frankfurt moves from “reflexivity” to “reflection” without explicitly discussing the shift.

You might think, as some current accounts suggest and Frankfurt later recognized,²⁴ that consciousness entails reflexivity, and that if Frankfurt thinks that reflexivity is key, then an updated, more detailed account of consciousness might be sufficient for personhood – or a distinguishing feature of entities that are persons. Frankfurt does not rest with reflexivity, however. He goes on to give an account of “volition,” pursuing a characterization of the “inner organization of the will,” “how we are to conceptualize ourselves,” and “what defines the identities we achieve.” He does this because he is interested not just in distinguishing entities that are persons from entities that are non-persons, but in articulating a fuller characterization of persons that departs from what he takes to be the dominant view. In contrast to “reason,” which he claims is said to be “most distinctive,” Frankfurt wants to show that “volition pertains more closely to our experience of ourselves and to the problems in our lives that concern us with the greatest urgency.”²⁵

Volition vs. reason. There are on the table two candidate identifying or distinguishing features for persons: reflexivity and reason. Why does Frankfurt choose to minimize reason? I think that Frankfurt may think that there is a plain-vanilla sense of reason in which any entity whose behavior is susceptible to intentional explanation can be said to “reason.” One takes up such a reason-attributing “intentional stance,” according to Daniel Dennett, when dealing with entities as “systems” “too complex to be dealt with effectively from the other [physical, mechanistic] stances...here one assumes...rationality in the system.” “The

²⁴ Harry G. Frankfurt, “Identification and Wholeheartedness,” in *The Importance of What We Care About* (1988).

²⁵ Harry G. Frankfurt, *The Importance of What We Care About: Philosophical Essays* (Cambridge: Cambridge University Press, 1988), p. viii.

success of the stance is...a matter settled pragmatically.”²⁶ The intentional stance, which attributes rationality to its object, is not yet the “personal” stance, however. Rather, it is an important precondition for the “personal” stance. This is a point on which Frankfurt agrees:

In maintaining that the essence of being a person lies not in reason but in will, I am far from suggesting that a creature without reason may be a person. For it is only in virtue of his rational capacities that a person’s capable of becoming critically aware of his own will and of forming volitions of the second-order. The structure of a person’s will presupposes, accordingly, that he is a rational being²⁷

This kind of reason is necessary, but clearly not sufficient to mark an entity as a person.²⁸ Reflexivity, by contrast, as Frankfurt is using it, indicates a level of complexity which qualifies an entity as a person, a level of complexity that the minimal “intentional systems,” the basic calculating agents Frankfurt calls “wantons,” creatures who act on their desires without reflectively considering whether they want to be moved by these desires, do not achieve.

But neither reason nor reflexivity (think consciousness), considered in themselves, seem to get at a further distinction, the active/passive distinction that Frankfurt says is “at the heart” of our existence as selves. Frankfurt has no account of

²⁶ Daniel Dennett, “Mechanism and Responsibility,” in *Free Will*, ed. Gary Watson (Oxford: 1982), p. 157. Note that Dennett equates the personal stance with the moral stance, while Frankfurt would maintain a distinction between the two.

²⁷ Harry G. Frankfurt, “Freedom of the Will and the Concept of a Person,” in *The Importance of What We Care About* (1988), p. 17.

²⁸ Frankfurt says that animals have the capacity to respond to reason “My own construal of reason and of rational behavior do not include any essential reference to beliefs”; “natural selection can be expected to see to it that many organisms...respond to reasons...” (Harry G. Frankfurt, “Reply to Eleonore Stump,” in *Contours of Agency*, ed. Sarah Buss and Lee Overton (2002), p. 62.). However Frankfurt also says, “Animals of many species have desires, but only animals of our species—or, perhaps, of a few others—are capable of seeing anything as a reason” (Harry G. Frankfurt, “Reply to T. M. Scanlon,” in *Contours of Agency*, ed. Sarah Buss and Lee Overton (2002), p. 184.). Furthermore, animals also lack the capacity for happiness because they lack the capacity to conceive of something as important to them (Frankfurt, *Necessity, Volition, and Love*, p. 158.).

how reflexivity itself might be “active.” Instead, he gives an account of “volition” in which reflexivity plays a central role.

I think it’s worth pausing a minute to compare Frankfurt with Kant. Frankfurt conspicuously follows Kant in placing primary importance on the will. Both Frankfurt and Kant think the most important fact about persons is their capacity for autonomy. Kant makes autonomy the source of value; Frankfurt makes it flow from his value analogue, love.²⁹ In Frankfurt’s view, we are most free when we act wholeheartedly on the basis of love. Frankfurt intends to give an account of autonomy that is purely phenomenal. Instead of a transcendental argument about what practical reason presupposes Frankfurt tries to work within the limits of what experience reveals. Like the existentialists, he rejects Kant’s conclusion that universal legislation and willing in accordance with the moral law are necessary for autonomy.³⁰ Like Nietzsche, (at first³¹) Frankfurt believes that autonomy is a matter of self-mastery, which doesn’t require any such universalizing or conformity. Kant believed that practical reason commits us to being universal legislators and to the moral law. Insofar as practical reason is identified with these presuppositions or entailments, Frankfurt, for his purposes, rejects it. That is, he claims that practical reason in Kant’s sense, which I’ll call Big-R Reason, is *not* necessary for autonomy

²⁹ At first Frankfurt claims that love has authority and the action in which love finds expression is autonomous because...love is active. Later, he realizes that this is circular and takes the stand that love is a precondition of activity. Harry G. Frankfurt, "Reply to Jonathan Lear," in *Contours of Agency*, ed. Sarah Buss and Lee Overton (2002).

³⁰ This rejection is sometimes thought of as the existentialist interpretation of Kant’s “reciprocity thesis.” For details see (draft) Ian Duckles, "Kant's Reciprocity Thesis: Towards an Existentialist Understanding," (2003). Citing Sartre, “Existentialism as Humanism,” Judith Lichtenberg suggests that the existentialists didn’t necessarily reject Kant’s conclusion about the dependence of autonomy on universal legislation and moral law.

³¹ He later decides that autonomy is a matter of wholeheartedness, which doesn’t require any such universalizing or conformity either.

or, in his terms, volition, or for personhood, or to account for the “importance of what we care about.”³²

So, having seen the importance Frankfurt places on the will and on reflexivity, now we’re in a better position to see where Frankfurt’s headed. He is rejecting both a calculative kind of reason as a distinguishing feature of persons, and a universally legislating, moral-law yielding, noumenal-will presupposing kind of Reason as a defining characteristic. Small-r reason does not distinguish persons; Big-R reason does not characterize them.

So what does Frankfurt offer instead? In what follows, I shall look at four major themes that occur in Frankfurt’s repeated attempts to articulate what constitutes freedom of will. His challenge is to be able to distinguish ways in which the will is determined or configured that constitute its freedom from those that do not. Because, following Hume, he eschews reason itself as a possible motive, he focuses primarily on making distinctions among desires in terms of whether or not we “identify” or “are identified” with them. His starting place is reflexivity, where he focuses on the hierarchical nature of desire. He understands freedom as acting in accord with “higher-order” desires. When his desire hierarchy threatens to evaporate into an infinite regress, he explores the possibility of capping it off with “decision,” as in a “decisive commitment” to one desire rather than another. Freedom here is associated with acting on desires to which we are decisively committed. He ultimately decides that this takes him in the wrong direction (i.e. in a circle), and he shifts focus from the structure of isolated pairs of motives to the state of the psyche as a whole. Here

³² Frankfurt speculates that even what I’m calling little-r reason (Frankfurt is talking about deliberation about ends) depends on volition – on making up one’s mind. Frankfurt, “Identification and Wholeheartedness,” p. 176.

freedom is associated with a certain coherence of motives – “wholeheartedness” – which enables action. As Frankfurt looks at the psyche as a whole, he finds he needs to draw another distinction between desires. This time, instead of first and second-order desires, he distinguishes between second-order desire and caring. Here freedom is associated one’s actions being necessitated by a particular form of caring that Frankfurt calls, “love.”

IDENTIFICATION AS HIGHER-ORDER DESIRE

Frankfurt begins by offering a minimalist account of the will that is arguably an improvement on early empiricist models.

According to Frankfurt, an entity’s will is its “effective desire” – the desire that in fact moves it to action on a given occasion. In this sense, animals, too, have wills. What persons have in addition, that nonpersons do not, is a capacity for forming attitudes about those motives. In particular, we can form desires about which (of our) desires we’d like to move us to action. This is Frankfurt’s well known “hierarchy.” The basic idea is that a person is free, by which Frankfurt means functioning distinctively and characteristically as a person, in other words, autonomously, not simply when he is moved by his “strongest” desire, like a dog chasing a rabbit, or even when he consciously chooses the option or path toward which he feels most drawn. Rather, he exhibits autonomy when the desire that moves him is the one by which, upon reflection, he desires to be moved. Another way to look at this is to see that, as Frankfurt understands it, a “first-order” desire is a desire to do, whereas a “second-order” desire – a desire about that desire’s efficacy – is a desire to be. That is, a second-order desire is the desire to be a person who is moved

by ‘this desire, but not that.’ Persons have the capability of caring about their motives; animals and other “wantons” don’t.

Frankfurt speaks of desiring to be moved by a particular desire in terms of reflectively “identifying” with or “endorsing”³³ that desire. To use one of his examples, even an addict may exhibit autonomy when he “willingly” shoots up if he endorses his compulsion to do so. Imagine a case in which an addict has to perform an important task and is only enabled to do so on this occasion by taking the drug – or imagine a socially admired compulsion, like “workaholism.” Alternatively, an addict might disapprove of his compulsion, regard it as alien – as a disease, for instance – and desire to rid himself of it. In this case, on at least some of the occasions on which he shoots up, both he and we will be inclined to understand his shooting up as his being overcome by a compulsion, as a case of his autonomy being compromised.

In this way, Frankfurt, in effect, proposes to overcome a limitation of belief-desire accounts of free action.³⁴ The limitation is that these basic belief-desire accounts lack the resources to account for the familiar phenomenon of volitional struggle and of the sense in which psychological compulsion robs us of our freedom – the sense in which, as in the second case above, an addict’s choice to take the drug, though intentional, may not be free. This limitation arises because a simple belief-desire account presupposes that action is the upshot of a motivational tussle in which the “strongest” desire wins. Therefore, by definition, the fact that one acted on a

³³ A term he later retracts. See replies to Watson and Bratman in *Contours of Agency*: Harry G. Frankfurt, "Reply to Gary Watson," in *Contours of Agency*, ed. Sarah Buss and Lee Overton (2002), Harry G. Frankfurt, "Reply to Michael E. Bratman," in *Contours of Agency*, ed. Sarah Buss and Lee Overton (2002).

³⁴ This point is made by Gary Watson. Watson, "Free Action and Free Will," p. 147.

given desire is evidence that that desire was the person's strongest. Consequently all the actions one performs are the actions one most strongly wanted to perform. This is a commonsense way of understanding our behavior, it is enshrined in the wisdom that "actions speak louder than words," but it begins to break down when we try to account for the sense in which we sometimes lose battles for self-control. Cases of addiction and compulsion suggest that one's "strongest" desires are not necessarily one's truest or most authentic. The simple belief-desire account apparently lacks the resources to make this distinction.

What Frankfurt originally proposes is a distinction between "effective" desires – desires that move us all the way to action – and desires with which we "identify." With Frankfurt's concepts of identification and of higher-order desiring we are supposed to be able to say why a person's acting on or being moved by his "strongest desire" by selecting the option or path toward which he feels most drawn may in some cases coincide with what he wants to want, and hence constitute autonomy or "volition," but in other cases acting on or being moved by his "strongest desire" may fly in the face of what he wants to want, or fail utterly to cohere with it. In that case, when he acts on his strongest desire, he does not exhibit autonomy.³⁵

³⁵ In order for a person's strongest desire to cohere with what he wants to want, what he wants to want must itself be coherent. Coherence among our desires – or at least among a core set of higher order desires constituting a "coherent affective or motivational identity" – is something Frankfurt explores later under the heading of "wholeheartedness." Harry G. Frankfurt, "The Faintest Passion," in *Necessity, Volition, and Love* (1999). According to Frankfurt, we have a powerful motive to achieve such coherence, a coherence he defines in contrast to "ambivalence." Arguably, we may also have a practical *requirement* for a degree of coherence among our desires, not to mention beliefs and actions, e.g., see Christine M. Korsgaard, "Personal Identity and the Unity of Agency: A Kantian Response to Parfit," *Philosophy and Public Affairs* 18 (1989). Frankfurt is less explicit about this practical requirement. While Frankfurt allows that one could accept being unalterably ambivalent, his primary claim is that there is something inherently self-defeating about ambivalence – that it is the motivational equivalent of irrationality. Therefore, he claims, "it is a necessary truth that we wholeheartedly desire to be wholehearted" (Unfortunately, ambivalence cannot be overcome cognitively and "cannot be overcome volunteristically." Frankfurt, "The Faintest Passion," p. 106.) As entities, then, with the

Frankfurt's introduction of higher-order desires into a more basic belief-desire model of action is supposed to enable us to distinguish between an entity's acting on or being moved by his strongest desire and his acting on or being moved by the desire upon which he reflectively wants to act, but without bringing in concepts like evaluation. Acting on one's strongest desires has historically seemed inadequate to distinguish persons from non persons. While we may value our freedom to do as we please, the thought that we are determined by our desires has needed to fit together with another thought – the thought that we can and do control their effectiveness. So, there is a further sense of freedom that corresponds to a further sense of control – self-control, or autonomy, which is both distinctive and characteristic of persons.

PROBLEMS WITH IDENTIFICATION AS HIGHER-ORDER DESIRE

Frankfurt's distinction between first- and second-order desires, desires to do and desires to be, among other things, was supposed to help capture the sense in which, if freedom is a matter of getting what we want, we can be prevented from getting what we want not only by external obstacles, but by our own motivational systems.³⁶ Here getting what we most want involves having the will we want. But in order to make sense of this kind of conflict within our motivational system, we have to distinguish among motives. If all motives are lumped together as desires it can be difficult to sustain a distinction between what we most want and what we are actually moved to do, since, in simple belief-desire terms, what we most want just is what we are actually moved to do.

capacity for forming attitudes about our motives, so motivated and so bound, we are free when we are moved by desires we wholeheartedly want to be moved by.

³⁶ This is Gary Watson's formulation. Watson, "Free Agency."

A classic way to draw a distinction between what we most want and what we are actually moved to do is to distinguish between reason and appetite or spirit, or reason and inclination. Frankfurt rejects this route. He originally distinguished among motives, between what I'll call "true" or authentic desires and merely effective desires, by ordering them. He distinguished between first-order desires, desires to do *x*, and desires about those desires, desires to *be* moved by the desire to do *x*. Those first-order desires by which one also wants to be moved are desires with which one thereby – by having the relevant higher-order desire, that is – "identifies." The higher-order desire thus authenticates the lower, such that being moved by these desires with which we identify expresses not just our will but our "volition."

A number of questions can be asked about this account. Perhaps chief among them is that if the higher-order desire authenticates the lower-order desire, what then authenticates the higher-order desire? Gary Watson presses this point in his article "Free Agency":

One job that Frankfurt wishes to do with the distinction between lower and higher orders of desire is to give an account of the sense in which some wants may be said to be more truly the agent's own than others (though in an obvious sense all are wants of the agent), the sense in which the agent 'identifies' with one desire rather than another and the sense in which an agent maybe unfree with respect to his own 'will'...But we can see that the notion of 'higher-order volition' is not really the fundamental notion for these purposes, by raising the question: Can't one be a wanton, so to speak, with respect to one's second-order desires and volitions?³⁷

Here, Watson points out that having second-order desires does nothing to save one from "wantonness." "The essential character of a wanton," according to Frankfurt, "is that he does not care about his will. His desires move him to do certain things,

³⁷ Ibid., pp. 107-8.

without its being true of him either that he wants to be moved by those desires or that he prefers to be moved by other desires.”³⁸

Watson argues that one can be a “wanton” with regard to one’s second-order desires, too. One can have second-order “acculturated attitudes,” for example about material success as a sign of divine election, which exist independently of one’s own best current judgment. Because of this independence and grounding in acculturation, these attitudes, like appetites, are “non-rational.” Acculturated attitudes “are often expressed in evaluative language” and generate “guilt.” “But, since conflict is possible here, to want something as a result of acculturation is not thereby to value it, in the sense of ‘to value’ that we want to capture.”³⁹ Frankfurt is aware of this difficulty, Watson acknowledges, but Frankfurt’s “answer (which I will explore below) that one just makes a ‘decisive commitment’” as an act of identification (to a pre-existing desire) is “unhelpful” and is “in any case different” from the notion “of a second- (or n-) order desire.”⁴⁰

Frankfurt partially concedes Watson’s point about the problem with higher-order desires. Here is Frankfurt’s summary:

On hierarchical accounts, a person identifies with one rather than with another of his own desires by virtue of wanting to be moved to action by the first rather than the second. For example, someone who is trying to quit smoking is identified with his first-order desire not to smoke, rather than his concurrent first-order desire for another cigarette, if he wants the desire not to smoke to

³⁸ Frankfurt, “Freedom of the Will and the Concept of a Person,” p. 86.

³⁹ Watson, “Free Agency,” p. 105.

⁴⁰ *Ibid.*, pp. 108-9. In the end, Watson’s complaint seems to be that it’s not the mere fact that a given desire is a *higher* order desire that does the work, it’s that the higher order desire tends to be the result of a certain kind of thinking assessment and belief – an all-things-considered judgment. Frankfurt’s higher order desire account seems plausible just because higher order desires result from something like one’s all-things-considered judgments about what one ought to do. It’s these judgments, these evaluations that do the *work* Frankfurt thinks second order desires do.

be the one that effectively guides his conduct. But what determines whether he identifies with this second-order preference? Considered in itself, after all, his desire to defeat the desire to smoke is just another desire. How can it claim to be constitutive of what he really wants? The mere fact that it is a second-order desire surely gives it no particular authority. And it will not help to look for a third-order desire that serves to identify the person with this second-order preference. Obviously, the same question would arise concerning the authority of that desire; so we would have to find an even higher-order desire; and so on endlessly. The whole approach appears to be doomed.⁴¹

Frankfurt understands the problem with his first hierarchical model as a regress problem.

In his initial attempt to stop this regress, Frankfurt devoted two articles (which I discuss below) to further developing the notion of identification as ‘decisive commitment.’ Watson had suggested that it was this notion that was doing the authorizing work in Frankfurt’s initial account, that ‘decisive commitment’ was distinct from the concept of a higher-order desire and that it was, so far, underdescribed.

IDENTIFICATION AS DECISION

The distinction Frankfurt is trying to make, that he needs for his account of autonomy, is a distinction between desires that in the “gross literal sense” belong to a person, and desires with which a person is more intimately “identified.”⁴² In “Identification and Externality,” Frankfurt gives up on making this distinction between desires that belong to a person and desires with which a person is

⁴¹ Frankfurt, "The Faintest Passion," p. 105.

⁴² I have been saying that Frankfurt needs to distinguish between effective desires and true desires. In his original article, he stipulated that he was working with a formal, or philosophical interpretation of desire, e.g. that which explains behavior, like Davidson’s “pro-attitude.” In this article, he is treating desires more as felt passions.

“identified” in terms of hierarchies of desires.⁴³ He begins instead by exploring the possibility that desires that in a “gross literal sense” belong to a person may still be “external” in the same sense that “a mere happening in the history of his body” is distinct from “a movement that a person makes.”⁴⁴

This analogy suggests that a more basic way to distinguish desires with which one is identified from desires with which one is not is in terms of desires with regard to which one is *active*, from desires with regard to which one is passive. But whether a desire is internal in the narrower sense (one with which we’re identified) or external (one with which we are not identified), then, is not just a matter of a higher-order desire one might have about it, because one’s higher-order desire may itself be passive, as in the “acculturated” attitudes above. This susceptibility produces a regress that “precludes explication of the concepts internality and externality by appealing merely to the order of attitudes.”⁴⁵

How is it that one is active with regard to some of one’s desires, then?

Frankfurt’s answer is that one identifies via “decision” and decision is necessarily active. In order to clarify what he means by “identifying” with a desire, and how this

⁴³ Harry G. Frankfurt, “Identification and Externality,” in *The Importance of What We Care About* (1988).

⁴⁴ According to Frankfurt:

There is in fact a legitimate and interesting sense in which a person may experience a passion that is external to him and that is strictly attributable neither to him nor to anyone else. Recognizing this need not prevent us from agreeing with Penelhum that such a passion is part of the person’s ongoing history. It may be noted, moreover, that declining to attribute to a person certain of the passions he experiences does not commit us to regarding those passions as altogether irrelevant in reaching a fair judgment concerning what we can expect from him. A passion is no less genuine, and its thrust is no less forceful, for being external to the person in whose history it occurs, any more than a bodily movement is less palpable in its occurrence or in its effects for being a movement that is not made by the person in whose body it occurs. *Ibid.*, p. 61-62.

⁴⁵ Frankfurt, “The Faintest Passion,” p. 66.

identification could be enough to authenticate the desire in question, to make it a “true” desire, independently of its strength or efficacy, Frankfurt explains what he means by “decision.” He uses the example of a person deciding when to stop checking his arithmetic calculations.

Now what leads people to form desires of higher orders is similar to what leads them to go over their arithmetic. Someone checks his calculations because he thinks he may have done them wrong. It may be there is a conflict between the answer which, for one reason or another, he believes may be correct; or perhaps he has merely a more generalized suspicion, to the effect that he may have made some kind of error. Similarly a person may be led to reflect on his desires either because they conflict with each other or because a more general lack of confidence moves him to consider whether to be satisfied with his motives as they are⁴⁶

This example illustrates the practical challenge posed by our reflective nature: a person is often in the position of needing to reflectively resolve his various attitudes into a single intention.

Unlike choice,

decision is something that we do to ourselves. In this respect it differs fundamentally from making a choice, the immediate object of which is not the chooser, but whatever it is that he chooses. This difference between deciding and choosing accounts for the fact that deciding to make a certain choice is not the same as actually making it (after all, the time or occasion of doing that may not yet have arrived), whereas deciding to make a particular decision (that is, deciding to decide things a certain way) cannot be distinguished from making the decision itself.⁴⁷

In making a decision by which he identifies with a desire, [he] *constitutes himself*.⁴⁸

The fact that the object of the decision is the self reveals how Frankfurt thinks he is now able to preserve the true desire/effective desire distinction. “All a decision

⁴⁶ Frankfurt, "Identification and Wholeheartedness," p. 169.

⁴⁷ Ibid., p. 172.

⁴⁸ Ibid., p. 170.

does is to create an intention; it does not guarantee that the intention will be carried out.”⁴⁹ In other words in deciding in favor of a desire, in identifying with it, one is creating an intention. We all understand that even the best intentions don’t always lead to action; therefore, if true desires are intentions, we can understand how they can be distinct from effective desires.

How are these intentions created? Frankfurt appears to be saying that general uncertainty or, more narrowly, concern about discrepancy – in the math example, between the results of one’s calculations now and results of future calculations; in the personal example, among one’s desires or between one’s ideals and one’s actions – prompts a cycle of re-checking one’s calculations or reflecting upon one’s motives. Assume we have the re-checking and reflection dynamic set up. How do we stop? How do we achieve a resolution? In the arithmetic case, how do we complete the action (calculation), given the possibility of discrepancy, given that we might be making a mistake? In the personal case, how do we know which conflicting desire to act on, given the possibility of discrepancy, given that we might be making a mistake?

Frankfurt says that the checking process may end arbitrarily, wantonly, or one might decide “*for some reason*” – ‘make up one’s mind’ to stop.⁵⁰ The decision is a “commitment that resounds endlessly”: one expects either the answer or the principle on which a decision to stop was made (e.g. cost-benefit) will be “endlessly” confirmed, that no further accurate inquiry would require one to change one’s mind. “A person is justified in terminating a sequence of calculations or reflections when he

⁴⁹ Ibid., p. 174.

⁵⁰ Ibid., p. 168.

sees no conflict to be avoided or resolved.”⁵¹ In other words, one decides “for some reason” when one has achieved a reflective resolution.

PROBLEMS WITH IDENTIFICATION AS DECISION

While it may be true that identification and decision are bound up together, there remains in Frankfurt’s own account a fundamental ambiguity about what constitutes decisive identification.

suppose that a person with two conflicting desires identifies with one rather than the other. This *might* cause the other – the desire with which the person does not identify – to become substantially weaker than it was ...But it need not...possibly the conflict between the two desires will remain as virulent as before...What the person’s commitment to the one eliminates is not the conflict between it and the other. It eliminates the conflict *within the person* as to which of these desires he prefers to be his motive.⁵²

But later he says,

it may be that energies tending toward action inconsistent with the intention remain untamed and undispersed however decisively the person believes his mind has been made up...In that case the decision, no matter how apparently conscientious and sincere, is not wholehearted,⁵³

suggesting that conflict within the *person* has not been eliminated after all.

Although Frankfurt more consistently avows the first of these possibilities – that is, he asserts that wholeheartedness and self-satisfaction are compatible with “virulent” psychic conflict – in Chapter IV, below, I’ll explore whether he’s established the grounds to do so.

Return, now, to Watson’s criticism. One way of putting Watson’s criticism is to say that although Frankfurt has given us a clear account of effective desire (or it’s

⁵¹ Ibid.

⁵² Notice that this restricts the scope of “person,” so that it is narrower than common usage would lead one to expect. Ibid., p. 172.

⁵³ Ibid., p. 174.

relatively easy for us to agree upon what an effective desire is) Frankfurt hasn't yet given a convincing account of true desire.⁵⁴ On Frankfurt's account we have conflicting desires, and desires about those desires, but no clear way of saying which desires are true, authentic, or authoritative. The idea of higher-order endorsing desires looks good until you realize that it implies that to be true/authentic (as opposed to wanton), a desire needs to be endorsed. Then you ask, what about those endorsing desires? What's endorsing them? If nothing's endorsing them, what prevents them from being wanton?

Frankfurt's "decision" answer is that a second-order desire is the result of a decisive identification with one or another first-order desire. Because decision is an action of the agent, the agent is not passive with regard to it. An agent decides and the result of that decision is an intention. Intentions are forward-looking and involve a commitment to suppress some motives in favor of others – a kind of anticipatory configuration of the will. Second-order desires, understood as intention-establishing decisions, don't need higher-order endorsement because the agent is not passive with regard to them. If the agent is not passive with regard to them, they cannot be external.

Why does Frankfurt ultimately drop decision as a means of explicating identification and thereby distinguishing true from merely effective desires? I think it might be that he began to see that if a decision is a decision only if it's wholehearted (second quote above), if you count as having decided only if your decision 'takes,' then it might look as if decision is more of a state than an action and

⁵⁴ Although Watson would *at least* insist on scare quotes around "desire," since he thinks a notion of evaluation is needed here, and he rejects the idea that evaluation can be reduced to desiring.

not so much up to us after all.⁵⁵ Furthermore, as Frankfurt replies to Scanlon, “there is... a fundamental and ineradicable error in the very attempt to explicate being active in terms of endorsement or in terms of any other activity. Such attempts are manifestly bound to be circular.”⁵⁶

In any case, Frankfurt comes up with what he thinks is a better solution.

IDENTIFICATION AS WHOLEHEARTEDNESS

The better solution runs as follows. Although first-order desires may need to be ratified (or minimally, “accepted”) higher up in order for action they effect to be autonomous, second-order desires do not need to be the object of higher-order desires (which leads to regress). Nor even do they need to be arrived at actively (decision). Rather, as a set, they need to be internally coherent such that when any first-order desire pops up there are not conflicting higher-order attitudes toward it such that the person “(a) cannot act decisively; or (b) finds that fulfilling either of his conflicting desires is substantially unsatisfying.”⁵⁷ This degree of internal coherence facilitates satisfying action he sees as expressing “wholeheartedness.”

Wholeheartedness refers to a unity of higher-order attitudes and inclinations.⁵⁸

This unity has nothing to do with passion: “what is at issue is the organization of the will, not its temperature.”⁵⁹ This unity is, however, compatible with “virulent psychic

⁵⁵ Bratman, *Faces of Intention: Selected Essays on Intention and Agency*, Bratman, “Identification, Decision, and Treating as a Reason.”

⁵⁶ Frankfurt, “Reply to T. M. Scanlon,” p. 220.

⁵⁷ Frankfurt, “The Faintest Passion,” p. 99.

⁵⁸ *Ibid.*, p. 100.

⁵⁹ *Ibid.*

conflict"⁶⁰ (just so long as it does not take place among higher-order desires).

Frankfurt says of the unwilling addict that the “addiction may defeat his will, but does not disrupt its unity.”⁶¹

Wholeheartedness is also “knowing” what one wants. But ambivalence – conflict among higher-order desires – cannot be overcome by getting additional information. It is not a “cognitive deficiency.” It may be that most ambivalence occurs in cases in which cognitive uncertainty and ambiguity is actually a factor. In other words, in order to know what one wants, one must also know what it costs and signifies and that is often impossible to assess with certainty. Frankfurt is saying that even if one has all the facts, one still might be unable to decide.

Wholeheartedness is “tantamount to the enjoyment of a kind of self-satisfaction.”⁶² Satisfaction is defined negatively as having no interest in making changes in such psychic elements as beliefs, intentions, feelings, or desires. It does not require a conscious assessment that one is satisfied. Such an assessment is neither necessary nor sufficient for satisfaction. “Satisfaction is a state of the entire psychic system – a state constituted just by the absence of any tendency or inclination to alter its condition.”⁶³ Frankfurt offers this as the solution to the regress problem: the negative definition is “important,” he says, “because it explains why there is no danger here of problematic regress.”⁶⁴

⁶⁰ Ibid. I have been questioning this.

⁶¹ Ibid., p. 99.

⁶² Ibid., p. 102. It is also, to some extent “tantamount” to loving, p. 106.

⁶³ Ibid., p. 104.

⁶⁴ Ibid.

Wholeheartedness, then, is equivalent to a kind of satisfaction, which Frankfurt defines negatively. Wholeheartedness, too, is defined negatively – as absence of ambivalence.

Here what is divided is neither a person's reason nor his affects, but his will. Insofar as someone is ambivalent, he is moved by incompatible preferences or attitudes regarding his affects or his desires or regarding other elements of his psychic life. This volitional division keeps him from settling upon or from tolerating any coherent affective or motivational identity. It means that he does not know what he really wants.⁶⁵

In a footnote, Frankfurt adds (as noted above): “there are degrees of the sort of conflict I am considering. In discussing ambivalence, I am concerned with conflict sufficiently severe that a person: (a) cannot act decisively; or (b) finds that fulfilling either of his conflicting desires is substantially unsatisfying.”⁶⁶

Frankfurt thus finally stops worrying about understanding “true” desires in terms of links with particular authorizing attitudes or actions. He decides that “satisfaction” can be authoritative and that satisfaction with our desires is the default case. Because our desires can conflict and because in creatures like us, such conflict typically generates reflective awareness and higher-order sorting out, unless a desire gets a bad review, it counts as one with which we are satisfied.⁶⁷ Another way to put this is to say that if on a given occasion one has doubts or has conflicting desires (to x, or not to x), the matter gets kicked up to volition (which desire we will officially accept as the one we want to move us in action). If the first-order conflict cannot be

⁶⁵ Ibid., p. 99.

⁶⁶ Ibid.

⁶⁷ Are we satisfied with desires of which we are unaware? Frankfurt seems to say ‘yes’ and ‘no’...yes, we can be satisfied and identified with unconscious and unreflective desires, but no, such desires, even though we might be identified with them, are wanton. In other words, ‘satisfaction’ is sufficient to authorize higher order desires, but not first-order desires.

resolved because of ambivalence at this higher level, we have not satisfaction but “volitional” ambivalence, which is one important threat to autonomy.⁶⁸

PROBLEMS WITH IDENTIFICATION AS WHOLEHEARTEDNESS

In what follows I’ll try to show in more detail why Frankfurt’s wholeheartedness solution is a departure from his earlier “hierarchical” and “decision” ways of construing identification. I will also discuss two basic problems with this solution. One problem is that as he develops his wholeheartedness solution, the concept of identification shifts, and as it shifts it becomes broader, so much so that identification itself no longer functions to mark the distinction between “true” desires and merely

⁶⁸ Early on, Frankfurt notes that a person can attempt to resolve ambivalence by interpreting a conflict between desires as contingent, and prioritizing the desires. The example Frankfurt uses is a conflict between a desire for work and a desire for leisure. He suggests that this is a case in which the conflict between the desires is not inherent in the desires themselves. It is intelligible to have both desires, and both desires can be satisfied sequentially; it’s just a matter of prioritization.

Frankfurt thinks some conflicts are deeper than this, and of a different nature. He calls them inherent conflicts. The example he gives is of wanting to both congratulate and insult one’s friend upon that friend’s winning an award. In that case, one makes a decision to exclude one of these incompatible desires, and in virtue of such a decision (assuming it takes), the incompatible desire is now “external” to the self in the same sense that an unintended bodily movement may be external to one’s history, Frankfurt wants to say.

Still, volitional ambivalence cannot simply be willed away or “overcome voluntaristically” Frankfurt later recognizes. Conflict may not be resolved, ambivalence may remain. Frankfurt claims that the absence of wholeheartedness with regard to competing desires is the absence of any true desire, suggesting that if we do manage to act in spite of ambivalence (and that action does not itself serve to resolve the ambivalence) the act was wanton. This interpretation might seem apt in the case in which someone ambivalently takes up a commitment to someone, a commitment in which wholeheartedness is presumed and required, and a commitment which his ambivalence rapidly undermines.

On the other hand, although satisfaction with our desires may be the default, that does not mean that we are typically satisfied with them. I think Frankfurt might agree that ambivalence is our lot in life, that wholeheartedness is rare. One way to achieve it might be to take on a commitment and grow into it. The possibility of being able to grow into a commitment is central to the self-invention perspective from which Frankfurt begins in his first article on this topic, where he takes freedom of the will as “having the will one wants.” Following his “satisfaction” solution, however, Frankfurt’s primary emphasis shifts to acceptance and, as we shall see below, to our existential dependence on having things we can’t help caring about, which he discusses as “volitional necessity” and “love.” In the end, Frankfurt stresses discovery over invention and building commitments from the “outside” in rather than inside out – at least in the sense that he recommends making commitments not on the basis of what we want to care about, but on what a clear-eyed assessment shows us we *can* care about. Such an assessment requires us to supplement immediate introspection with a third person view of our behavior.

effective ones. The second problem is that neither of Frankfurt's alternatives, namely, wholeheartedness considered in itself, or "caring" clearly functions to mark out the distinction, either. In the next chapter, III, I'll try to show how the concept of identification shifts with the introduction of wholeheartedness, the difficulties that creates, which I'll illustrate in a discussion of the case of the unwilling addict, and how these difficulties might be, but ultimately aren't solved by "caring." In the following chapter, IV, I'll attempt to connect those difficulties to an ambiguity about standards of internality. I'll suggest that Frankfurt has more than one standard in play and that "identification" is mixed up with both of them.

III. EXPLICIT TO IMPLICIT: IDENTIFICATION AND CARING

I believe that in the course of modifying his hierarchical account to stop a regress problem, Frankfurt actually loses the ability to differentiate merely effective desires from “true” desires – or free action from free willing.

This worry is inspired by noticing that as Frankfurt moves from an emphasis on hierarchy to an emphasis on wholeheartedness, he also moves from explicit to implicit interpretations of “identification” and, later (to be discussed below), “caring.” Identification and, later, caring are what serve to authenticate and thus authorize our desires. The problem with implicit interpretations of these authorizing functions is that they are read off behavior, much as “strongest” desires or “effective” desires are, and I suspect that that might lead us back to where we started – unable to account for volitional struggle or the sense in which we can lose battles for volitional control, and unable to distinguish our effective desires from our “true” desires.

Here’s how I think the move from explicit to implicit interpretations of “identification” goes. In the case of identification, it looks as if Frankfurt moves from an “invention” account wherein identification is an explicit (although mysterious) act of ratification, to a “discovery” account, wherein (in spite of the wholeheartedness criterion) identification need be no more than resigned acceptance of one’s effective desires; furthermore, one may be identified with desires of which one is not even aware.

It is not clear that this diminished sense of identification captures the sense in which we are “active” that is at the “heart” of our existence as selves. Because this diminished sense of identification is so inclusive, it doesn’t seem to help us much

with the effective desire/"true" desire distinction. It certainly does not distinguish between desires that we regard as important to us and those we don't.

To see how the shift works, I think it might help to view it in terms of the two main bases of attitude attribution: first person, or "inside out," and third person or "outside in." As far as I can tell, there are the two basic approaches to "identification" that correspond to these two main bases of attitude attribution, inside out and outside in. Frankfurt starts inside out and he moves outside in.

Inside out. One approach to identification, "inside-out," attempts to ground an account of which desires are an agent's own in terms of a first-person act or attitude of identification. This is the approach Frankfurt takes in his hierarchical and decision accounts. For example, in an early essay, Frankfurt says the following:

Now a person is active with respect to his own desires when he identifies himself with them, and he is active with respect to what he does when what he does is the outcome of his identification of himself with the desire that moves him in doing it. Without such identification the person is a passive bystander to his desires and to what he does, regardless of whether the causes of his desires and of what he does are the work of another agent or of impersonal external forces or of processes internal to his own body.⁶⁹

Here it sounds as if identification itself is a kind of action. When one "identifies" with one's desires, "they are not merely desires he happens to have or to find within himself, but desires that he adopts or puts himself behind."⁷⁰ Now this

⁶⁹ Frankfurt, "Three Concepts of Free Action," p. 54.

⁷⁰ *Ibid.*, p. 53.

to the extent that a person identifies himself with the springs of his actions, he takes responsibility for those actions and acquires moral responsibility for them; moreover, the questions of how the actions and his identifications with their springs are caused are irrelevant to the questions of whether he performs the actions freely or is morally responsible for performing them.

Below Frankfurt explains how Locke's Devil/Neurologist may supply an entity with stable character program, but that the implanted character, while it would extend some responsibility to the Devil/Neurologist, would not mitigate the responsibility of the programmed entity. It's worth noting

kind of identifying action of “adopting” or “putting oneself behind,” or, later “deciding” is pretty assertive in the old-fashioned existential sense:⁷¹

When the decision is made without reservation, the commitment it entails is decisive...The decision determines what the person really wants by making the desire on which he decides fully his own. To this extent the person, in making a decision by which he identifies with a desire, constitutes himself...It is these acts of ordering and rejection—integration and separation—that create a self out of the raw materials of inner life.⁷²

But there is a sense of first person “identification” that is also more passive, which Frankfurt emphasizes in later work. Beginning with “Faintest Passion,” he seems to imply that this more passive sense is sufficient to fulfill the authenticating function of identification.⁷³ He characterizes this identification not in terms of active decision, but in terms of “satisfaction” where satisfaction with a (higher-order) desire (a desire to be) is simply a matter of doubts and conflicts that do not arise (“neutral attitude of acceptance”). Insofar as identification is a conscious phenomenon (which is a conceptual truth on an inside-out account), this more passive identification (wherein I am both conscious and accepting of motives as mine) is less a matter of

that this adopting or putting-oneself-behind “action” can be set up by a Lockian Devil/Neurologist if he:

provides his subject with a stable character or program, which he does not thereafter alter...and that the subsequent mental and physical response of the subject to his external and internal environments are determined by this program rather than by further intervention on the part of the D/n. [The subject] may become morally responsible, assuming that he is suitably programmed, in the same way that others do: by identifying himself with some of his second-order desires...

⁷¹ Inside out illustration: “When you deliberate, it is as if there were something over and above all of your desires, something which is *you*, and which *chooses* which desire to act on” Korsgaard, *The Sources of Normativity*, p. 100. Frankfurt’s account is parallel to Korsgaard’s, but, unlike Korsgaard, he doesn’t see the gap between reflective awareness and action as needing to be bridged by reason.

⁷² Frankfurt, “Identification and Wholeheartedness,” p. 170, Frankfurt, *The Importance of What We Care About: Philosophical Essays*, p. 170.

⁷³ But the active sense shows up as recently as Frankfurt, “Reply to T. M. Scanlon.” in *Contours of Agency*.

assertion and self-invention and more a matter of recognition and discovery. For instance, a person who grows up among bold adventurers, who admires their deeds but lacks their temperament, may finally come to accept or identify with his own overpowering desires for safety. He might do this in a mood of self-contempt, castigating himself for his cowardice, or he might begin to see himself as prudent. In either case, by recognizing desires for safety as his, he identifies with them.

Outside in. In “Faintest Passion” Frankfurt also starts to shift from explaining what it is “to identify” to speaking in terms of “being identified with.” He switches emphasis from something like an act of identification to something closer to an account of what it is to be identified. He looks at the self as a system of psychic elements that must function together in a coordinated fashion.

Here’s how Frankfurt originally sliced things up: “First-order desires” are essentially impulses to action. Formally, they are desires to do *x*. “Second-order desires” are essentially self-assertions. Formally, they are desires to *be* [moved by one’s desire to do *x*].⁷⁴ “Wholeheartedness,” introduced in “Faintest Passion,” refers both to a motive and to a practical constraint on our self-assertion. While desires to *do* may be reined in by desires to *be*, desires to *be* are limited by a wholeheartedness requirement (what we can consistently will, where that is understood empirically rather than theoretically).

On Frankfurt’s account in “Faintest Passion,” one is identified with desires with which one is satisfied or upon which one wholeheartedly acts. One can be identified with desires with which one is satisfied whether one recognizes them or

⁷⁴ Frankfurt, “Freedom of the Will and the Concept of a Person,” pp. 13-14.

not. The desires one identifies with, i.e., the desires on which one acts with self-conscious accepting awareness, will be a subset of those with which one is identified. These desires on which one wholeheartedly acts are the “true” desires.

The original sense of identification was the self-assertion sense illustrated in quotes above, wherein (it now looks) from a desire to do, one reflexively generated a desire to be that either passed or failed the wholeheartedness test. The new, more passive, sense of identification, then, might be described as the recognition of which desires pass the test. Passing the test of wholehearted willing is what qualifies desires as authentically one’s own. Identification is the recognition of these desires. If this is the case, it looks as if identification is not itself doing the authenticating. In other words, it is not identification that is doing the authenticating but the fact that one’s desires form something like a coherent set.

Recent remarks of Frankfurt seem to leave room for such a reading:

four elements or stages are conceptually distinguishable within the process, even though they may not be either separated or separable in fact. The agent (i) provisionally suspends or brackets his relationship to D [“The fact that the agent is reflecting on D entails that he has already put his relationship to it in question”] (ii) reflects on D, (iii) forms a higher-order desire or attitude toward D, and (iv) identifies himself with D or alienates himself from it.⁷⁵

We could read (iii) “forms a higher-order desire” as the original, self-assertion sense of identification and (iv) “identifies or alienates” as the new, self-recognition sense.

Often, assertion and recognition will occur together. This is what happens in paradigm cases of identification, and these are cases that Frankfurt originally had in mind. Consider being moved to a New Year’s resolution by the desire to stop smoking. When such a higher-order desire fails the wholehearted willing test, when

⁷⁵ Frankfurt, “Reply to Michael E. Bratman,” pp. 86-7.

it becomes apparent that this desire is not a true commitment, assertion and recognition split, the assertion fails to hold, revealing that an act of assertion is not sufficient to authenticate one's motive.

In his early examples Frankfurt may have had in mind that "assertion" characterizes only the "success" cases, the cases in which the reflexively generated attitude is at least robust (wholehearted) and probably also effective (a volition). In the success cases, in the cases in which my New Year's resolution coheres with my other higher-order desires, we recognize that in making our resolution, what we have done is to have identified ourselves with a desire. This is the paradigm case. In the incomplete cases, the cases in which our resolution never had a chance, what we have is a reflexive mechanism that has failed to tip the balance, to get behind a desire we can wholeheartedly act upon. Identification is now the recognition, the cognizance, the discovery,⁷⁶ of what one can wholeheartedly will. "We can only be what nature and life make us, and that is not so readily up to us."⁷⁷

WHOLEHEARTEDNESS AND THE UNWILLING ADDICT

While there is some truth captured in Frankfurt's new understanding of identification, the shift also creates some difficulties, which I think can be seen if we look closely at the case of the unwilling addict.

Frankfurt discusses three addicts: *wanton*, unwilling, and willing. He distinguishes between the first two – the unwilling addict, who opposes his addiction, and the wanton addict, who is heedless of it – as a way of driving home the point that

⁷⁶ Ex post facto. And this can be faulty. What Frankfurt says: we can "decide" to go with one alternative, but it may not take. (Frankfurt, "The Faintest Passion," p. 101.)

⁷⁷ Ibid.

having second-order desires, caring about which desires move us in action is “essential to being a person.”

“Persons” are not wantons. A wanton is a creature that does not self-consciously “identify.” A wanton lacks capacity for or utterly fails to generate second-order desires (but nevertheless “is identified” with whatever desires/motives/attitudes we can attribute to him on the basis of his behavior). Persons do generate second-order desires. This allows conflicting desires occurring simultaneously to be resolved *personally* (rather than mechanistically, as on the Hobbesian model). Hence, it makes sense to ask which of a person’s attributable motives represent his “true” or authentic desires. When persons act on such motives they enjoy not only *freedom of action*, “doing what one wants to do,”⁷⁸ but also *freedom of the will*, wherein a person has “the will he wants.”⁷⁹ Wantons are capable of freedom of action, but not freedom of will. A person enjoys freedom of will when his effective desire and higher-order desires are in alignment.⁸⁰ In such a case, the person acts on what I’ve been referring to as a “true” motive. What the third addict – the *willing* addict, the addict who endorses his addiction – illustrates is freedom of will, or *autonomy*. Even when, because of his addiction, he is not free to have whatever will he might want, by identifying with his addictive desire, a person may nevertheless act freely and act of “his own free will.”

⁷⁸ Frankfurt, “Freedom of the Will and the Concept of a Person,” p. 20.

⁷⁹ *Ibid.*, p. 24.

⁸⁰ Originally Frankfurt claimed “And it is in the discrepancy between his will and his second order volitions, *or his awareness that their coincidence is not his own doing but only a happy chance*, that a person who does not have this freedom feels its lack” (my emphasis, *Ibid.*, p. 20.). Now it is all about the configuration.

Frankfurt can, of course, distinguish between the willing addict, who identifies with his addiction, and the wanton addict, who simply doesn't care. The wanton addict lacks second-order desires, hence doesn't identify at all. The limitation of this contrast, though, is that since a wanton is not a person anyway, you might say that the contrast between the wanton and either the willing or the unwilling addict doesn't tell us what happens when a *person's* autonomy breaks down. For that, perhaps, we need to look at the contrast between the willing and unwilling addicts, both of whom are persons.

In particular, Frankfurt needs to be able to distinguish between the willing addict shooting up, willing freely, and the unwilling addict shooting up, acting freely. That is, the unwilling addict shooting up, it might be said, is still technically "acting freely," acting on what we might *ordinarily* say is his strongest desire, even if it is not what he "most wants." What Frankfurt says, though, is that the unwilling addict is "helplessly violated by his own desires," a "passive bystander" to his addiction – technically, it is not he who shoots up.

I think this is too strong. It implies that a *person* either acts freely *and* autonomously, or doesn't act at all – suffers some sort of behavioral seizure. Furthermore, Frankfurt does not uniformly take this hard line. He says, for instance, that "When a *person* acts, the desire by which he is moved is either the will he wants or a will he wants to be without,"⁸¹ suggesting there is an action out there to be discussed and accounted for. If the willing addict wills freely as he shoots up, as Frankfurt claims, and the unwilling addict at least acts freely as he shoots up, as I

⁸¹ Frankfurt, "Freedom of the Will and the Concept of a Person," p. 19.

have suggested, then, in order for Frankfurt to distinguish among kinds of freedom involved in different cases of acting on an addictive desire, he needs to be able to distinguish the case of the willing addict shooting up, willing freely, and the unwilling addict shooting up, acting freely.⁸²

The contrast between the willing and unwilling addict is shooting up is stark if you grant Frankfurt that, because the unwilling addict fails to identify with his addictive desire, it's literally not the addict who's shooting up. But wholeheartedness (or, technically, "satisfaction") is the new criterion of identification. How does identification as wholeheartedness work to distinguish these two addicts shooting up?⁸³

Let's grant for now that acting wholeheartedly entails acting not on a desire with which I am only implicitly satisfied – a desire with which I am merely identified

⁸² If, however, acting on a bare first-order desire is wanton, I think that acting akratically on an addictive desire is arguably a kind of wantonness, because in the case of the unwilling addict, the addictive desire, while it is the object of a higher order attitude, lacks higher order endorsement. Furthermore, I think that the unwilling addict shooting up is better understood as a kind of complex ambivalence, and if ambivalence can result in wantonness (on Frankfurt's "no volitions of the second order" criterion), which I believe he does allow, then it would make sense to read the unwilling addict's shooting up as a kind of wantonness. Howard Rachlin, "Complex Ambivalence," in *The Science of Self Control* (Cambridge: Harvard University Press, 2000). Rachlin develops his conception of complex ambivalence to describe phenomena like addiction, wherein, for any time-slice, t, taking a drink is more rewarding than staying sober, but over T, staying sober is more rewarding.

⁸³ Now there may be some ambiguity about whether acting wholeheartedly is different from acting on a desire with which one is satisfied. Acting on a desire with which one is satisfied doesn't actually entail that one explicitly reflects upon or approves the desire upon which one acts, it just means that one harbors no psychic elements that clash with/disapprove of the desire. This could be a problem in the sense that someone might object that acting on a desire with which one is implicitly satisfied is not sufficient for autonomy. It certainly clashes with traditional rationalistic accounts. In the spirit of retaining this rationalistic intuition, Frankfurt at times gives a seemingly more explicit interpretation of satisfaction as being the satisfaction of a self-reflective entity and therefore apparently involving something like an explicit assessment of being, overall, satisfied. However, he states that such an assessment is neither necessary nor sufficient for actually being satisfied. Furthermore, such an overall assessment still allows the possibility of acting on desires with which one is still only implicitly satisfied. This is where "wholeheartedness" comes in. Frankfurt defines wholeheartedness in terms of satisfaction ("tantamount to a certain self-satisfaction"), that is, to be satisfied, overall, that this rather than that desire is effective... To act wholeheartedly upon a desire therefore paradigmatically (though perhaps not necessarily) involves some explicit awareness, assessment, and approval of the desire in question.

– but on a desire with which I am explicitly satisfied, on which I act wholeheartedly, with which I *identify* On Frankfurt’s original view, the simple hierarchical account, identification is sufficient for authentication, for free *willing*, for volitional acts. However, on this new, satisfaction account, “identification” doesn’t work quite the same way. He makes this quite explicit in a recent reply to Gary Watson.

In the first place, Frankfurt insists that mere “resigned acceptance” suffices for identification.

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. What I have had in mind when I have employed the notion of endorsement is something that makes no claim or judgment whatever, and that is more accurately specified as the altogether neutral attitude of acceptance. A person may be led to accept something about himself in resignation, as well as in approval or in recognition of its merit. The fact that he accepts it entails nothing, in other words, concerning what he thinks of it.⁸⁴

In the second place, Frankfurt goes even further, claiming that on his view, “I may identify with desires of which I do not approve.”⁸⁵ Finally, he claims that I may identify with desires I do not consider important: “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.”⁸⁶

With these three admissions, Frankfurt is conceding that identification, even if interpreted explicitly, does not distinguish between attitudes that are characteristic of me and attitudes that are important to me. What does this mean? Attitudes that are characteristic of me are attitudes that might be attributed to me by others, based on

⁸⁴ Frankfurt, "Reply to Gary Watson," p. 160.

⁸⁵ *Ibid.*, p. 161.

⁸⁶ *Ibid.*

my behavior. They are patterns of reaction that have proved useful for predicting my behavior. They may also be attitudes I attribute to myself, ruefully admitting, say, that I am a drug addict. They may be desires I attribute to myself in order to make sense of my past behavior and to help me decide what to do next. For example, I may turn down an invitation if it entails something I know I won't enjoy, or will trigger my addictive desires.

Admitting that a desire is characteristic of me in this sense is something Frankfurt now equates with identification, but if this is the case, then what? Then, as I've said before, identification cannot perform the authenticating function. That is because identification, even if we restrict it to first person, explicit, conscious attitude attribution, fails to distinguish attitudes that are characteristic of me from attitudes that are important to me. Frankfurt needs this distinction in order to distinguish free action from free willing, the unwilling addict from the willing addict.

Here's why. If identification is a matter of accepting attitudes that are characteristic of me, regardless of whether I approve of them or whether I regard them as important, then, for one thing, paradoxically, it looks as if identification actually is a necessary condition of the "externalization" on which the unwilling addict's very status as unwilling depends. It seems "identification" is the first step of 12-step programs, which go on, I believe, to then "externalize" the addiction as a disease or something the person does not have power over. If you think that 12-step programs are effective in ultimately freeing people from their addictions, and that they result in greater personal autonomy, then explicit, conscious identification, in the resigned acceptance sense, of the desire you're hoping to rid yourself of is the

essential first step. Frankfurt's current interpretation of identification (resigned acceptance) is consistent with this. The problem is that it doesn't help us to understand, in the case of the unwilling addict, why my desire to overcome my addiction is more important than my desire to placate it. This broader concept of identification is no longer enough to mark the distinction. In this broader conception, by rueful admission, I identify with my addiction, yet I also identify with my desire to overcome it. The fact that I can be identified with both attitudes, even though they conflict, is entailed by Frankfurt's insistence that I may identify with a desire I disapprove of, insofar as disapproval consists of having an opposing desire with which I am also identified, which seems to be the case on Frankfurt's view.

As a willing addict, I act on a desire with which I also identify, my desire to ingest the drug. Minimally, I admit that I have this desire and that it's characteristic of me.⁸⁷ As an unwilling addict, I also act on a desire with which I identify, the very same desire, my desire to ingest the drug. (Although, as I said earlier, Frankfurt originally denies that in this case I am acting at all. He claims I am a "passive bystander," but that seems too strong.) When I'm an unwilling addict, there is another desire I'm identifying with, which is the desire to overcome my addiction, to act on any of a set of desires other than the desire to ingest the drug. While it's true that this desire marks a difference between these two cases, the difference can no

⁸⁷ Actually, when Frankfurt introduces the willing addict, he emphasizes that the willing addict not only identifies with his addiction, but cares about it – he would work to maintain it should it threaten to subside. Frankfurt seems to need to endow identification with the properties he later identifies as belonging to concern because he says that "A person's will is only free if he is free to have the will he wants. And it is in the discrepancy between his will and his second order volitions, or his awareness that their coincidence is not his own doing but only a happy chance, that a person who does not have this freedom feels its lack." Frankfurt, "Freedom of the Will and the Concept of a Person," p. 20. In other words, unless a person's fixed effective desire is overdetermined, that is, unless a higher order desire would "step into the breach" to ensure that this motive stays effective, then the person acting on this fixed motive, even if she identifies with it, might not be acting of her own free will.

longer be made out in terms of identification. The difference might be made out in precisely in terms of how the willing and unwilling addicts are identified with their addictive desires, but that is another question.

You might think that Frankfurt could make the distinction between these two addicts shooting up directly in terms of wholeheartedness. Intuitively, it makes sense. The willing addict is wholehearted about his desire to ingest the drug; the unwilling addict, though he identifies in the resigned acceptance sense with his desire to ingest the drug, also harbors a “higher-order” desire to overcome his addiction as well as a first-order desire not to take the drug.⁸⁸ He is identified, in turn, with this higher-order desire by virtue of the fact that he is implicitly satisfied with it. He is identified then with both the addictive desire and the desire to overcome the addiction. He also identifies with both of these desires (presuming we can translate a desire to overcome an addictive desire to a first-order desire). Insofar as these desires conflict, it seems they are not his “wholehearted” desires. So one way to distinguish the willing and unwilling addict might be to say that when the willing addict shoots up, he does so wholeheartedly, and when the unwilling addict shoots up, he does so ambivalently. (In spite of Frankfurt’s claim in “Faintest Passion” that “The addiction may defeat his will, but it does not disrupt its unity.”⁸⁹)

The possibility that the willing addict shoots up wholeheartedly and the unwilling addict shoots up ambivalently does seem to mark a distinction between the

⁸⁸ This is another niggling thing about Frankfurt’s account. He would say that the addict has a first order desire not to take the drug, but I’m not sure that’s accurate. First order desires ‘not to x’ where x is some powerful biological urge seem contrived. It is not as though there is an equal and opposite urge to ‘not x’. To the extent there’s a desire to ‘not x,’ it seems higher order. More likely, in each instance where the problem urge flares up there’s a higher order desire that I act on one of my other available desires in that situation.

⁸⁹ Frankfurt, "The Faintest Passion," p. 99.

two cases, but there is still a problem. The problem is that on Frankfurt's satisfaction account of identification, an account he's developed to stop the regress of ever higher order desires necessary to authenticate an attitude as one's own, wholeheartedness is a criterion of identification. To identify or be identified is just to be wholehearted. But I've just noted (quotations above) that Frankfurt *also* says that identification need be no more than resigned acceptance and that we can identify with desires of which we disapprove.

These somewhat deflated identifications could count as forms of wholeheartedness on Frankfurt's view. After all, Frankfurt distinguishes "struggle" between first-order desires from higher-order "ambivalence" and claims that "struggle" is perfectly compatible with wholeheartedness. An example of struggle just is the struggle of overcoming an addiction wherein Frankfurt treats the desire to overcome the addiction as a wholehearted higher-order attitude toward a first-order desire not to take the drug, which is in competition with another first-order desire to take the drug. Frankfurt wants it to be the case that conflict between a first order desire and higher-order desire does not threaten satisfaction or wholeheartedness, which applies mainly to the set of higher-order desires which take first-order desires as their objects. On the other hand, as noted earlier, Frankfurt sometimes suggests that the distinction between "struggle" and "ambivalence" is less than clear:

it may be that energies tending toward action inconsistent with the intention remain untamed and undispersed however decisively the person believes his mind has been made up...In that case the decision, no matter how apparently conscientious and sincere, is not wholehearted⁹⁰

⁹⁰ Frankfurt, "Identification and Wholeheartedness," p. 174.

I believe a case can be made that the distinction between first-order and higher-order desires is also less than clear (see “Two standards of internality” below).

The point is that Frankfurt allows that wholeheartedness is something less than complete psychic unity. Certainly if Frankfurt is understanding identification as a kind of wholeheartedness, and he wants to allow that one can identify with desires of which one does not approve, wholeheartedness must be something less than complete psychic unity (whatever that could be). This seems plausible. He may also want to allow that wholeheartedness, and therefore identification, is a matter of degree. This also seems plausible.

One consequence, however, is that the distinction between the willing addict shooting up and the unwilling addict shooting up is not a matter of whether either addict is identified with his addictive desires. In the new, deflationary sense of identification, both are. The distinction between them becomes the degree to which they are identified.

In insisting that resigned acceptance counts as identification, in allowing that “I may identify with desires of which I do not approve,” and since I can identify with a desire without regarding it as important,⁹¹ Frankfurt is conceding that the concept of identification, even if interpreted explicitly, does not absolutely or by itself distinguish between attitudes I acknowledge may be characteristic of me and attitudes that are important to me.

It’s important to notice that acting on attitudes that are characteristic does not by itself distinguish persons from wantons. People act on characteristic desires, but

⁹¹ Harry G. Frankfurt, “On Caring,” in *Necessity, Volition, and Love* (1999), pp. 160-1.

so do wantons. Frankfurt has defined wantons as entities that don't care (in some cases because they are unaware) about which of their desires are effective in action. Wantons simply act on, or are driven by, their desires. Persons are sometimes (many times) driven in the same way, unreflectively. The idea that I might additionally recognize and identify with such a desire by means of resigned acceptance doesn't seem to help. Identification here does not indicate that I care about which desire moves me or that I regard any particular desire as important.

In order to distinguish between persons and wantons Frankfurt needs to be able to distinguish between desires that are merely characteristic of an entity and those that are important to it. Stipulating that the entity acting on a characteristic desire has the complexity of a person doesn't yet explain how that complexity bears on the desire in question. In fact, in spite of our intuitive sense that wholeheartedness entails something like explicit identification, it's possible to argue that, on Frankfurt's view, episodes of wholehearted willing in persons need not involve explicit identification. What I have in mind here is that just as he shows we attribute desires to wantons and consider them "identified" with those desires (conscious awareness and/or reflection not being required here) so too we tend to attribute such desires to persons. More to the point, we also tend to attribute unifying higher-order desires to person, e.g. an organizing desire for control, achievement, love, without requiring that those higher-order desires be conscious or explicit.

Perhaps the real bottom line, though, is that Frankfurt himself does not rest with his "satisfaction" solution or his "wholeheartedness" fix as you might expect if he felt they were sufficient for expressing his original intuition about the importance

of volition. He feels compelled to introduce a caring supplement. But as I aim to show below, caring, too, can collapse into effective desire.

CARING

Toward the end of “Faintest Passion” Frankfurt revisits his original assertion, introduced in “Freedom of Will and the Concept of a Person,” that “a natural and useful” conception of free will is that “a person’s will is free to the extent that he has whatever will he wants.” He explains that if this – if having whatever will one wants – refers to direct, voluntaristic control – free will is unreal. “The concept of reality is fundamentally a concept of something which is independent of our wishes and by which we are therefore constrained...A person’s will is real only if its character is not absolutely up to him.”⁹² A better interpretation, he suggests, requires not that “he originate or control what he wills, but that he be wholehearted in it. If there is no division within a person’s will, it follows that he has the will he wants.”⁹³

Frankfurt also wants certain kinds of wanting to carry authority. He holds that a “person acts autonomously only when his volitions derive from the essential character of his will,”⁹⁴ but “what autonomy requires is not that the essential nature of the will be a priori [contra Kant], but that the imperatives deriving from it carry genuine *authority*.” To articulate this idea of authority Frankfurt makes a distinction between not just orders (as in hierarchies, even hierarchies with cohesive upper

⁹² Frankfurt, “The Faintest Passion,” p. 101. See also “Choosing agent,” below.

⁹³ Ibid.

⁹⁴ Frankfurt, *Necessity, Volition, and Love*, p. 132.

levels), but *kinds* of wanting.⁹⁵ He introduces the idea of caring as a kind of wanting. He asserts that caring is “substantially equivalent” to “regards it as important.”⁹⁶ “[P]eople necessarily consider whatever they care about to be important to them; and, conversely, they necessarily care about anything that they consider important to them.”⁹⁷ Something is important, in turn, if it is “needed,” and it is needed if its lack causes “harm.”

I am not aware of Frankfurt addressing the following point explicitly, but when considering wholeheartedness with regard to particular objects, actions, or motives, it seems that the concept of wholeheartedness cannot by itself express the idea of importance. While wholeheartedness paradigmatically is also a manifestation of concern, hence, importance, if anything, it may be easier to be wholehearted in our attitudes and willing about trivial matters than important ones – as in wholeheartedly preferring grapefruit for breakfast.

Wholeheartedness enabled Frankfurt to expand the scope of his discussion from hierarchical relations among single desires to coherence among desires considered as a set. A coherent set is “tantamount” (as Frankfurt likes to say) to personal unity. Caring, now, enables Frankfurt to extend this unity through time.

⁹⁵ Representations of what people want are thought to have justificatory power and can certainly be rhetorically effective; therefore a good account of government’s role and demands of morality needs a more nuanced account of wanting. Liberalism in particular aims to maximize the space for people to get what they want, justifying it with the idea that this leads to happiness. Because the connection between wants, preferences, and happiness is notoriously unreliable, unless we can draw distinctions among kinds of “wanting” by recourse to something like a concept of caring, we won’t be able to account for the relationship that that’s supposed to hold between getting wants or preferences satisfied and happiness. Happiness does not equal getting what you want or even getting what you prefer. (This is a close paraphrase of pp. 156-157, Frankfurt, “On Caring.”) Frankfurt also explicitly states that he means to use “desire” and “want” interchangeably (Frankfurt, “Freedom of the Will and the Concept of a Person.”)

⁹⁶ Frankfurt, “On Caring,” p. 155.

⁹⁷ *Ibid.*, p. 156.

Although Frankfurt's definition of importance in terms of need might at first get us thinking about urgent cases – thinking synchronically – what caring really enables Frankfurt to do is to expand from synchronic to diachronic unity. What he says is:

Suppose we cared about nothing. In that case, we would be creatures with no active interest in establishing or sustaining any thematic continuity in our volitional lives...Of course, we would still be moved to *satisfy* our desires...our capacity for higher-order desires and higher-order volitions might remain fully intact. Moreover, some of our higher-order desires and volitions might tend to endure and thus to provide a degree of volitional consistency or stability in our lives. From our point of view as agents, however, whatever coherence or unity might happen to come about in this way would be merely fortuitous and inadvertent. It would not be the result of any deliberate or guiding intent on our part.⁹⁸

Being “committed” to a desire, then, involves a “disposition to be active in seeing to it that the desire is not abandoned or neglected,”⁹⁹ and having such a disposition is distinct from having higher-order desires or volitions.¹⁰⁰ Finally, by emphasizing that caring has a certain existential importance to us, apart from what it is we care about, Frankfurt implicitly commits to the idea that persons are essentially diachronic entities. Caring is the “foundational activity through which we provide continuity and coherence to our volitional lives.”¹⁰¹

Frankfurt seeks to distinguish caring from desiring something badly on the one hand, and valuing on the other.¹⁰² Frankfurt imagines this possible objection.

⁹⁸ Ibid., p. 162.

⁹⁹ Ibid.

¹⁰⁰ There's a chance Frankfurt's reply to Lear undercuts this point. Frankfurt, "Reply to Jonathan Lear."

¹⁰¹ Frankfurt, "On Caring," p. 162. Because of its existential importance, caring has intrinsic value for us. Frankfurt asserts that it is better for us to care about what is truly valuable, but the value of caring is not entirely derived from the value of what we care about.

¹⁰² What I want/desire
What I prefer
What I judge valuable
What I will pursue for its own sake

Isn't desiring something badly the same as regarding it as important? After all, if we desire something badly enough, its lack will result in frustration at least, and isn't that harmful? Doesn't someone who wants something badly enough "necessarily care about avoiding the frustration of his desire."¹⁰³ His response is that it's important to draw a distinction between caring about the object of one's desire and avoiding frustration. In some cases we'd just as soon take a pill to extinguish the desire as achieve our object.

Caring about an object is also distinct from judging valuable. You can judge an object, like the taste of ice cream, intrinsically valuable, even pursue it, but not care about it.¹⁰⁴ Judging valuable is "recognizing that it *qualifies* to be *desired for its own sake* and *pursued as a final end*." Caring, on the other hand, is regarding as important to self.¹⁰⁵

Caring also has instrumental value for us. Something can be important to us without our being aware of it. The example Frankfurt uses is vitamins.¹⁰⁶ Although we might not be aware of it, we need these nutrients. Lack of these nutrients harms us by eventually undermining something we can't help caring about. (For most of us,

vs.

Decide or discover (note equivocation) what I care about

¹⁰³ Frankfurt, "On Caring," p. 158.

¹⁰⁴ I suspect this only works with the trivial. Or maybe judging valuable is just judging that the object in question could be legitimately important to someone, but then perhaps the taste of ice cream would not meet this criterion, i.e. it could not play this role – importance – in anyone's life. Normally, however, we don't bother to say something valuable unless we judge that it's legitimately important to most people, i.e., that it is bound up with something they can't help caring about, whether they recognize that fact or not.

¹⁰⁵ Frankfurt, "On Caring," p. 158.

¹⁰⁶ Harry G. Frankfurt, "On the Usefulness of Final Ends," in *Necessity, Volition, and Love* (1999), p. 87.

this would be our health.) If we fail to regard *anything* as important to us, if we fail to care, we'll neglect our needs, leaving them to vagaries of instinct and luck.

There's a potential ambiguity in how Frankfurt uses "needs" here. Given what he says about evolution's seeing to it that animals are responsive to reasons, I think he's committed to an objective conception of needs.¹⁰⁷ In other words, even if I cared about nothing at all, I'd still have needs in this biological sense. However, he seems to be working up to a conception of *persons* as, essentially "caring" entities: "our essential natures as individuals are constituted...by what we can't help caring about."¹⁰⁸ Most of the time when Frankfurt talks about importance and needs, he's indicating needs derived *from* concern.

One of caring's "modes" is love. "Among the things that we care about there are some that we cannot help caring about; and among the things that we cannot help caring about are those that we love."¹⁰⁹ Frankfurt focuses on two features of caring as love: (1) its "disinterestedness" and (2) the fact that our concern for the flourishing of our beloved is itself "volitionally constrained." We are disinterested when we love

¹⁰⁷ Frankfurt says that animals have the capacity to respond to reason "My own construal of reason and of rational behavior do not include any essential reference to beliefs"; "natural selection can be expected to see to it that many organisms...respond to reasons..." Frankfurt, "Reply to Eleonore Stump," p. 62. However Frankfurt also says, "Animals of many species have desires, but only animals of our species—or, perhaps, of a few others—are capable of *seeing* anything as a reason." (My emphasis.) Frankfurt, "Reply to T. M. Scanlon," p. 184.

This objective conception *may* be what he thinks underlies morality, but he may also think morality has to be derived from the distinctive needs of *persons*, beginning with the foundational need for caring.

Elsewhere, discussing Kant, he allows that the authority of the moral law is supposed to derive from its theoretical necessity, but notes that the actual practical effectiveness of the moral law in determining the conduct of someone who fully understands the law is, nevertheless, contingent. It depends upon whether his respect for the moral law is strong enough to compete with his inclinations. Harry G. Frankfurt, "Autonomy, Necessity, and Love," in *Necessity, Volition, and Love* (1999), pp. 140-41.

¹⁰⁸ Frankfurt, "Autonomy, Necessity, and Love," p. 138.

¹⁰⁹ Frankfurt, "On Caring," p. 165.

because considerations that bear on the flourishing of what we love move us directly. We “identify” with our beloved. We don’t have the “one thought too many” problem¹¹⁰ nor do we need to take our love itself and whatever benefits it confers upon us as a reason, rather, our love is manifest in the reasons we take ourselves to have. This point is tricky: in Frankfurt’s view persons are entities for whom loving is a kind of existential imperative, therefore, our need to love *is* a reason; however, paradoxically, “what serves the self-interest of the lover is, precisely, his selflessness.”¹¹¹

To illustrate this point, Frankfurt considers the following example. The fact that money will help a needy person is a *reason* for both the person who acts out of love and acts out of duty. In the duty case, however, the fact that the money will help a needy person’s being a reason is contingent upon a motive, caring, or desire to do what duty commands. In the love case, the motive is built in. “The immediacy of the linkage between loving and what counts as a reason for doing things that help the beloved is part of what *essentially constitutes* loving.”¹¹² The lover does not infer from the fact that the money will be helpful and the fact that he loves the needy person to the conclusion that he has a reason to help the needy person. He does not take his love as a reason. Rather, he takes the person’s neediness directly as a reason and from that *we* infer love.

¹¹⁰ Bernard Williams, "Persons, Character and Morality," in *Moral Luck* (Cambridge: Cambridge University Press, 1981).

¹¹¹ Frankfurt, "On Caring," p. 174.

¹¹² *Ibid.*, p. 176.

Secondly, our concern for our beloved is “volitionally constrained” because the concern for our beloved is not under our direct volitional control, and, in any case, we are unwilling to modify it.

Love is essentially a somewhat non-voluntary and complex *volitional* structure that bears both on how a person is disposed to act and upon how he is disposed to manage the motivations and interests by which he is moved. Thus, love not only shapes a person’s conduct with respect to whatever it is that he loves. It also guides him in supervising the design and the ordering of his own purposes and priorities.¹¹³

Paradigm case: concern of parents for their infants and small children, but the “beloved” can also be abstract – even a “moral ideal.”

Frankfurt thinks that the kind of self-regulation we care about when we are thinking that there is more to freedom than simply being the cause of our actions (as when we are trying to change or come to grips with our bad habits, or to live up to our ideals), can be described without recourse to evaluation, an evaluative point of view, or Reason, which might be thought to provide the form and lever for change. His idea rather is that this form of self-regulation springs from discovered commitment (caring), grounded in our actual behavior – commitment that is susceptible to strategies of modification we might employ, but which nevertheless is not under our direct control. The case of love is the case in which, despite the presence of counter-incentives or considerations, we discover a second fact, which is that we have not the slightest tendency to employ such modifying strategies. This discovered commitment we find ourselves unwilling to modify, love, functions as a grounding personal necessity, generating categorical reasons for us, constituting our volition, our personal

¹¹³ Ibid., p. 165.

and individual essence, and getting reason itself off the ground.¹¹⁴ “The essence of a person...is a matter of contingent volitional necessities by which the will of a person is as a matter of fact constrained;”¹¹⁵ “A person acts autonomously only when his volitions derive from the essential character of his will.”¹¹⁶

PROBLEMS WITH CARING

To make this distinction regarding importance, which the concept of wholeheartedness is not sufficient to express, Frankfurt introduces the concept of caring. But in spite of his saying that caring is “substantially equivalent” to “regards as important,” he describes caring in terms of dispositions, stopping short of saying caring entails belief. He claims that one can care despite one’s sincere, straightforward disavowals. This interpretation suggests that just as whatever in fact moves one is one’s strongest desire, so whatever in fact moves one is what one cares about, particularly since caring is a matter of disposition and can exist despite straightforward disavowals.

Frankfurt says that: Caring is “substantially equivalent” to “regards as important”¹¹⁷; “people necessarily consider whatever they care about to be important

¹¹⁴ Carol Rovane develops an idea that I think is similar to what Frankfurt has in mind. She claims that “an individual person’s commitment to being rational, in the sense of being committed to achieving overall rational unity within itself, is necessarily a conditional commitment to unifying projects [loves, Frankfurt would say] that require such overall rational unity.” Rovane, *The Bounds of Agency*, p. 247.

¹¹⁵ Frankfurt, “Autonomy, Necessity, and Love,” p. 138.

¹¹⁶ *Ibid.*, p. 132. What seems to ground the necessity is the same thing that warrants our being able to say that someone cares or loves implicitly – effective desire.

¹¹⁷ Frankfurt, “On Caring,” p. 155.

to them; and, conversely, they necessarily care about anything that they consider important to them."¹¹⁸

That caring entails belief. The characterization, "necessarily consider whatever they care about to be important to them" sounds a lot like a belief. It sounds as if Frankfurt is suggesting that where there is caring there is also the belief that something is important to one. However, Frankfurt does not say that "necessarily consider important to" is *the same as* caring. What Frankfurt says is that "necessarily consider important to" is "substantially equivalent to" caring. I take it from this that he does not mean to reduce caring to the belief that something is important to one. So, notice that while Frankfurt is careful not to reduce caring to a belief (about an object and one's desire for it), caring seems to entail a belief (about an object and one's desire for it).

That caring involves a disposition to sustain a desire. Caring entails a belief that is borne out if there is a "disposition" to sustain one's desire for the object one ostensibly cares about, should that desire threaten to subside. This part about the disposition to sustain a desire is a little tricky to understand.

The puzzling example Frankfurt uses to illustrate this idea is that of a person who loves music – a person who, for the sake of doing a friend a favor, turns down a chance to go to a concert, but whose desire to attend the concert persists, and it persists not because of some "volitional inertia" but because of some "activity" on the music-lover's part. Frankfurt chooses to explain this activity, this commitment to sustain the desire to go to the concert in terms of higher-order desiring. He says that

¹¹⁸ Ibid., p. 156.

the music-lover's caring is a matter of his identifying with a "higher-order desire" that his first-order desire to go to the concert "not be extinguished or abandoned." Interestingly, Frankfurt suggests that this "higher-order" desire need not be "fully conscious or deliberately explicit;" the bottom line is that the music lover's "caring about going to the concert implies that he is disposed to support and sustain his desire to go to it."¹¹⁹

Difficulty in distinguishing a simple disposition from a disposition to sustain a disposition. You might wonder how we could distinguish between a dispositional account of having a desire per se and a disposition to sustain that disposition. The desires that have proven most vital to human survival over the course of our species history are manifested as powerful dispositions that are very difficult to override even in circumstances where they are no longer functional. For example, our cravings for foods that are fatty, salty, or sweet, persists in an environment in which the greater danger is that we will consume too much of these foods. This looks like dispositional evidence of care, wherein the desire perennially reemerges, no matter what, as if some agency were sustaining it. The funny thing is, though, that Frankfurt intends caring to be a second-order phenomenon, but these sorts of perennial dispositions are often paradigmatic cases of what Frankfurt used to treat as first-order desires. They don't seem to be reflective. Frankfurt's favorite example of love is his love for his daughters, but this sort of attachment is precisely that sort of instinctual passion that is non-reflective and first-order.

¹¹⁹ Ibid., p. 159-61.

An interpretation of caring: Dispositions around which we organize our lives.

What Frankfurt could say in reply is that he has built his life around his attachment to his daughters and if he were just to wake up one day and find that attachment gone, he would be lost, bereft, and would desperately try to recover it. Perhaps the most obvious example wherein a powerful first-order disposition naturally wanes and in some cases people struggle or, in Frankfurtian terms, display a "disposition" to maintain it, is sexual passion. (Contrast with Bertrand Russell's story of discovering during a bike ride that he didn't love his wife anymore, and abruptly dropping her.) Caring, then, could be something that originally develops out of a dawning appreciation of instinctual drives and gratifications without which we are lost (reflective awareness is the beginning of conscious commitment).

Complication: higher-order desire is not sufficient for caring – belief about an object's importance is needed. At first it might look as if higher-order desiring and caring are the same, but notice that having a higher-order desire does not seem to be sufficient for caring, at least in the "regards as important" sense Frankfurt emphasizes. For example, while I'm making plans to indulge my admittedly bad habit, I may have a second order desire to be moved by the desire to engage in the habit, which I may nonetheless not "regard as important." This could simply be a matter of my not wanting to be frustrated, and Frankfurt does explicitly distinguish between caring about avoiding frustration and caring per se, but what about this concern to avoid frustration? A concern to avoid frustration might be manifest in a desire that *all* my first-order desires persist through their satisfaction. If regarding something as important requires discrimination, then it would seem that this blanket

higher-order desire – that my all first-order desires persist through satisfaction – would in fact smother concern. It also seems that I could positively want a desire to persist – as a source of minor pleasure, say – without regarding it as important. For instance, I might desire that my desire for coffee ice cream persist. If this example shows that higher-order desire – even a higher-order desire that a given first-order desire persist – is not sufficient for caring, then it looks as if the judgment that something is important to me, that I need it, that I'll be harmed by absence, etc., is essential to caring.

Further complication: it looks as if caring can entail contradictory beliefs.

On the other hand, Frankfurt also says that you can care even when you see the caring is irrational, wish you didn't, and "despite strenuous attempts to stop." This suggests that caring can persist in the face of my judgment that something is not important to me. This looks like a contradiction. It looks as if in order to distinguish caring from second-order desiring, we have to stipulate that caring necessarily involves the judgment or belief that something is important to me. However, if caring can persist in the face of my belief that that very thing is not important to me (that it's irrational, that I have a reason to stop), it is truly unclear what Frankfurt means by caring.

Furthermore, although there's certainly an ordinary sense in which we can care in spite of ourselves – falling in love against what one takes to be one's better judgment, as Elizabeth Bennett does in "Pride and Prejudice" – it's very hard to see how Frankfurt can distinguish this kind of caring from the addictive desire of the unwilling addict.

Perhaps Frankfurt needs to develop the thought that caring is substantially equivalent to “regards as important” and to keep trying to characterize the reflexive character of that, and to drop the idea that what persists in the face of all reason and attempts to stop really is caring.

In order to distinguish between wanton willing and volition, Frankfurt needs caring to involve explicit belief. I have been interpreting Frankfurt's definition of caring, "substantially equivalent" to “regards as important," as entailing an explicit belief that something is important (to the self). I think Frankfurt wants to allow that "regards as important" could be understood as an implicit belief. But even if you have an overall dispositional account of belief, you still would be able to distinguish implicit from explicit beliefs. Both would show up in behavior, but an explicit belief would show up directly in linguistic behavior and would therefore be susceptible to modifications that an implicit belief would not. In other words, whether the belief that something is important to one is implicit or explicit can make an observable difference. A person who holds a belief implicitly will behave differently from a person who holds the “same” belief explicitly. (At the very least, the person holding the implicit belief might say, “ah ha!” when the belief becomes explicit.)

Consider again the wanton, a creature that, according to Frankfurt, acts on its desires without the inclination or ability to consider whether it wants those desires to be effective in action, whether it wants to be moved by those desires, whether it cares about those desires in the sense that it has the disposition to sustain them. From this characterization of wantonhood, it looks as if one way of making the distinction between the wanton and the person is that the wanton "cares" implicitly, in the sense

that the wanton has certain fixed dispositions, whereas the person cares explicitly, knows he has certain dispositions and is glad he has them, believes them to be reasons responsive, to be good, and learns to sustain them. It looks as if caring must be explicit, involve an explicit belief that something is important to one in order for "caring" to mark the distinction between the wanton and the person, between wanton willing and volition.

If this sketch is roughly accurate, then it looks as if caring, the way Frankfurt sometimes insists on using it, particularly when he's talking about "necessity," may not help us to distinguish effective desires from one sense of true – "important" – desires because caring just collapses into effective desire.¹²⁰

EXPLICIT TO IMPLICIT – SOME CONCLUSIONS

The task in this chapter has been to say how Frankfurt moves from an explicit to a more implicit interpretation of identification, and how authentication, hence authorization, at first follows identification, then caring. My worry has been that with implicit interpretations of identification and caring, Frankfurt might not be able to say how it is that one can be prevented from getting what one most wants by one's own motivational system. The challenge was for him to be able to give a convincing account of autonomy without resorting to Reason. He's building a desire-based account. In order to have an account that's true to the phenomenon of volitional struggle, he needs to be able to distinguish true desires from effective desires. He needs to be able to say why the desire that actually did move us was not expressive of

¹²⁰ Or, perhaps more accurately, standing disposition.

our “true” selves. With identification and caring at bottom implicit, he may not be able to do that.¹²¹

If implicit satisfaction is all that is required for identification or if we can identify with desires of which we disapprove, then acting on desires with which we identify or are identified seems insufficient to characterize an expression of personal volition. I think there are several points here. I’ll try to lay them out.

(1) Acting on desires with which we identify or are identified in the broad sense may amount to an expression of personal volition, but only in the sense that I recognize I am responsible for my action in the sense that I was the cause. It was me acting on a desire I recognize as mine and not some alien force. That it was my action is intelligible to me. This is an important point, but it fails to get at the self-

¹²¹ The argument I tried to make was that acting wholeheartedly is not enough to distinguish wanton acts from volitional acts. (Originally, a hierarchical identification was supposed to mark the distinction. There was something that seemed intuitively right about that, but Frankfurt’s formulation led to a regress problem.) In order to account for volitional struggle, which Frankfurt needs to do if he’s going to fulfill his promise of offering a convincing account of the will that doesn’t rely on reason, it is not enough for Frankfurt to distinguish entities that are wantons from persons. He has to distinguish, in the case of persons, episodes of wanton willing (free action) from exercises of volition (free willing). A wanton is an entity that does not care about which desires move it in action. Indeed a wanton can’t care because it does not have the reflexive self-awareness that’s a prerequisite for caring. This suggests in the case of a person, an entity possessed of such a reflexive capacity, an episode of wanton willing is an episode in which the reflective-caring breaks down, and the person acts unconsciously or impulsively/compulsively against his reflective-concern. This is the model Frankfurt used initially and initially “identification” in the form of a higher order desire was the authorizing reflective-concern. With the satisfaction solution, this shifted so that a person was identified with desires with which he, as a reflective being, was implicitly satisfied or desires upon which he acted wholeheartedly.

I am still not sure whether Frankfurt means to distinguish between satisfaction and wholeheartedness. In order to stop the regress, though, satisfaction is necessarily, at bottom, implicit. Frankfurt says it’s a matter of psychic elements that do not appear. Wholeheartedness, however, seems stronger. I think it’s probably fair to interpret Frankfurt as assuming that acting wholeheartedly entails explicit satisfaction. Acting wholeheartedly upon a desire is sufficient for identification, but, it seems, not necessary. Acting on a desire with which one is implicitly satisfied I think is all that is necessary, so long as one is an entity with the relevant psychic complexity. Frankfurt insists that identification need be nothing more than resigned acceptance and can include identification with desires of which one actually disapproves. This latter hardly sounds like wholeheartedness (and it even begins to erode satisfaction).

regulative aspect of volition -- the extra something beyond reflective awareness, the activity that's at the heart of our existence as selves.

(2) Acting on desires with which we identify or are identified may fail to support a distinction between effective and true desires that seems to be needed to account for the phenomenon of volitional struggle. As long as I recognize an effective desire, I am identified with it and therefore exercising my volition. What about cases in which what I actually do and what I most want diverge? It looks as if this account must say that I am identified with both desires. Identification itself does not mark the difference between them.

(3) What these come down to is the point that acting on desires that I acknowledge are characteristic of me is not the same thing as acting on desires that are important to me. Acting on desires that are merely characteristic of me does not distinguish me from the wanton. Wantons act on desires that are characteristic of them, too. Acting on a desire that I *consciously* acknowledge to be characteristic adds another layer of complexity, but it doesn't seem to be enough. If it were enough, Frankfurt could rest with consciousness as the distinguishing characteristic of persons, but he's very clear that even if it's perhaps distinguishing, it fails to fully characterize what he thinks we regard as important about persons. Consciously acknowledging a desire as mine is not the same thing as caring about it. I can acknowledge desires as mine that I wish I didn't have. It's caring about the desire that distinguishes persons from wantons (in the sense Frankfurt cares about). Being identified or identifying with a desire in this minimal sense is not sufficient for caring about it, for regarding it as important.

If acting on desires I minimally identify with, or desires I acknowledge are characteristic, is not sufficient for distinguishing me from the wanton (getting at the heart of the difference, rather than the reflexivity that enables it), then it seems that caring or regarding as important, not (or not merely) identification, is key to volition. Exercising volition would seem to be acting on desires that are important to me. The problem with caring is that insofar as it is consistent with disavowals and attempts to overcome, caring doesn't mark a distinction between important desires and characteristic desires, either.

IV. TWO STANDARDS OF INTERNALITY

TWO KINDS OF COHERENCE

I argued above that with Frankfurt's wholeheartedness solution to the regress problem, identification becomes more a matter of discovery than invention, as in the discovery of what one can wholeheartedly will, and I noted that Frankfurt has said that identification is matter of acceptance rather than existential assertion, as in an acceptance that may be resigned or even disapproving. I noted that this strange result, combining wholeheartedness with resignation – even disapproval – suggests that Frankfurt thinks that wholeheartedness is something less than total psychic unity. There is more to the story. I think that Frankfurt actually has a couple of standards of internality, and that part of what is confusing about his account is that identification starts out as a way to mark the narrower version and ends up marking the broader version. Another way of putting this is to say that there is some ambiguity about what gets included in the set of attitudes that constitutes either wholehearted unity or psychic ambivalence.

Frankfurt claims that wholeheartedness is equivalent to “satisfaction.”

Formally, as pointed out above, Frankfurt indicated that satisfaction is implicit, not a positive assessment that one is satisfied, but an absence of conflict.

Being genuinely satisfied is not a matter, then, of choosing to leave things as they are or of making some judgment or decision concerning the desirability of change. It is a matter of simply *having no interest in* making changes. What it requires is that psychic elements of certain kinds *do not occur*...¹²²

¹²² Frankfurt, "The Faintest Passion," pp. 104-5.

Frankfurt goes on to emphasize that although “satisfaction is the state of the entire psychic system,” it is also the satisfaction of a certain *kind* of system, a self-conscious, self-reflective system, and that satisfaction requires a person’s (first-person) “appreciation of his psychic condition.”¹²³

[This] absence [of interest in making changes, which is a sign of satisfaction] must nonetheless be reflective. In other words, the fact that the person is not moved to change things must derive from his understanding and evaluation of how things are with him. Thus, the essential non-occurrence is neither deliberately contrived nor wantonly unself-conscious.¹²⁴

This overall sense of reflective satisfaction, derived as it is from “his understanding and evaluation of how things are with him” might seem to be explicit (though at the same time, not “deliberately contrived”).¹²⁵ Frankfurt, however, makes it clear that a personal assessment that one is satisfied is neither necessary nor sufficient for satisfaction.¹²⁶ Does this mean that we can be satisfied, for instance, with desires we are not thinking about, or of which we are unaware? Frankfurt seems to say ‘yes’ and ‘no.’¹²⁷ Yes, we can be implicitly satisfied and identified with unconscious and unreflective desires,

It is possible, of course, for someone to be satisfied with his first-order desires without in anyway considering whether to endorse them. In that case, he is identified with those first-order desires.¹²⁸

¹²³ Ibid., p. 105.

¹²⁴ Ibid.

¹²⁵ Frankfurt uses “reflective” though maybe “reflexive” would be more accurate here.

¹²⁶ Frankfurt, “The Faintest Passion,” p. 104.

¹²⁷ *Self-conscious serenity is neither necessary nor sufficient for satisfaction.* One implication of this is that I can’t know introspectively if my identification “takes.” At the same time as he’s suggesting “satisfaction is a state of the entire psychic system,” he seems to suggest that we can be wrong about whether we are satisfied, about whether a given desire is in fact a wholehearted commitment. Satisfaction is constituted by the entire psychic system – he includes unconscious – so self-conscious serenity is not sufficient for satisfaction in the psychic system that contains the desire in question.

¹²⁸ Frankfurt, “The Faintest Passion,” p. 105.

But no, such desires, even though we might be identified with them are wanton:

But insofar as his desires are utterly unreflective, he is to that extent not genuinely a person at all. He is merely a wanton.¹²⁹

In other words, while implicit satisfaction is sufficient to authorize higher-order desires (after all, this is how Frankfurt gets around the regress problem) first-order desires must seemingly be approved or accepted via direct reflection.

When Frankfurt introduces “wholeheartedness” he speaks of it with respect to a specific psychic element, as opposed to a state of the “entire psychic system,” suggesting that wholeheartedness is a unity that could perhaps exist as an island of satisfaction within the larger psychic system. In what does wholeheartedness toward a psychic element, an attitude, consist?

It consists in his being fully satisfied that they, rather than others that inherently (i.e. non-contingently) conflict with them, should be among the causes and considerations that determine his cognitive, affective, attitudinal, and behavioral processes.¹³⁰

In this sense, Frankfurt suggests, I may discover that I can “wholeheartedly” declare my addictive desire “outlaw.” Doing this, Frankfurt suggests, is sufficient to de-identify with this desire, to purge it from myself.

But how is this whole differentiation and self-assembly process supposed to work? Overall reflective satisfaction, or, more to the point, dissatisfaction, flags potential problems or conflicts, stimulating higher-order review,¹³¹ followed by

¹²⁹ Ibid., p. 106.

¹³⁰ Ibid., p. 103.

¹³¹ One thing that’s not clear is whether our being reflective beings entails automatic generation of higher order attitude in response to conflict so that when a system is in a satisfied state it is identified with its highest level attitudes [negative]. Or whether this wholehearted desire for wholeheartedness is something other than a redescription, in positive terms, of this automatic generation. In other words, is the wholehearted desire formally equivalent, or is there some substantive difference and

acceptance (which can involve prioritization of desires), rejection (deciding to exclude a desire), or ambivalence. You are identified only with what you wholeheartedly will, on the one hand, or with any attitude that can reasonably be attributed to you until you can “wholeheartedly” reject it, on the other. If wholeheartedness, the unity of higher-order desires and the first-order desires they license, just is overall reflective satisfaction, then if my satisfaction is based on a reflective awareness of how things are with *me*, since the addiction I’ve excluded is no longer “mine,” I need no longer consider it in my overall (satisfied) assessment. This seems wrong – particularly given the definition of satisfaction as “simply *having no interest in making changes.*” It also seems wrong to say that I consider it, but am satisfied, since it is safely outlawed. (This is the move the ambivalence/conflict distinction relies upon.) What seems right is that the addictive desire will continue to be included in my overall reflective assessment, and that it will continue to throw up red flags. An initial tension between two or more of what are in the ordinary sense my desires does not disappear when the desire is ruled “outlaw” and externalized. It becomes a tension between the desires or urges I actually have, that “assail” me, and the desires I’d like to or ought to have, my ideal self-image.¹³²

TWO KINDS OF INTERNALITY

This lack of resolution, I think, stems from two kinds of internality.

[positive] work that it does. [Is the difference between a reasons-responsive and a volitionalist account?]

¹³² This is a distinction that Geoffrey Sayre-McCord and Michael Smith think is missing from Frankfurt’s account.

The first is a broader internality which consists of whatever desires can reasonably be ascribed, by ourselves or others, for purposes of assigning responsibility,¹³³ predicting, or making sense of behavior. This internality will include effective desires as well as “true” desires. These are all desires with which one is “identified.”

The second is a narrower internality which consists of the highest-order desires with which we are implicitly satisfied and so identified, and the first-order desires they license. These are what I’ve been referring to as “true” desires. When our effective desires, the desires that have propelled us into actions, are aligned with these higher-order “true desires,” our action is autonomous.¹³⁴

The reason Frankfurt has these two standards of internality is that he does not first make an internal/external distinction to locate an “agent” in order to establish causal responsibility, and then ask, of the attitudes and motives attributable to the

¹³³ Watson indicates that it’s a well-known fact that the bar for freedom is lower than the bar for responsibility. “Introduction,” Gary Watson, *Free Will*, ed. Gary Watson, Second ed. (Oxford, UK: Oxford University Press, 2003), p. 18. I have been assuming the opposite: that freedom is developed by being assigned and/or taking responsibility where one initially has no control or where consequences were unintended. Dennett discusses this: Dennett, *Elbow Room*, p. 143-44. On the other hand, minors may be “free” in many of the senses we care about, but are not held fully responsible for their actions.

¹³⁴ *Alternative internalities*. One way of construing internality that comes up over and over again is what you might call self-attribution. Whatever attitudes one attributes to oneself in the course of understanding one’s past actions and what it makes sense to do next. Judging by what Frankfurt says in reply to Moran, I think he’d be friendly to this. Velleman pushes this as the self of autonomy and practical reason (J. David Velleman, “Self to Self,” *Philosophical Review* 105 (1996).); Joseph Raz says something like this (Raz, “We Are Ourselves: The Active and the Passive.”); David Wiggins says things along these lines in an essay he revised for Watson’s second edition of *Free Will*. Dennett quotes Kant as saying that the motive that drives practical thought is a “concern for acting in accord with reason” or “what animates practical thought is a concern for acting in accordance with reasons” (Daniel C. Dennett, *Freedom Evolves* (New York, NY: Viking Penguin, 2003), p. 286.) which, though it is more than this, might also be construed as a concern to make sense of one’s actions; Rogers Albritton, denying that addictions pose a threat to the will per se, says that will is just the capacity to form intentions or decide in favor of one course of action or another (Watson’s summary, introduction to, “Free Will” 2nd edition);

Maybe there are really three internalities: (1) anything that happens in psychic history that there is any basis at all for attributing to one, even if only as passing experience; (2) self-attribution; (3) wholehearted or volitionally necessary willing.

agent, which are truly expressive or have authority. Rather, I think Frankfurt presumes a rough and ready internal external/distinction (the broad internality above) and offers his own internal/external distinction as something of an answer to the authenticity/authority question (the narrow internality above).

He originally draws a parallel between the problem of saying which bodily movements are actions versus which just happen to me and moves to the problem of saying which desires are mine versus which just happen to me.¹³⁵ This could be restated as: Of those desires that precipitate behavior, which precipitate behavior that is autonomous and hence constitute my volition? His first answer had to do with decisive identification. His second answer has to do with, negatively, a stability that is indirectly reflective (satisfaction), and positively, wholeheartedness, caring, and love.¹³⁶

¹³⁵ According to Frankfurt:

There is in fact a legitimate and interesting sense in which a person may experience a passion that is external to him and that is strictly attributable neither to him nor to anyone else. Recognizing this need not prevent us from agreeing with Penelhum that such a passion is part of the person's ongoing history. It may be noted, moreover, that declining to attribute to a person certain of the passions he experiences does not commit us to regarding those passions as altogether irrelevant in reaching a fair judgment concerning what we can expect from him. A passion is no less genuine, and its thrust is no less forceful, for being external to the person in whose history it occurs, any more than a bodily movement is less palpable in its occurrence or in its effects for being a movement that is not made by the person in whose body it occurs. Frankfurt, "Identification and Externality," p. 61-62.

¹³⁶ Internality, for Frankfurt, marks the distinction between what might be described as wanton agency or mere willing – action that lacks reflective endorsement – and personal volition, which he also (confusingly) refers to as willing. One consequence of drawing the distinction here, between wanton willing and personal volition, is that strictly speaking it does not allow for *intrapersonal* conflict, only conflict between the person and, let's say, the person's human host. It is not clear how to think about the person and the host, on Frankfurt's account.

Is this a consequence of Frankfurt's shift from a strictly hierarchical model to a coherence model, or is it a problem on the hierarchical account as well? The problem is more conspicuous on a coherence account because the coherence account shifts emphasis from episodes to entities.

You might think that I should say the conflict occurs between the person and his wanton host, but the fact that a person-seed (reflexive capacity) transforms a wanton host is something Frankfurt makes clear in "The Faintest Passion," where he says that a desire's failing to perturb a reflective *system* is enough to count it as internal to the person. Previously, on the original hierarchical model,

But if wholeheartedness is the criterion of internality, there is a problem.

Wholeheartedness is strictly defined as an absence of ambivalence and "Ambivalence is constituted by conflicting volitional movements or tendencies, either conscious or unconscious, that meet two conditions": (1) "inherent;" (2) "internal," not "alien."¹³⁷

Let's look at this second condition. In order to count as ambivalence, and not the "virulent conflict" that is compatible with wholeheartedness, the opposing tendencies must occur within the self. Internally. But the boundaries of the self are precisely what are at issue here. In order to solve the regress problem, Frankfurt decided that we no longer have to positively "identify" with our second-order desires (desires to be). Rather, we just need to be satisfied with them, which shows up as a lack of interest in making changes. But suppose there is ambivalence (to be or not to be...). That ambivalence constitutes a lack of satisfaction. But, by the wholeheartedness criterion of identification, satisfaction is what authenticates these second-order desires – grants them internal status. This line of reasoning seems to show that ambivalence is conceptually impossible, since ambivalence undermines the criterion of internality. Internality is just the absence of ambivalence. In other words, looks as if in order to be involved in ambivalence, as opposed to mere "conflict," a desire must qualify as "internal." Yet in order to be "internal," in order to have the possibility of being involved in ambivalence, a desire must be wholehearted, which seems to exclude the possibility of ambivalence.¹³⁸

what marked a wanton was lack of concern about first order desires. Now it seems to be lack of any awareness altogether: "...insofar as his desires are utterly unreflective, he is to that extent not genuinely a person at all. He is merely a wanton."(Frankfurt, "Faintest Passion, p.106"Frankfurt, "The Faintest Passion," p. 106.)

¹³⁷ Ibid., p. 99.

¹³⁸ Here is what Watson says:

WHOLEHEARTEDNESS AND IDENTIFICATION

What is going on? Frankfurt has two standards of internality in play. He does not clearly or explicitly distinguish them. As an interpretation of identification, wholeheartedness is supposed to be an explication of the narrower standard, just as higher-order desiring and decision, earlier, had been. But wholeheartedness actually just presupposes the narrower standard. Wholeheartedness is defined negatively as absence of ambivalence, and ambivalence Frankfurt stipulates, is “internal.” Given that absence of ambivalence is a matter of being able to act decisively and with satisfaction, it is hard to see how such a standard could exclude as external desires with which we struggle. Frankfurt just stipulates that it does.¹³⁹

When Frankfurt later says that we can identify with attitudes of which we disapprove, the equation of wholeheartedness, identification, and the narrow standard of internality begins to unravel. If we can identify with desires of which we disapprove, then wholeheartedness, as an interpretation of identification, is either not unified enough to perform its earlier authenticating function (pick out “true” desires from effective ones) and Frankfurt loses his ability to account for autonomy in terms

The idea of volitional struggle as opposition to one’s own identifications appears to be ruled out by the doctrine of “The Faintest Passion,” according to which identification “is constituted...by an endorsing higher-order desire with which the person is satisfied” (Ibid., p. 105.), where “satisfaction” entails “an absence of restlessness or resistance” (Frankfurt, “The Faintest Passion,” p. 103.). On the satisfaction criterion, neither that against which one struggles, nor that on behalf of which one struggles could be said to be defining *complete* or *full* identification. Elsewhere, Frankfurt speaks of ambivalent agents as “*in part*” identified” (my emphases). Frankfurt, “Autonomy, Necessity, and Love,” pp. 137-8. Thus cases of volitional struggle would be instances of conflicting partial identifications. And this in turn could be a matter either of contrary caring, of a divergence between caring and endorsement. (Gary Watson, “Volitional Necessities,” in *Contours of Agency*, ed. Sarah Buss and Lee Overton (2002), p. 159.).

¹³⁹ But then he equates wholeheartedness with “satisfaction.” Technically, according to the structure of the argument, that is, satisfaction should have the same scope, refer to the same set of desires. In fact, I think it tends toward the broader standard. When Frankfurt says that it refers to the “entire psychic system” and is a matter of mental events that do not occur, he should be referring just to higher order desires, but it is very tempting to read it more broadly.

of wholeheartedness, or identification has split from wholeheartedness and migrated from the narrow standard to the broad standard of internality, leaving behind its association with authorization.^{140 141}

Suppose that Frankfurt allows identification and wholeheartedness to come apart such that it is wholeheartedness that marks narrow internality, not identification.¹⁴² One problem (above) is that wholeheartedness is made out in terms that presuppose internality – as in the absence of *internal* ambivalence. But Frankfurt would likely reply that the distinction between higher and first-order desires is still a way of getting at internality. One's higher-order desires are internal, as are the effective desires they license.

The original problem with the first-order/higher-order distinction as a way of establishing internality was that one could ask what licenses the higher-order desires, given that Frankfurt is worried about wantonness of unlicensed first-order desires.

A further problem with leaning too heavily on the first-order/higher-order distinction for establishing internality is that there are reasons for doubting that this distinction itself holds. For instance, how is it possible to distinguish desires to *be* from desires to *do* when the desires to do can be complex and ambitious, like desiring to circumnavigate the globe in a balloon, and desires to be can be simple, like

¹⁴⁰ A person may identify “mindlessly” and be criticized for that, but the desires are nonetheless his own. Frankfurt, “Reply to Richard Moran,” p. 223.

¹⁴¹ If both, if identification has migrated from the narrow standard to the broad standard of internality, taking wholeheartedness along with it, then Frankfurt has entirely lost his narrow internality and his criterion for distinguishing true from merely effective desires. I tried to make this point out earlier in terms of an inability to distinguish characteristic from important desires. The fact that Frankfurt needs to introduce caring, even though that does not quite solve the problem either, might be evidence that this is how it goes.

¹⁴² Evidence for this appears in Frankfurt's reply to Scanlon “we have relatively few desires with which we do not – either wholeheartedly or ambivalently – identify ourselves” Frankfurt, “Reply to T. M. Scanlon,” p. 185.

desiring that my desire to consume another spoonful of ice cream be effective? I think this problem stems from the fact that Frankfurt has stipulated that the object of a first-order desire is an *action*.¹⁴³ Arguably, desires to do are already implicitly desires to be. That is, first-order desires perhaps could've been defined by Frankfurt as a desire *that* something be the case, that the world be a certain way, but Frankfurt defines them more narrowly, as a desire *to* *x*. Insofar as a desire to *x* can also be described as a desire *that* I *x*, it seems that this implies a desire to be a person who *xes*. While there is something intuitively appealing about Frankfurt's distinction between "first-order" and "higher-order" desires, I also think it's fair to conclude that he has yet to formalize this distinction in a way that's completely convincing.

If identification and wholeheartedness split, it looks as if Frankfurt may now be left addressing the original regress challenge with the reply that first-order desires don't necessarily have to be reflectively endorsed, yet they have to enable one to act reflectively (to some degree), decisively and with substantial satisfaction. As long as one is acting, let's say, reflectively, that is, self-consciously, decisively, and with substantial satisfaction, one is acting autonomously. Desires that effect such actions are true; desires that effect unconscious, incomplete, or substantially unsatisfying actions are merely effective. Identification, in the *broad* sense, is necessary for self-conscious action, free action, but not sufficient for autonomous action, free willing.

This sketch gives us a sense of how it is that merely effective desires compromise autonomy: they precipitate unconscious, incomplete, or unsatisfying action. It's easy to see how addictions and compulsions fit the bill. Would such a

¹⁴³ Frankfurt, "Freedom of the Will and the Concept of a Person," p. 13.

sketch, if it could be filled out, help us to distinguish persons from wantons, as was Frankfurt's original aim? Would it help us to say how I (my true desires) can be defeated by my own motives (effective desires)? This was why we were interested in identification in the first place. I'll address these questions and some others below.

V. WHAT MIGHT WORK

So far, I've been focusing on whether and how "identification" helps Frankfurt to distinguish effective desires from true desires. The distinction between effective and true desires is important both in accounting for volitional struggle and distinguishing persons from wantons. I argued first that Frankfurt has been moving from explicit to implicit interpretations of identification and, later, caring, and that on the implicit interpretations, he is hard pressed to maintain the distinction between effective and true desires. I pointed out that this is related to the fact that there are (at least) a couple of standards of internality at work, and that Frankfurt does not always clearly distinguish between them. I argued secondly that Frankfurt has had difficulty articulating a sense of identification that can be used to establish a criterion of internality without at the same time presupposing it.

One way of construing how Frankfurt distinguishes effective desires from true desires is by how they move us to action – wholeheartedly or not – where wholehearted action is action that is conscious, decisive, and satisfying. This way of drawing the distinction involves essential reference not just to a subset, but to the broadest set of attitudes – throwing them all into one pot, so to speak, and rejecting the conflict/ambivalence distinction. This approach seems appropriate as the issue is autonomy or self-rule, which arguably applies to the person on the broadest construal.¹⁴⁴ But the distinction may still be a difficult one to draw. When an

¹⁴⁴ It would be interesting to compare this overall wholeheartedness view to Carol Rovane's view. Rovane gives an account of personhood which includes an account of "self-change." She contrasts her view with Frankfurt's early view: He claims that actions are free when they are in accord with one's desires, and one's will is free when it is in accord with one's *higher-order* desires – for

effective desire is not wholehearted, it is not always clear what the significance is. A seeming breakdown, weakness of will, could be the beginning of a change, of growth, of possibly even greater autonomy – Huckleberry Finn’s weakness in helping Jim, for instance.

Weakness of will in the face of addiction is a problem because it’s an instance of non-wholeheartedness, of conflict. But internal conflict doesn’t seem to be enough to characterize the problem of addiction. Perhaps it’s because the conflict is chronic. But it also seems that even among cases of chronic conflict, which we might agree is undesirable, there are better and worse cases. The better case is the person who resists “temptation,” or what she accurately judges to be bad. The worse case is when the person is chronically conflicted, but gives in. Frankfurt wants to say that the person who resists chronic temptation *is* wholehearted. This is his struggle/ambivalence distinction, but this distinction depends up on a criterion of internality that Frankfurt has thus far been unable to establish. On the other hand, our intuitions about these cases may also have to do with which conflicts are stable and which spin out of control, and Frankfurt could conceivably still use wholeheartedness to address these intuitions (see “Degenerate desires” below).

Insofar as Frankfurt allows true desires, they are wholehearted desires, but truth in the sense of wholeheartedness, on the present interpretation, is a matter of

example, second-order desires by which an agent can, in principle *change* it’s lower order desires. However it should not be thought that the account of reflective rational agency that has been developed in this book shares the same *hierarchical* structure that Frankfurt attributes to personal agency. He supposes that the direction of self-change is always from higher-order desire to lower-order change, whereas this account simply posits a general commitment to overall rational unity, in the light of which a person might have reason to revise *any* of its attitudes, no matter whether they are of the first, second, or *n*th order. Thus the direction of self-change is not from higher or lower order; it is rather from a state of less rational unity to a state of greater rational unity. Rovane, *The Bounds of Agency*, p. 236, n. 12.

degree. There is certainly something intuitively appealing about understanding wholehearted desires as true. Frankfurt might be able to make out volitional struggle as a struggle for wholeheartedness, and I think there is something right about this as well. When we struggle to bring ourselves to do something, whether or not we succeed, the struggle often isn't just to do a deed, but, if it's something that will come up again, to be able to it with some ease or peace of mind. Certainly this is the case with changing our habits. But even with a one-shot event, the decision to marry, for instance, we surely hope not just to resolve the issue one way or another, to the point of being able to take decisive action, but to do so without lingering ambivalence.

This much seems right, but what happens when we apply this all-inclusive wholeheartedness to the original problem? The picture I presented earlier was of a simple belief-desire account being unable to account for volitional struggle because it had no way of distinguishing strongest desires, what we most want, from effective desires, what we actually do. It seemed that the phenomenon of not doing what we most want is all too familiar and that it is just perverse not to recognize this, yet it seems the simple belief-desire model cannot do the job. What I said was that in order for this type of account, which Frankfurt is committed to, to be plausible, it must offer some distinction between effective desires and "true" desires. I mentioned different ways truth might be made out – notably as authentic or authoritative. Now suppose a desire's being true is not an either/or thing, but a matter of degree. How will this affect a characterization of volitional struggle?

One thing that might look problematic is that in the cases at issue, the cases where you want to be able to identify a true desire as opposed to a merely dominant

desire, these are the cases in which, by the all-inclusive wholeheartedness criterion, there is no true desire. On the early hierarchical account we could say that the desire the person identified with was the true desire, whether or not it was effective. This seemed useful and to have some intuitive appeal. It accounted for how our autonomy might be compromised by our own desires. We could identify with one desire, a desire to kick the habit, our true desire, yet be moved to action by our addictive craving, which on this account is also in some crucial way external to the true self, like an event in the body's history might differ from an intentional movement.

With the satisfaction solution, and wholeheartedness as "truth," in the case of struggle, neither desire to kick the habit nor the addictive craving is "true." When the craving wins, autonomy is compromised because the act was ambivalent. When the desire to kick the habit wins is autonomy compromised? It seems strange to say it, but from another angle, the whole person (note that the person is enlarged here) is not – you might say – free, until his struggle diminishes substantially. This seems right.

So, perhaps even successful struggle to keep the addiction under control can't be wholehearted, at least at first. Tossing the cigarettes across the parking lot as my late father-in-law did isn't yet free. Given the insistent nature of the desire neither smoking nor stopping is satisfying. There's a sense (desires-needs) in which that may actually be true. That is, even a successful struggle to keep an addiction under control may not be sufficient for freedom. When our desires diverge sharply from our needs, no matter what our other attitudes, our autonomy is compromised. Frankfurt has not taken that tack, but he could. He would just have to say that there is more than one sense of freedom and that sometimes they come apart.

What Frankfurt actually says, which also sort of works: “A natural and useful” conception of free will “a person’s will is free to the extent that he has whatever will he wants.” Another interpretation requires not that “he originate or control what he wills, but that he be wholehearted in it. If there is no division within a person’s will, it follows that the will he has is the will he wants.”¹⁴⁵

On the account we’re left with, in the absence of wholehearted action there may be no true desires. In any case, the picture is not of an inner agent or agency that is overcome (as you might get on the Platonic model Watson was suggesting). For Frankfurt, post “Faintest Passion,” true desire/effective desire doesn’t seem to be a psychic struggle of the kind that takes place in a classic tri-partite soul: reason, spirit, appetite; or ego, superego, id, but rather a contrast, perhaps not best described as a struggle, between self and absence thereof.

Does this overall, as opposed merely to higher-order unity wholeheartedness set the bar for autonomy too high? Let’s suppose that as long as one is acting self consciously, decisively, and with sustained satisfaction, one is acting autonomously. Let’s assume for now that this is a sufficient condition. Is it necessary? It seems to set the bar pretty high. That Frankfurt may have thought so is perhaps indicated by the fact that he tries to lower the bar by making wholeheartedness compatible with “virulent struggle.” I’ve said why I don’t think that particular attempt really succeeds. What are the alternatives? Leaving it high might be fine if Frankfurt is willing to say we have different conceptions of freedom for different purposes.¹⁴⁶

¹⁴⁵ Frankfurt, "The Faintest Passion," p. 101.

¹⁴⁶ See Frankfurt, "Three Concepts of Free Action."

Speaking of the rigors of wholeheartedness, John Middleton Murray reportedly has said, “For a good man to realize that it is better to be whole than good is to enter on a strait and narrow path compared to which his previous rectitude was flowery license.”¹⁴⁷ Wholeheartedness, as a threshold for autonomy, may imply that we are scarcely ever autonomous, or fully autonomous. That is a worry if we need to certify that someone acted autonomously before we can hold him responsible for his deeds. Another worry is that wholeheartedness and hence autonomy are more readily achieved by those with the narrowest concerns and simplest psychologies. The simplest psychology that crosses the threshold from wanton to person might be capable of the greatest wholeheartedness, and therefore autonomy or freedom. The worry in part is that the simplicity comes at a cost of awareness of one’s interests (what’s important to one), compromising one’s ability to look out for them. While we might not usually worry or feel pity toward animals for lacking this capacity, we do sometimes worry about persons’ relative simplicity. We sometimes pity or scorn them, live in fear of their incompetence, and/or corrupt ourselves by exploiting it. What worries about responsibility and simplicity may show is that autonomy, wholeheartedness, and volition are, in part, determined and sustained by forces that extend beyond the boundary of any individual person, and that is in fact what Frankfurt suggests, saying in regard to the opposite problem – the problem not of simple-mindedness, but of volitional indeterminacy,

It may be, as Saint Augustine supposed, that a thoroughly unified will comes only as a gift of God. Still, the extent to which people suffer from volitional indeterminacy is not entirely independent of the social, political, and cultural

¹⁴⁷ Quotation attributed to John Middleton Murray in Parker Palmer, *Let Your Life Speak* (San Francisco, CA: Joey-Bass, Inc., 2000). I have also seen it attributed to “Anonymous.” I have not been able to track down the original.

conditions in which they live. Those conditions may either facilitate or impede the development of unambivalent attitudes, preferences, and goals.¹⁴⁸

WHY SATISFACTION IS NOT PASSIVE

I've just given a rather impressionistic account of how Frankfurt might be able to capture some of our intuitions about volitional struggle with "wholeheartedness" sans the identification that was supposed to establish a narrow internality. That is, I have been speaking not of wholeheartedness as an island of unity in a sea of conflicting desire, but simply as an overall unity. I have not yet addressed the question of whether this simplified, all-inclusive wholeheartedness will enable us to distinguish the person from the wanton. I am not sure what to say about this. Frankfurt says that "to be a person, as distinct from simply a human organism, requires a complex volitional structure involving reflective self-evaluation."¹⁴⁹ I think this is right, but I have been investigating how well Frankfurt's articulation of this structure holds up, and, as far as I can tell, it doesn't hold up as well as he would like. I've been focusing on whether and how "identification" helps Frankfurt to distinguish effective desires from true desires. He has had difficulty articulating a sense of identification that can be used to establish a criterion of internality without at the same time presupposing it. With what's left, the way I've told the story, he may not be able to distinguish the person from the wanton. Not without saying more. To illustrate, here is what Frankfurt originally says about the wanton:

The essential characteristic of the wanton is that he does not care about his will. His desires move him to do certain things without its being true of him

¹⁴⁸ Frankfurt, "The Faintest Passion," p. 102.

¹⁴⁹ Ibid., p. 103.

either that he wants to be moved by these desires or that he prefers to be moved by others.¹⁵⁰

Compare this to what Frankfurt has recently said about identification:

what I have actually intended to convey by referring to ‘endorsement’ is not that the agent *approves* of what he is said to endorse, or that he considers it to *merit* his support, but nothing more than that the agent *accepts* it as his own. The sense in which he accepts it as his own is quite rudimentary. It is free of any suggestion concerning his basis for accepting it and, in particular, it does not imply that he thinks well of it.¹⁵¹

A related challenge, or maybe a different way of putting the same challenge for Frankfurt’s present satisfaction solution, is for him to say how it is that this basic authenticating mechanism captures the sense in which we are active that’s at the heart of our existence as selves.

The satisfaction solution looks passive because satisfaction is characterized negatively, as a matter of mental elements that do not occur. Unless my desires happen to generate conflict among themselves, they are, by default, authoritative. But these authoritative desires are not up to me.¹⁵² Desires, as “passions,” are even etymologically linked to “passivity.” The worry is, I think, less about the absence of a choosing self, though that is the purer worry. The stronger worry, the fear, is, in theory, satisfaction could end up authorizing harmful or degenerate desires. In order to avoid the feared outcome – that satisfaction ends up authorizing harmful or

¹⁵⁰ Frankfurt, "Freedom of the Will and the Concept of a Person," p. 16.

¹⁵¹ Frankfurt, "Reply to Michael E. Bratman," p. 89.

¹⁵² Bratman worries about this and reaches back for a solution, decision, Frankfurt has rejected:

Suppose I am not moved to change things and then simply stop there and ‘leave things as they are’...If in such a case one has not actually decided to leave things as they are, one has not, I think, identified with how things are with one...While decisions may be needed for identification, it may not be all that is needed. Perhaps what is needed is, in part, a decision with which one is ‘satisfied.’ Bratman, *Faces of Intention: Selected Essays on Intention and Agency*, p. 195.

degenerate desires – I think the intuition is that desires must be answerable to some external authority, reason or facts about the world. At a bare minimum, the circle needs to be wider, to include beliefs, awareness of needs, of identity as a member of a moral community.¹⁵³

CHOOSING SELF

At various places, Frankfurt tries to address these worries. At one point, addressing the first of these worries, the worry about the choosing self, he suggests that our higher-order desires have authority because we have control over them.

The volitional attitudes that a person maintains toward his own elementary motivational tendencies are entirely up to him. Passions such as jealousy and craving merely provide him with psychic raw material, as it were, out which he must design and fashion the character and structure of his will... Whether a person identifies himself with these passions, or whether they occur as alien forces that remain outside the boundaries of his volitional identity, depends upon what he himself wants his will to be¹⁵⁴

This claim represents his own lingering attachment to the chooser self that is well-represented in the first hierarchical model of autonomy, but it doesn't sit well with the way he develops the source of authority in the satisfaction solution. Jonathan Lear pointed this out, and Frankfurt acknowledged that the satisfaction solution's model of authority is more an existential dependence model than an unmoved mover model. "Following the dictates of love does not constitute acting autonomously because the love is active rather than passive," Frankfurt acknowledges. "It constitutes acting autonomously simply because love is essential to the nature of the self."¹⁵⁵

¹⁵³ Dennett, *Freedom Evolves*, p. 285. Here Dennett, discussing Velleman's discussion of Frankfurt, observes the "requirement that we not make ourselves too small."

¹⁵⁴ Frankfurt, "Autonomy, Necessity, and Love," p. 137.

¹⁵⁵ Frankfurt, "Reply to Jonathan Lear," p. 294.

Frankfurt argues in an article about the importance to us of ideals¹⁵⁶ that the notion of an unconstrained choice, of a will that's purely a chooser, a will that can distance itself equally from every concern, a will that is radically free, is essentially incoherent. He claims in particular that a will that is free even with regard to its own character – a will that could choose and enact its own motives at any moment with no internal resistance whatsoever – has no basis on which to choose. He goes so far as to say that such a will would undermine the very conditions of its existence. He thinks that the concept of reality just is a concept of that which resists the will, that which is not under its immediate control.¹⁵⁷ The idea is that if the will itself was under the will's own immediate and arbitrary – arbitrary because unconstrained – control it also could not exist. Similarly, the very idea of choice or decision presupposes preexisting motives which both enable and constrain the choice. If Frankfurt is right about this, then grounding volitional authority in preexisting desire is not only unproblematic, it's inevitable.¹⁵⁸ The preexisting desire is not by itself active, nor is it actively acquired, but it is a condition of activity, and Frankfurt also thinks, its core.

Frankfurt also works with this theme in his essay on the importance to us of final ends.¹⁵⁹ The having of final ends that the will cannot simply will away, that the will can at best influence only indirectly, is a necessary condition for meaningful choice, decision, and action. Frankfurt calls this “volitional necessity” and he goes on to explicate it further in “Autonomy, Necessity, and Love.” In “The Importance of

¹⁵⁶ Harry G. Frankfurt, "On the Necessity of Ideals," in *Necessity, Volition, and Love* (1999).

¹⁵⁷ Frankfurt, "The Faintest Passion," p. 100.

¹⁵⁸ Unless, of course, Kant is right that the will is a causal power, and as a causal power is determined by a law, its own law, which turns out to be the moral law.

¹⁵⁹ Frankfurt, "On the Usefulness of Final Ends."

Final Ends,” Frankfurt takes issue with Aristotle’s formulation that action aims at the good. Frankfurt says this is misleading, that it’s persons, not actions, who have ends and aims.

He argues that having aims, final ends is important because, even if we lack distinct final ends, we still can’t help caring what happens to us (a claim about human/person nature), and aiming will help us hit the target, take care of our inevitable concerns. This is one way in which final ends have instrumental as well as intrinsic value. But final ends are instrumentally valuable in another way, a way that’s less obvious. Final ends allow us to engage in purposeful activity, which is itself a final end for us, for persons. Frankfurt fears that this point may get lost when we take Aristotle’s formulation of ends belonging to actions. The ability to engage in purposeful action is the ability to live a meaningful life. (He hastens to add that meaningful lives can be hard and miserable. Meaning is not the only consideration...) So, while (according to Frankfurt anyway), on Aristotle’s account the bridle maker’s activity is only instrumentally valuable, not an activity to be taken up for its own sake, Frankfurt argues that it is intrinsically valuable precisely because it is instrumentally valuable. That is, if purposeful activity is intrinsically valuable to us, then the instrumental value of an activity gives it intrinsic value.

While we may not be able to help loving what we love, and are in that sense passive, being bound in this way, having final ends we can’t help having, is essential to our being persons with active, meaningful lives. Does this account of the intrinsic value – the existential value, really – of instrumental activity enabled by ends we can’t help having, in turn, enable us to distinguish the person from the wanton? It’s

not clear. As long as we are invoking Aristotle we might note that on Aristotle's view, the wanton, like all other natural entities, would be subject to its own final cause. If we decline to tangle with Aristotle and his metaphysics, we could go back to Dennett and his intentional stance. From the perspective of Dennett's intentional stance, we assume that wantons have beliefs about the world and desires they're aiming to satisfy. If this perspective proves useful for predicting wantons' behavior, we'd also have to admit that it's useful to think of wantons as having ends and engaging in instrumental activity. They, even more so than us, have desires they can't help having.

DEGENERATE DESIRES

I don't believe that Frankfurt addresses what I'm calling the "degenerate desire" concern directly, or as such, but there are places he indicates that the circle of satisfaction is broader than it sometimes appears.¹⁶⁰

Someone could complain that the plausibility of Frankfurt's account rests in part on optimism about human nature. Such a complaint might reflect worries that if a person's reasons are contingent on his concern, that contingency makes morality too weak. I think that this is not necessarily the case. I think that *some* such worries can be addressed from within a Frankfurtian perspective. Frankfurt can make the case that by far most people's concerns are such that being moral is important to them, whether they recognize it or not. Frankfurt's view allows that reasons may apply to

¹⁶⁰ Problems for my assertion that the circle is broader than it sometimes appears: (1) in reply to Wolf, Frankfurt discusses the freedom of the wholehearted compulsive; (2) also in reply to Wolf, while discussing Hitler, Frankfurt asserts that an immoral life can be a good one (Harry G. Frankfurt, "Reply to Susan Wolf," in *Contours of Agency*, ed. Sarah Buss and Lee Overton (2002), p. 248.); (3) in reply to Lear, Frankfurt discusses the freedom of the slave who loves his master; also Frankfurt lays out love's authority in a way that does not block degenerate desires.

us whether we acknowledge them or not. All that's required is that there's *something* we care about. If there's nothing we care about, Frankfurt hints that we would not qualify as persons. He toys with the idea of saying that the distinctive quality of persons is that they care about themselves: "Perhaps caring about oneself is essential to being a person. Can something to whom its own condition and activities do not matter in the slightest properly be regarded as a person at all?"¹⁶¹ This really is a reprise of his person/wanton distinction in "Freedom of the Will and the Concept of a Person." Frankfurt could argue that as a matter of contingent but reliable fact, the concerns people have are concerns that make morality important to them. Being moral is therefore important on the practical grounds Hume's "sensible knave" would appreciate.

Of course the problem with the sensible knave is not that he doesn't appreciate these practical grounds, it's that he appreciates only these practical grounds. If the first part of the worry about making reasons contingent on caring makes morality too weak because it seems too easy to escape caring about morality by taking the stance of the amoralist, the second part is that even if we back up a minute and recognize that, actually, on Frankfurt's view, everyone capable of personal interaction is going to have a reason to be moral, there is no guarantee that on a view like Frankfurt's, moral reasons for persons are always going to be overriding. I suspect that at bottom this concern is animated by a fear that if we ground reason in preexisting concern and if people persist in their recalcitrant

¹⁶¹ Frankfurt, "On the Usefulness of Final Ends," p. 90.

disregard for morality we (enlightenment rationalists) might not have the authority to force them to behave while respecting them as citizens of the kingdom of ends.

But that is not all there is to it. So far, this has been a prudential argument, but Frankfurt is not committed to thinking that the normative force of morality is strictly prudential. Caring and love are existentially necessary for us on Frankfurt's view. Morality can also be a direct object of this love. It can be a final end for us, and for all Frankfurt's said, it may prove to be the best. That is, caring about morality, or "the good," for its own sake may at the same time prove to be the best route to achieving and sustaining volitional unity.¹⁶²

Of course, I might not care about morality or the good, or anything beyond satisfying my immediate appetites. Where does Frankfurt "broaden the circle"? As I mentioned previously he broadens the circle primarily with his discussion of importance.^{163 164}

¹⁶² See Watson's article on Kant and Happiness for a similar idea, but keep in mind that although Frankfurt allows this as a possibility, he'd probably be *very* skeptical about it. Gary Watson, "Kant on Happiness in the Moral Life," *Philosophy Research Archives* 9 (1984).

Also, Frankfurt's reply to Herman:

we act morally when we are moved by love for a certain kind of world or a certain kind of life. The moral law may be in a sense objective, because it is an objective matter what sorts of conduct are required in order to promote the realization of that world. However, the objectivity of the reasons that these requirements generate is grounded in nothing but the subjective necessities of love. Harry G. Frankfurt, "Reply to Barbara Herman," in *Contours of Agency*, ed. Sarah Buss and Lee Overton (2002), p. 277.

¹⁶³ On the one hand, he admits this concept (importance) is circular. On the other hand he talks about animals as being reasons responsive, suggesting that the related concepts of importance and harm, could also be made out in biological terms. He says that animals have the capacity to respond to reason "My own construal of reason and of rational behavior do not include any essential reference to beliefs"; "natural selection can be expected to see to it that many organisms...respond to reasons..." Frankfurt, "Reply to Eleonore Stump," p. 62.

Frankfurt's vitamin example is pertinent here. He holds that to care about something is to regard it as important. But something can be important even if we don't directly recognize its importance to us – even if we are entirely unaware of its importance. He illustrates this concept with the example of vitamins. Vitamins are important to us whether or not we recognize the role they play in our health. He also holds that when something is important to us, we can be harmed by its degradation; therefore, we can also be said to "need" it. In the human case, Frankfurt wants to make

Given that on Frankfurt's view persons normally will have reasons to be moral, given that these reasons need not be strictly prudential, and for the most part, probably aren't, it becomes harder to imagine much scope for clear-eyed, wholehearted wickedness. I guess the point to make here is that in broadening the circle, Frankfurt is implicitly bringing in the "clear-eyed" criterion. In a reply to Susan Wolf, Frankfurt admits that if the life Hitler chose did not damage anything he cared about then it was a "reasonable" life for Hitler to choose. "It is possible, I am sorry to reveal, that immoral lives may be good to live."¹⁶⁵ The way Frankfurt tells the story, Hitler may have had a good life. A clear-eyed criterion would emphasize a certain amount of epistemic responsibility. If Hitler's "good life" entailed self-deceit or willful ignorance, then it may not have been such a reasonable choice after all.

Closely related, but not quite the same issue, I think, is the concept of degenerate desires. While we may enjoy the freedom to act on them, while we may wholeheartedly embrace them, they are desires that lead us down the primrose path of diminishing returns...and worse. How could Frankfurt block this? How could he say that wholehearted addictive or other degenerate desires compromise autonomy, especially given that such desires can provide focus to a life that was previously lacking it, temporarily "increasing" autonomy by diminishing ambivalence? In order to have a reason not to take the drug it need only be the case that taking the drug

harm out in terms of care such that we are only harmed when we care – vitamins derive their importance by impinging on something we do (consciously) care about. However, if animals are also reasons responsive, then it would seem that conceptions of harm and importance can be made out independently of concern. Although the explanation would still be teleological and refer to some interdependently defined conception of an organism's functioning and its interests, it would not be subjective in the sense that Frankfurt is pushing for.

¹⁶⁴ Frankfurt, "On the Usefulness of Final Ends," p. 87.

¹⁶⁵ Frankfurt, "Reply to Susan Wolf," p. 248.

interferes with something I care about – whether I recognize this fact or not. How could Frankfurt get from my having a reason I don't recognize to an assertion that my autonomy is compromised, even though I am wholehearted? The threat to my autonomy is based on a kind of hypothetical – not if I cared I'd have a reason, but since I do care, if I knew about the drug's destructive effects on the object of my concern, I'd be motivated, my wholehearted desire to take the drug would fracture into ambivalence, and my taking of the drug would not be an exercise of autonomy.

I guess the question is whether wholeheartedness and autonomy refer to actual or hypothetical satisfaction. This is ambiguous. If it is hypothetical, then Frankfurt can say how addictive desires, even if wholeheartedly embraced, compromise autonomy. I believe that Herman reads Frankfurt's satisfaction as closer to hypothetical and Frankfurt strongly approves of Herman's account of his view.¹⁶⁶ But the cost might be too high, because it might imply that whenever I act on a desire I have reason to thwart, I am not exercising my autonomy. If this were the case, life being the complicated mess it is, I might live an entire life without once exercising autonomy.

In any case, in his early articles, Frankfurt heads in the opposite direction, showing how an effective addictive desire can be an expression of volition if it is endorsed. More recently, in his reply to Susan Wolf, he addresses worries Wolf raises about connections between wholeheartedness and “zealotry, fanaticism or...close-mindedness.” Frankfurt allows that “the paranoid and the cleanliness fanatic” may have lives that are “warped by irrational forces and for this reason are

¹⁶⁶ Herman, "Bootstrapping."

deficient and undesirable,” but he insists that “the fact that their lives are bad implies nothing concerning whether they are free.”¹⁶⁷ In other words, as long as a person’s life is organized around desires he wholeheartedly endorses Frankfurt wants to say that he is free. Frankfurt here seems to be forgetting about what he says earlier about things being important to me whether I am aware of them or not. Maybe there is a sense in which such a person is free, but insofar as lives organized around the desires in question tend not to be sustainable, and, in the end, self-defeating, there is also a sense in which such a person is not free. The threat of degenerate desires to autonomy is best understood over time, diachronically, and he does bring that in later with “care,” but Frankfurt is never very explicit about this dimension.

¹⁶⁷ Frankfurt, "Reply to Susan Wolf," p. 251.

AUTHORITY

[In 'Autonomy, Necessity, and Love'] I distinguished two modes of love – active and passive – and I argued that being ruled by love counts as autonomous only insofar as the love that rules is active rather than passive. I should not have made this a condition of autonomy.

Following the dictates of love does not constitute acting autonomously because the love is active rather than passive. It constitutes acting autonomously simply because love is essential to the nature of the self.¹⁶⁸

Is the sense of authority Frankfurt is working with just a discovery and/or acceptance of the inevitable, of what I cannot control? His reply to Lear in the second quoted passage suggests so. Frankfurt cites Spinoza and pointedly instructs Lear that Freud (Lear's hero) is in the same camp. Frankfurt suggests in reply to Lear and elsewhere that morality, which he associates with a theoretical necessity, might be an independent source of authority, but it seems to me that when he does this, he's sketching a possibility about which he's personally agnostic. In his reply to Herman, he says, "there really is no authority for us other than the authority of what we care about."¹⁶⁹ For Frankfurt it seems that recognizing authority is closely related, if not identical, to recognizing the limits of our power, which is closely related to, if not the same thing as, recognizing "reality."¹⁷⁰

¹⁶⁸ Frankfurt, "Reply to Jonathan Lear," pp. 293-4.

¹⁶⁹ Frankfurt, "Reply to Barbara Herman," p. 276.

¹⁷⁰ Frankfurt, "The Faintest Passion," p. 100.

VI. CONCLUSION

I began this essay by trying to understand what role identification may play in an account of autonomy. I was attracted by the possibility of understanding autonomy in a way that is not quite so “top down” as autonomy is typically understood in traditional theories. That interest led me to focus on Frankfurt’s theory, since his account of identification’s role in autonomy has been the most influential. As I looked at Frankfurt’s theory more closely, however, I realized that his views have been shifting, and that, at least until quite recently, that fact has been generally underappreciated in the literature. I decided that it would be useful for me to try to “pin Frankfurt down” to the extent that I was able. What I discovered is that both the nature of identification and the role identification plays in Frankfurt’s account of autonomy shifts – partly by design, but not entirely, I think.

Frankfurt uses identification as a way of distinguishing mental states, desires in particular, that are internal to a person such that when a person is moved to action by them, it constitutes an exercise of volition, or autonomy. He begins with two senses of internality, and it is the narrower in which he is interested and for which identification is supposed to function as a criterion. The way I framed this earlier was by saying that the narrow sense picks out the “true” desires from merely “effective” desires – desires which in some broader sense belong to the person, even though when she acts on them, she’s not acting autonomously. Examples of desires that are sometimes *merely* effective include addictions and compulsions.

Frankfurt interprets identification, which is initially the criterion for narrow internality, first in terms of higher-order desiring, and then decision, and he speaks of

identification as “endorsement.” He later interprets identification in terms of wholeheartedness and satisfaction, and speaks of identification as, minimally, “acceptance,” disavowing his earlier “endorsement” language. But identification in the acceptance sense seems to track only the broader sense of internality, the one associated with free action, but not free willing, volition, and autonomy.

As identification as a criterion of the narrow internality – the internality that we associate with autonomy – moves out, caring and wholeheartedness move in. Caring, however, on Frankfurt’s account, is read off behavior just as effective desires are read off behavior, so there seems no way, on his account, to distinguish desires that constitute caring from desires that are merely effective. Wholeheartedness, as an interpretation of identification, is supposed to be an explication of the narrower standard, just as higher-order desiring and decision, earlier, had been. But wholeheartedness actually just presupposes the narrower standard: wholeheartedness is defined negatively as absence of ambivalence, and ambivalence Frankfurt stipulates, is “internal.”

In the end, Frankfurt doesn’t seem to have an account of what it is to act on a true or authentic desire that can clearly differentiate between acting on a desire that’s true or authentic from acting on a desire that’s merely effective. For instance, to act out of concern rather than mere desire seems on Frankfurt’s view to amount to acting upon a desire that’s stable over time. A distinction between stable and unstable motivations is important, but, by itself, can’t do all the work Frankfurt wants it to do. If I am right, it’s not clear how it can, for example, distinguish what we might consider a genuine concern from, in some cases, an unwanted addiction. To act

wholeheartedly, to act from a wholehearted desire rather than a desire that's merely effective, is, technically, to act on a desire that faces no "internal" opposition. This definition, though, just pushes the problem of identifying internal desires a step back. Alternatively, perhaps understanding wholehearted action as action that's reflective, decisive, and substantially satisfying may point to a way of discriminating among effective desire, of indicating desires that are effective in a deeper sense – desires that truly serve the agents ends, or serve the agents' true ends, for instance. However, I don't yet see how this can be made out without saying more than that the agent just has some final ends or other.

Frankfurt's account of the kind of agent we're interested in, the person, is an account of a creature that not only has, but consciously discovers some final ends or other – ends that are contingently determined, but, despite their contingency, not directly escapable. The only thing that is necessary is that the person be at least dimly aware of having some such end or other, and, I would add, that that end also *function* as an end – that it prove organizationally viable, relatively stable and compelling. Although Frankfurt hints that a person's biological needs will naturally play a role in determining which ends would so function, he wants to avoid saying anything substantive about these ends. In the end, though, in order to make the distinction he is after, he may have to.

What I'm driving at is that the true desire/effective desire distinction may depend upon a theoretical conception of persons or selves and their needs (and an acknowledgement that persons are creatures with such conceptions, however informal). Frankfurt does offer a theoretical conception of persons, but he also seems

to want to deny that such a conception could say anything interesting or substantive about needs, because in the case of *persons* (as opposed to humans), needs derive entirely from an individual's contingent (although binding) concern. However, even if we grant Frankfurt that the full truth of any theoretical conception about persons will depend upon the ends around which persons' attitudes and actions are in fact, or could in fact, be stably organized, and we grant that the facts about these ends cannot be determined a priori – from a bare concept of personhood, say – it's still the case that we can and do make some fairly reliable predictions, both about our own "true" ends and the "true" ends of others. What's more, "love," in Frankfurt's sense, actually requires us to do so. On Frankfurt's own account, loving our children entails our having *some* conception of their needs, which derives from some conception of their true ends. These conceptions may be fragmentary and wildly inaccurate, of course. That is the key point. These conceptions can be mistaken. Our self-conceptions, both individual and collective, may necessarily be works in progress, but it seems to me that they play a critical role in the distinctions among our effective desires that Frankfurt is trying to make. Frankfurt begins to get at this with his various versions of identification, and he does, in passing, allow that we can be mistaken about what we love, but I don't think he pursues this point far enough. Perhaps we "identify" in the relevant way with desires that move us in action, in part by means of fallible *beliefs* about which of these desires (or ends toward which they are directed) are "true." Alternatively, perhaps an identifiatory "accepting awareness" of our effective desires causes them to change on their own in reliable and interesting ways. There are many possibilities. Frankfurt, however, keeps his

investigation focused on the priority of desire. Too much emphasis on this priority, though, has a tendency to undermine the true desire/effective desire distinction that is central to his account of autonomy.

- Anscombe, G. E. M. *Intention*. London: Basil Blackwell, 1957. Reprint, Cambridge: Harvard University Press, 2000.
- Bratman, Michael E. *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge, UK: University of Cambridge, 1999.
- . "Hierarchy, Circularity, and Double Reduction." In *Contours of Agency*, edited by Sarah Buss and Lee Overton, 65-85, 2002.
- . "Identification, Decision, and Treating as a Reason." In *Faces of Intention: Selected Essays on Intention and Agency*, 185-206, 1999.
- Buss, Sarah, and Lee Overton, eds. *Contours of Agency: Essays on Themes from Harry Frankfurt*. Cambridge: The MIT Press, 2002.
- Dennett, Daniel. "Conditions of Personhood." In *The Identities of Persons*, edited by Amélie Rorty. Berkeley, CA: University of California Press, 1976.
- . *Elbow Room*. Cambridge, MA: The MIT Press, 1984.
- . "Mechanism and Responsibility." In *Free Will*, edited by Gary Watson, 150-73. Oxford, 1982.
- Dennett, Daniel C. *Freedom Evolves*. New York, NY: Viking Penguin, 2003.
- Duckles, Ian. "Kant's Reciprocity Thesis: Towards an Existentialist Understanding." 2003.
- Epictetus. *Epictetus: The Discourses as Reported by Arrian, the Manual, and the Fragments*. Translated by W. A. Oldfather. Edited by G. P. Goold. 2 vols. Vol. 1, *The Loeb Classical Library*. Cambridge, MA: Harvard University Press. Reprint, 1989.
- Frankfurt, Harry G. "Autonomy, Necessity, and Love." In *Necessity, Volition, and Love*, 129-41, 1999.
- . "The Faintest Passion." In *Necessity, Volition, and Love*, 95-107, 1999.
- . "Freedom of the Will and the Concept of a Person." In *The Importance of What We Care About*, 11-25, 1988.
- . "Identification and Externality." In *The Importance of What We Care About*, 58-68, 1988.

- . "Identification and Wholeheartedness." In *The Importance of What We Care About*, 159-76, 1988.
- . *The Importance of What We Care About: Philosophical Essays*. Cambridge: Cambridge University Press, 1988.
- . *Necessity, Volition, and Love*. Cambridge: Cambridge University Press, 1999.
- . "On Caring." In *Necessity, Volition, and Love*, 155-80, 1999.
- . "On the Necessity of Ideals." In *Necessity, Volition, and Love*, 108-16, 1999.
- . "On the Usefulness of Final Ends." In *Necessity, Volition, and Love*, 82-94, 1999.
- . *The Reasons of Love*. Princeton: Princeton University Press, 2004.
- . "Reply to Barbara Herman." In *Contours of Agency*, edited by Sarah Buss and Lee Overton, 275-8, 2002.
- . "Reply to Eleonore Stump." In *Contours of Agency*, edited by Sarah Buss and Lee Overton, 61-4, 2002.
- . "Reply to Gary Watson." In *Contours of Agency*, edited by Sarah Buss and Lee Overton, 160-4, 2002.
- . "Reply to Jonathan Lear." In *Contours of Agency*, edited by Sarah Buss and Lee Overton, 293-8, 2002.
- . "Reply to Michael E. Bratman." In *Contours of Agency*, edited by Sarah Buss and Lee Overton, 86-90, 2002.
- . "Reply to Richard Moran." In *Contours of Agency*, edited by Sarah Buss and Lee Overton, 218-26, 2002.
- . "Reply to Susan Wolf." In *Contours of Agency*, edited by Sarah Buss and Lee Overton, 245-52, 2002.
- . "Reply to T. M. Scanlon." In *Contours of Agency*, edited by Sarah Buss and Lee Overton, 184-8, 2002.
- . "Three Concepts of Free Action." In *The Importance of What We Care About*, 47-57, 1988.

- Herman, Barbara. "Bootstrapping." In *Contours of Agency*, edited by Sarah Buss and Lee Overton, 253-74, 2002.
- Kant (1724-1804), Immanuel. *Ethical Philosophy: The Complete Texts of Grounding for the Metaphysics of Morals and Metaphysical Principles of Virtue*. Translated by James Ellington. Indianapolis: Hackett, 1983.
- Korsgaard, Christine M. "Personal Identity and the Unity of Agency: A Kantian Response to Parfit." *Philosophy and Public Affairs* 18 (1989): 101-32.
- . *The Sources of Normativity*. Cambridge: Cambridge University Press, 1996.
- Locke, John. *An Essay Concerning Human Understanding*. new abridged ed. London: Everyman, [1690] 1993.
- Martin, Raymond. *Self Concern: An Experiential Approach to What Matters in Survival*. Cambridge, UK: Cambridge University Press, 1998.
- Metzinger, Thomas. *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Edited by Thomas Metzinger. Cambridge, MA: The MIT Press, 2000.
- Palmer, Parker. *Let Your Life Speak*. San Francisco, CA: Joey-Bass, Inc., 2000.
- Quartz, Stephen R., and Terrence J. Sejnowski. *Liars, Lovers, and Heros: What the New Brain Science Reveals About How We Become Who We Are*. New York: HarperCollins, 2002.
- Rachlin, Howard. "Complex Ambivalence." In *The Science of Self Control*, 57-81. Cambridge: Harvard University Press, 2000.
- Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.
- Raz, Joseph. "We Are Ourselves: The Active and the Passive." In *Engaging Reason: On the Theory of Value and Action*: Oxford, 2000.
- Rovane, Carol. *The Bounds of Agency*. Princeton: Princeton University Press, 1998.
- Strawson, P. F. *Individuals*. London: Methuen, 1959.
- Velleman, J. David. "Identification and Identity." In *Contours of Agency*, edited by Sarah Buss and Lee Overton, 91-123, 2002.
- . "Self to Self." *Philosophical Review* 105 (1996): 39-76.
- Watson, Gary. "Free Action and Free Will." *Mind* (1987): 145-72.

- . "Free Agency." In *Free Will*, edited by Gary Watson, 96-110, 1982.
- . *Free Will*. Edited by Gary Watson. Second ed. Oxford, UK: Oxford University Press, 2003.
- . "Kant on Happiness in the Moral Life." *Philosophy Research Archives* 9 (1984): 79-108.
- . "Volitional Necessities." In *Contours of Agency*, edited by Sarah Buss and Lee Overton, 129-59, 2002.
- Williams, Bernard. "Persons, Character and Morality." In *Moral Luck*, 1-19. Cambridge: Cambridge University Press, 1981.
- Yaffe, Gideon. *Liberty Worth the Name: Locke on Free Agency*. Princeton, NJ: Princeton University Press, 2000.