REFLECTIVITY, CARING AND AGENCY

By
Zoltan Wagner

Submitted to
Central European University
Department of Philosophy

In partial fulfillment of the requirements for the degree of Doctor of Philosophy in Philosophy

Supervisor: Professor Ferenc Huoranszki

Budapest, Hungary
2011

Undersigned declares that this dissertation contains no material accepted for any other degrees in any other institution

_____ March 3, 2011

Undersigned declares that this dissertation contains no material previously written and/or published by another person, except where appropriate acknowledgment has been made in the form of bibliographic reference.

_____ March 3, 2011

Abstract

As a solution for the problem of free will and determinism, Harry G. Frankfurt developed his hierarchical theory of desires. This controversial theory led to fruitful debates not only in connection with the problem of free will but in other areas as well including personal and practical identity, rationality in action and the problem of agency in general. Thus, this dissertation is not focused on the problem of free will but on the role of second-order desires in agency. More particularly, it discusses the importance of two features of human agency that Frankfurt emphasized most. First, humans are reflective, that is, they can evaluate their own desires. Second, humans are capable of caring about things, that is, some things are important for them in a special way. Frankfurt understands both of these crucial abilities as based on our ability to have second-order desires. In agreement with Frankfurt, I will argue that both reflectivity and caring are very important and basic abilities, and that they are connected to each other. However, I will disagree about their relation and about the definition of caring as dependent on second-order desires. I will argue that caring in the most basic sense is not a reflective ability but a first-order evaluation which is centrally important to action. I will also try to show that the reason humans have second-order desires is that they care about things. Thus, our second-order desires should be based on our personal values; in other words, what we care about. As a consequence, the most important claim of this dissertation is the following: though humans are reflective creatures, it is not reflectivity that is the most important feature of human life. Rather, caring, as a first-order evaluation is more important and our ability to reflect about our motivation depends on being able to care about things in this sense. At the end of the dissertation I will briefly discuss a serious problem for any caring based theory of action: the problem of self-control and self-management. Since we do not directly control by decisions what we care about, it seems that we cannot have these abilities. However, I will try to sketch a method of self-control and self-management which is consistent with the importance of the redefined concept of caring and which is based on self-trust and openness. This method will add one more argument against the importance of reflectivity: second-order desires do not help us in managing our desires; rather, they can interfere with self-control and become self-defeating strategies.

Acknowledgements

First of all, I would like to thank my supervisor, Professor Ferenc Huoranszki, who helped me clarify my often confused ideas. I would not have been able to complete this dissertation without his guidance.

I would also like to thank Gray Watson who helped me a lot in understanding basic questions of the philosophy of action and Harry G. Frankfurt's philosophy. The three-month research period I spent at the University of California in Riverside (2006/2007 Fall) – funded by the Doctoral Support Research Grant of Central European University – was an invaluable contribution to this dissertation.

I am also very grateful to the Institute of Philosophy of the Slovak Academy of Sciences and Martin Muransky and Tibor Pichler for giving me an opportunity to do research, give talks and publish papers in the Slovak Republic. My research in Bratislava was funded by the Slovak Mobility Fellowship (2008/2009 Fall) and the Visegrad Mobility Fellowship (2010/2011 Fall).

And last, but of course, by no means the least, I would like to thank Harry G. Frankfurt for reading an earlier version of *Chapter 6* and sending me detailed comments. Also, this dissertation would never have been written without the inspiration of his ideas.

TABLE OF CONTENTS

Introduction

This dissertation was first motivated by my interest in the problem of free will and determinism. Among the diverse suggestions to solve this problem I was particularly attracted to one: Harry G. Frankfurt's theory of free will built on his hierarchical theory of desires. Besides trying to explain the concept of free will with the help of this view, he claimed that we are persons and humans by virtue of our ability to have *second-order desires*. But what is exactly a second-order desire? When one has a desire like that, one wants to have or not to have a certain desire.[1] Frankfurt's claim about free will was that when one acts in accordance with one's second-order desires one has freedom of the will, but when one acts on a contrary desire one has not. But apart from their significance for freedom, second-order desires reveal something very important about our attitude to our own desires and wants. They are the ultimate mechanism through which our ability to evaluate ourselves, to use Frankfurt's term, *reflective self-evaluation* works.

Thus, the hierarchical theory of desires is not simply a theory about free will but a more general theory about human reflectivity and its role in agency. As a result, besides free

[1] As we will see, Frankfurt uses the term "to desire" in a broad sense, interchangeably with "to want".

will, determinism and moral responsibility, Frankfurt's ideas reinterpret a range of other topics including the problem of action in general, rationality, practical reason, practical and personal identity. Frankfurt's controversial views led to a lively debate in the course of which he introduced new concepts into the discussion such as commitment, identification, satisfaction, wholeheartedness, and most lately, caring and love. Obviously, as controversial as it is, Frankfurt's concept of second-order desires has proved to be very heuristic and led to interesting claims in all the areas listed above.

I found two claims particularly interesting: first, Frankfurt thinks that we are agents and persons because we have second-order desires. We evaluate our own selves through second-order desires which make active participation in the process of decision-making possible. A creature without them is a *wanton*, someone whom we cannot regard as a person. Second, his recent views imply that we are agents because we are able to *care* about things. Human beings are special because they care about things: they find some things in their life important, some persons and things matter them in a special way. I will argue in my dissertation that the two abilities – reflective self-evaluation and caring – are closely tied together. However, I will disagree with Frankfurt about the way they are connected. For Frankfurt, the ability to care presupposes reflectivity. I will try to show that the relationship should be understood in a quite different way: our second-order desires should be based on what we care about, that is, it is reflectivity that should be based on caring. In what follows I will give a brief summary of the key ideas of the dissertation and I will also give a short overview of each chapter.

When I talk about second-order desires, I will assume that humans do have this kind of reflectivity about their motivation. Of course, it is possible to doubt that there are second-order desires. One might simply say that there is no such thing as "to want to want to do

something" or "not to want to want to do something". Sure, it might be awkward to put it this way. But everybody is familiar with situations in which we become aware of our desires and our attitude towards them. In such cases we might think: "I wish I wanted this more" or "I wish I did not want to do this so much". For example, when faced with the fact that he has not made enough progress with his work, a student might wish that he wanted to focus on his work more and that he did not want to waste his time on the things that distract him, say lazing around in bed or going to the movies. Or to take a typical case in which second-order desires might appear, an addict who has a strong desire to take the drug might say to himself: "I wish I did not want to take the drug so badly". These are typical instances of having the type of reflective concern about our own motivation that Frankfurt talks about. No doubt, one can describe these scenarios with a vocabulary different from that of the hierarchical theory of desires. For example, one might say that the student judges that it would be better to work more instead of wasting his time and that the addict judges that it would be better not to take the drug and stop his self-destructive way of life.[2]

I will not try to prove that we get closer to the truth if we describe such situations with the vocabulary of second-order desires rather than practical judgments or our belief in objective values. What I will try to do is to describe certain problems and complexities of human agency with the help of the vocabulary of second-order desires and caring. I will work under an assumption, Humean in nature: human beings are not rational in the sense that they are not moved by pure reason. Hume believed that "reason alone can never be a motive to any action of the will."[3] As it is well-known, he also declared that "[R]eason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve

[2] A similar view was developed by Gary Watson. I will discuss it in *chapter 1*. See Watson, Gary, "Free Agency", in *Agency and Answerability* (Oxford: Claredon Press, 2004), 13-32, reprinted from *Journal of Philosophy* 72 (1975), 205-20.
[3] Hume, David, *A Treatise of Human Nature* (Oxford: Oxford University Press 2000), 265.

and obey them".[4] But there is an important difference between Hume and Frankfurt: the latter does not talk about the rule of the passions. Rather, he stresses that some of our desires are reflective, that is, they belong to the second-order and that they are needed for decision-making which is consequently not a rational process. As we will see later he on he developed a controversial view on practical reason built on his concept of caring and love. In his theory of human nature, the role of reason is merely instrumental, thus actions are not the result of reasoning and rational judgments.[5]

Though reflectivity is an important human capacity, I will argue that our second-order desires have to be based on something. In agreement with Frankfurt, I will argue that there is a particular ability which is very important for agency: the ability to care about things. That we are able to care about things means that we find some things particularly important to us: they matter to us; we evaluate them in a personal way. I will argue that it is caring on which our second-order desires should rest. What we care about is crucially important to how we make our decisions and how act: it is centrally important to establishing priorities and organizing our agency. But though I accept the central importance of caring, I will criticize Frankfurt's concept of caring and try to develop a different one. According to Frankfurt, caring is a capacity that requires reflectivity in the sense of second-order desires. However, this does not seem to be quite correct. Though there is a reflective sense of caring, there is a more important, more "primitive" meaning of the term as well, which simply means a positive evaluative attitude towards the object of care.[6] Thus, caring is primarily a first-order phenomenon, which, given its importance for agency, has a tendency to generate second-order desires.

[4] Ibid., 266.
[5] Nomy Arpaly developed an important theory of agency in which she also questions the importance of rationality in agency. See Arpaly, Nomy, "On Acting Rationally Against One's Best Judgment", *Ethics* 110 (April 2000), 488-513 and also her book, *Unprincipled Virtue: An Inquiry into Human Agency* (Oxford University Press, 2002).
[6] For a recent similar concept of caring and criticism of Frankfurt, see: Jaworska, Aginieszka, "Caring and Internality", *Philosophy and Phenomenological Research* 74, No. 3 (May 2007), 529-568.

The redefined concept of caring implies that if we stress the significance of caring, reflectivity becomes of secondary importance. What makes us agents and human beings is not only our ability to reflect on our desires but our ability to care about things, to value them in a particular way. Another important consequence of a caring based theory is the shift from a voluntary to a non-voluntary element in our psychology.[7] We do not directly control what we care about and this raises difficult problems in connection with self-control and self-management which I will discuss towards the end of the dissertation.

Before I begin, let me give a brief overview of the chapters. The first part of my dissertation consists of the first two chapters and it discusses Frankfurt's classic hierarchical theory of desires. In *chapter 1* I will discuss this theory and the concept of free will based on it. I will consider some of the objections directed against this view, most importantly Gary Watson's criticism. I will argue that Watson's criticism is right inasmuch as we need a certain basis for our second-order desires.

In *chapter 2* I will look at in detail how Frankfurt developed his views further. To answer his critics, he introduced new concepts, including commitment, identification and satisfaction. These concepts lead to an elaborate and sophisticated view of reflective self-evaluation. But I will argue that though these concepts explain a lot of complexities, in the end they cannot escape the objection discussed in *chapter 1*: they seem to have no firm basis.

In the second part of my dissertation I will discuss Frankfurt's recent views based on the concept of caring. In *chapter 3* I will discuss some general points about the problem of caring. Since this concept is used in several other contemporary theories, I will differentiate Frankfurt's concept from them. I will show that the significance of caring is that it can provide a basis for our second-order desires. I will discuss and criticize in detail Frankfurt's concept of caring and its significance for reflective self-evaluation, free will and autonomy.

[7] Stefan Cuypers discussed this shift in detail in a recent paper: Cuypers, Stefan E, "Autonomy beyond Voluntarism: In Defense of Hierarchy", *Canadian Journal of Philosophy* 30, No. 2 (June 2000), 225-256.

After discussing the concept of caring in detail, in *chapter 4* I will look at what Frankfurt thinks to be the most important mode of caring: love. As I will show, Frankfurt's concept of love is too wide, and a lot of claims he has in connection with love might be said to refer to caring in general. Several important problems will be discussed in the chapter including practical reason, subjectivity about values and reasons for action.

In the last part of my dissertation I will try to develop a revised theory of agency on the basis of a revised concept of caring and reflectivity. In *chapter 5* I will differentiate between *deep caring* and *reflective caring* and I will redefine the concept of caring as a multi-order phenomenon which involves both first-, and second-order desires. I will argue that deep, first-order caring is more important, and that our reflectivity is based on it.

After discussing the redefined concept of caring, I will move to the concept of love in *chapter 6*. I will especially focus one of Frankfurt's most controversial claims according to which love liberates us. Thus, the problem of freedom comes back to the picture, though not in the sense of free will. Rather, I will differentiate between freedom as self-expression and freedom as self-transcendence and claim that Frankfurt's view belongs to the first category. I will criticize Frankfurt's view and show that love liberates us by making self-transcendence possible.

In *chapter 7* I will consider a very important problem of any view that treats caring as centrally important for agency. We do not control directly what we care about and at first this leads to skepticism about self-control and self-management. I will briefly sketch a theory of self-management which – instead of Frankfurtian reflectivity – is built on a type of self-trust: on an ability to form beliefs of a certain kind.

This sketch already reveals the most basic claim of my dissertation: though humans are importantly reflective agents as Frankfurt argues, it is not reflectivity in itself what makes us agents and human beings. Because of his complex understanding of reflectivity and caring

Frankfurt fails to provide a Humean picture of agency. As Jaworska notes, Frankfurt attempted "to delineate the organization of agency peculiar to human beings, while avoiding the traditional overintellectualized emphasis on the human capacity to reason about action".[8] Nevertheless, as she adds, in the end Frankfurt himself overintellectualized human agency. At least, his complex picture of reflectivity, and his concept of caring, as requiring reflectivity lead to a complex structure that seems to be too complicated to accommodate average human agency. And this is the significance of the redefined concept of caring: caring is not a sophisticated reflective phenomenon but a very basic feature of human beings. Thus, the revised theory of caring is more in vein with the original Humean assumption about action and deliberation.

[8] Jaworska, "Caring and Internality", 566-67.

Part I: The Classical Theory

Chapter 1: Frankfurt's classical theory

1.1. The theory of free will based on the hierarchical theory of desires

Frankfurt's seminal paper "Freedom of the Will and the Concept of a Person"[9] offers an original solution to the free will problem and at the same time it contains the basis of his philosophy of action and his theory of practical identity. In this paper he introduced the *hierarchical theory of desires* which is at the center of thinking about human psychology and action.

Before starting to analyze the theory it is necessary to point out two important features of this view that make it rather unusual as compared to mainstream views of free will and moral responsibility. First, Frankfurt's theory of free will is *neutral* with regard to the problem of determinism[10]. It is mainly concerned with the psychology and not the metaphysics of free will. Second, Frankfurtian free will is *not* a condition of moral responsibility.

[9]Frankfurt, Harry G., "Freedom of the Will and the Concept of a Person", in Frankfurt, Harry G., *The Importance of What We Care About*, (Cambridge: Cambridge University Press, 1988), 11-26, reprinted from *Journal of Philosophy* 98, no. 23 (January 1971), 5-20.
[10] Ibid., 25.

To look at the first characteristic, philosophers often treat the problem of free will as a *metaphysical* problem. On the one hand, incompatibilists claim that if determinism is true, one cannot possess free will. According to well know incompatibilist arguments as the *consequence argument*,[11] laws of nature and past events jointly make free will, as an ability to do otherwise, impossible. On the other hand, compatibilists claim that one can be free even in a deterministic world, as long as one's actions are determined in the right way. However, though it might be labeled as a compatibilist, Frankfurt's hierarchical theory is not concerned with metaphysics and he points out that the conception of freedom of the will based on it is neutral with regard to the problem of determinism. That is, he claims that it is *conceivable* that one is casually determined and is yet free *and* that it is also conceivable that one becomes free by chance. What is more, he thinks that it is conceivable that the state of affairs he describes as freedom of the will comes about in some third way, or "a way other than by chance or as the outcome of a sequence of natural causes".[12] We can call the conjunction of these claims the *conceivability claim* according to which free will as a psychological phenomenon can conceivably exist in a variety of deterministic and nondeterministic worlds.[13]

Though this claim makes Frankfurt a compatibilist, we have to keep in mind that the main issue here is not the metaphysics, but the psychology of freedom. As regards the metaphysical debate then, this view accepts the conceivability claim only as an assumption, but the main goal is not to defend it but to describe the psychological conditions of freedom. Defending the conceivability claim as a metaphysical claim would be a different matter, and

[11] See for example Inwagen, Peter van, *An Essay on Free Will*, (Oxford: Claredon Press, 1983).
[12] Frankfurt, "Freedom of the Will", 25.
[13] See also David W. Shoemaker, who has developed a Frankfurtian view on freedom: "Whereas the classical compatibilist claimed that freedom is not only compatible with determinism but requires it, if freedom instead is simply a matter of one's will embodying a certain structure, than freedom is actually compatible with determinism or *indeterminism*. Whether one comes to have a properly structured will through deterministic or indeterministic means becomes irrelevant." Shoemaker, David W., "Caring, Identification and Agency", *Ethics* 114, No. 1 (October 2002), 116 (italics in the original).

if one were to do that, one would count as some kind of "hard compatibilist".[14] But since this primary focus of this thesis is not the problem of free will and determinism but to defend a hierarchical account of agency I will not discuss this question.

For the same reason I will also not discuss the so-called 'Frankfurt type of counterexamples' which try to demonstrate that the Principle of Alternate Possibilities (PAP) is false. In his paper "Alternate Possibilities and Moral Responsibility" Frankfurt defined PAP the following way: "[A] person is morally responsible for what he has done only if he could have done otherwise".[15] He tried to show that this principle is false by presenting cases in which the agent could not have done otherwise and yet can be regarded as having free will and morally responsible. There is an extensive and important literature on this topic, and it is still at the center of the debate on free will and determinism. However, since I am not primarily concerned with metaphysics of free will but with problems of agency in general, I will not discuss this issue.

The second difference between Frankfurt's view on free will and mainstream theories is that though philosophers generally treat free will as the primary condition of moral responsibility, Frankfurt states that "it is not true that a person is morally responsible for what he has done only if his will was free when he did it".[16] He introduces the case of the *willing addict*, whose will is not free but nevertheless might be responsible for taking the drug.

[14] I take this expression from Gary Watson who thinks that the conditions of free and responsible agency are non-historic, that is, it does not matter how they come about. This even makes it possible that one becomes free and responsible as a result of design by a "devil/neurologist". See Gary Watson: "Soft Libertarianism, Hard Compatibilism", in Watson, *Agency and Answerability*, 197-215, originally *Journal of* Ethics 3 No. 4 (1999) 351-65. Frankfurt also accepts this possibility in "Three Concepts of Free Action", in Frankfurt, *The Importance of What We Care About*, 47-57, reprinted from *Proceedings of the Aristotelian Society* 49 (1975), supplementary volume, 95-125.

[15] Frankfurt Harry G. "Alternate Possibilities and Moral Responsibility" in *The Importance of What We Care About*, 1-11, reprinted from *Journal of Philosophy* 96 No. 23. (December 1969), 829-839.

[16] Frankfurt, "Freedom of the Will", 23-24. Frankfurt discussed the problem of responsibility in another paper, "What We Are Morally Responsible for", in *The Importance of What We Care About*, 95-103, reprinted from *How Many Questions? Essays in honor of Sidney Morgenbesser*, ed. L.S. Cauman, Isaac Levi, Charles D. Parsons, Robert Schwarz (Indianapolis: Hackett, 1983).

It could be argued that Frankfurt's theory discusses *ideal agency*.[17] Since not many people can achieve an ideal, it would obviously be too strong as a condition of moral responsibility. Thus, Frankfurtian free will is *not* a condition of moral responsibility. Of course, Frankfurt believes that an agent satisfying the conditions that he talks about is morally responsible, but it is not necessary to reach this ideal in order to be justly held morally responsible. This is connected to Frankfurt's thesis that freedom comes in degrees. As he puts it in a later paper, "freedom is necessarily susceptible to variations in degree".[18] This is opposed to classic views on the subject, according to which one either possesses freedom or one does not. Frankfurt is not alone in claiming that freedom is a quantitative concept. This view is accepted by other notable thinkers including Gary Watson. For him, freedom depends on the degree one's motivational system overlaps with one's evaluational system, or in other words on the degree one's motivation and actions are dependent upon one's values. He claims that only God is totally free in this sense; human beings are "only more or less free agents, typically less."[19] Watson's point shows that freedom in this sense cannot be the condition of moral responsibility. It is rather like an ideal that we cannot reach. It is not required to reach the highest degree of freedom in order to be morally responsible for an act.

After these preliminary remarks, let us look at Frankfurt's early theory of free will in detail. As the title of his paper suggests, he believes that being a person and having free will is closely tied together. This means that the criteria of being a person and that of having free

[17] Some philosophers of action recently used other expressions with a similar meaning, including "agency par excellence" and "full-blooded action". These require one to fully exercise his capacities to actively participate in his agency. See for example Velleman: "What Happens when Someone Acts?" in Velleman, David J. *The Possibility of Practical Reason*, (Oxford: Oxford University Press, 2000) 123-143, and Mele, Alfred R. *Motivation and Agency*, (Oxford: Oxford University Press, 2003). See also Gideon Yaffe, who thinks that ideal agency or "agency at its best" is the basis of free will. Yaffe, Gideon, "Free Will and Agency at its Best", *Nous* 34, No. 4, Supplement 14, 2000.

[18] Frankfurt, "Concerning the Freedom and the Limits of the Will", in *Necessity, Volition and Love* (Cambridge: Cambridge University Press, 1999), 76, reprinted from *Proceedings and Addresses of the American Philosophical Association* 59, No. 2 (November 1985).

[19] Watson, "Free Agency", In *Agency and Answerability* (New York: Oxford University Press, 2004), 31. Originally *Journal of Philosophy* 72 (1975).

will are also closely connected. The main feature of persons and free agents lies in the structure of their will. Average human beings are persons "by default" and they are capable of achieving free will because, as Frankfurt puts it, they are capable of forming "second-order desires" or "desires of the second order".[20] Our capability to form such desires is the basis of our capacity for *reflective self-evaluation*.[21]

Now, what makes these desires so special that they are a basis of an important capacity? The primary difference is that while first order motivation is directed at possible courses of action, second-order desires are directed at desires themselves. As Frankfurt puts it, "besides wanting and choosing and being moved *to do* this or that, men may also want to have (or not to have) certain desires and motives".[22]

It is important to note here that the verbs "to want" and "to desire" are used interchangeably in Frankfurt's writings.[23] Of course, one could object to this usage that it makes sense to say that though I desire to smoke, I do not actually want it because I want to lead a healthy life. In this case, "to want" and "to desire" seem to be different, indicated by the fact that one can fail to want what one desires and *vice versa*. However, for the sake of simplicity, Frankfurt suggests the following usage: "A wants to X may mean to convey that it is his desire that is motivating or moving A to do what he is actually doing or that A will in fact be moved by this desire (unless he changes his mind) when he acts.[24] Whether I smoke a cigarette or I go jogging, in the end I do what I do because of a desire that motivates me; I do what I do because I want to do it. Of course, the motivation that in the end moves me still might conflict with some of my other desires but that does not mean that the verbs "to desire" and "to want" pick different kinds of phenomena.

[20] Frankfurt, "Freedom of the Will", 12
[21] Ibid.
[22] Ibid., Italics in the original.
[23] Ibid., 12-13, footnote 2.
[24] Ibid., 14.

Someone might still say that treating the two terms as synonyms is very unhelpful since I might desire something very strongly but nevertheless fail to want it, and vice versa. For example I might desire smoking a cigarette very strongly, but it does not imply that I want to smoke a cigarette. Actually, I do not want to do it because I want to quit. Frankfurt would say the following about this case: I am in a conflict about smoking; I desire or want both to smoke and not to smoke. That is, I have competing desires, and I have not yet committed myself to one of them on a higher, reflective level. One might still say that we can differentiate between two types of cases. In the first case, though I want to quit for the reason that I want to live a healthy life, I also want to smoke, because, as it is expressed in Jack London's autobiography, *John Barleycorn*, I think that a life without smoking, drinking and adventures is not a man's life. In this case there is a real conflict in which it is not yet clear which life I want to commit myself to. In the other I want to quit smoking because I am committed to a healthy life, but I still have a strong desire to smoke without actually wanting to smoke. One might argue that in this second case the contrast between what I desire and what I want is very important because there is a conflict between what I desire and what I want, and it is inappropriate to say that this is merely a conflict of desires. The fact that I do not want to smoke reflects a judgment and consequently my wanting not to smoke is more than a desire. Gary Watson made a similar point in 'Free Agency' when he claimed that it is not the hierarchy of desires what is important but evaluations which are prior to higher order desires and volitions.[25]

Watsonian evaluations have a distinct nature from desires but Frankfurt does not believe that our motivation has two distinct sources, one being rational and the other non-rational. Desires have often been regarded as not simply non-rational but rather, somehow by nature being opposed to reason. But Frankfurt's way of using the term suggests a much

[25] Watson, "Free Agency", 30.

broader meaning. Higher-order desires and volitions are the basis of the capacity of reflective self-evaluation and thus play an important role in his theory of agency. Consequently, for him, desires are not necessarily blind forces but can be "intelligent" in the sense that some of them guide our actions and express our authority about what to do.

The usage Frankfurt proposes also helps us identify the agent's "will". To put it simply, an agent's will is his *effective* desire, or in other words the desire that moves him all the way to action.[26] Since one's will is a desire that motivates one in action, one's will is identical with some of one's first-order desires. Our ability to form second order desires can have a special role because besides being able to want or not want to have certain desires, we are also capable of wanting a certain desire to move us all the way to action; that is we can want a certain desire to be our *will*. This very ability is the essence of being a person, and a creature without it is a "wanton".[27] A wanton might be a very clever and reasonable creature but without this ability he cannot count as a person. This shows that it is not reason but the will that makes us persons. Though being a person presupposes rationality as well,[28] it is the structure of the will which is more essential to being a person.

With this we get to a very essential claim of Frankfurtian philosophy of action. He claims that reason is of secondary importance. It is rather like a tool than the source of normativity; its role is only instrumental, in itself it does not move us.[29] This is in accordance with the Humean claim that pure reason does not move us, it is only instrumental. Frankfurt maintains that the crucial feature of agency lies in the structure of the will and not in reason itself. A second-order volition itself is an a*ct of will*. As we will see Frankfurt later introduced several other types of acts of will, including *commitment*, *endorsement* and *identification*. It is

[26] Ibid.
[27] Ibid., 16.
[28] Ibid., 17.
[29] For a detailed discussion of Frankfurt's view on practical reason see *The Reasons of Love* (Princeton: Princeton University Press, 2004). I will discuss it in more detain in *chapter* 4.

common in these that they are all manifestations of the capacity of reflective self-evaluation, but they are acts of will and not reason.

Frankfurt tries to illustrate the difference between a person and a wanton by differentiating two possible drug addicts. He assumes that they are both physiologically addicted but there is a crucial difference in their psychology, namely that while one of them, the unwilling addict hates the fact that he is addicted and struggles against it, the wanton addict lacks this kind of concern. The unwilling addict, besides having a conflict between his first-order desires (a desire to take the drug and a desire to abstain), also forms a higher-order volition by which he favors one of his desires (the desire to abstain) and wants it to be his will. Even though his second-order volition is ineffective, he has a concern about what should move him to act. In contrast, the wanton addict simply acts on his momentary desires; when he wants it, he takes the drug, when he does not, he abstains.

Now, it is the unwilling addict's concern with his motivation what makes him human and a person. He is not a "passive bystander"[30] with regard to his own desires but he actively takes sides in the conflict and tries to influence what moves him to act. It is the wanton's *indifference*, or lack of self-concern and his passivity that deprives him of his humanity and personhood. The crucial difference between a person and a wanton is that while a person is "moved by either the will he wants or a will he wants to be without", when wanton acts, neither is the case.[31] When a person acts, it is either *because* he wants to be moved by a certain desire or he acts on it *against* his will. And since a wanton cannot form higher-order volitions, neither of these can be the case.

As we have seen above, according to Frankfurt, personhood is based on a certain volitional condition, namely on the capacity to form second-order volitions. However, this capacity is also a condition of free will. As I mentioned earlier, the conditions of personhood

[30] Frankfurt
[31] Frankfurt, "Freedom of the Will", 18.

and free will are related in Frankfurt's thinking. To see the exact relation let me quote Frankfurt at length here:

> It is only because a person has volitions of the second order that he is capable of both enjoying and of lacking freedom of the will. The concept of a person is not only, then, the concept of a type of entity that has both first-order desires and volitions of the second order. It can also be construed as the concept of a type of entity for whom freedom of its will may be a problem. This concept excludes all wantons, both infrahuman and human, since they fail to satisfy an essential condition for the enjoyment of freedom of the will. And it excludes those suprahuman beings, if any, whose wills are necessarily free.[32]

The condition of personhood (the ability to form second-order desires and volitions) in itself does not make us free but it makes us capable of enjoying freedom. A profound point in this link might be that human action is essentially problematic, since the very capacity that makes us different from (and supposedly superior to) animals can also be the source of both freedom and the lack of it. Though personhood is given to us thanks to our reflective capacities, free will is not: it takes a lot of effort to achieve it. The possibility of becoming free can be a source of a host of problems. For example, as in the case of the unwilling addict, we might end up doing something against our will. If the addict is unable to execute his decision, this will lead to the loss of his free will. He might become weak willed or he might deceive himself about what he really wants most. Here we can note that the capacities that make free will possible open up their counterparts: all the negative states that make human life problematic. Therefore, the capacity for reflective self-evaluation can be used to explain many important complexities of human psychology that do not arise for simpler creatures. Of course, this suggests that the capacity is not a perfection in itself; indeed, in some cases it might be a curse rather than a blessing.

Having free will depends on our ability to adjust our motivation to our second-order volitions. Thus, it is different from free action, which is neither sufficient nor necessary for

[32] Ibid., 19.

having free will.[33] Being able to do freely as one wants is not enough for free will, and being deprived of the ability does not take away our free will since we can still form second-order volitions. However, Frankfurt suggests a definition of free will analogous to that of free action. Just as free action is "the freedom to do what one wants to do", enjoying free will means that the agent "is free to will what he wants to will or to have the will he wants".[34]

Now, it seems to be very already difficult to create harmony between one's first-order motivation and second-order volitions, but it might be even more difficult to form these volitions themselves. Besides the ones listed above, yet more psychological complexities arise here: one can deceive oneself what kind of second-order desires one really has, and one can have higher-order desires that are in conflict with each other. *Indifference* and *ambivalence* are also related states which I will discuss later on in detail.

As we have seen lots of complexities and difficulties might arise in connection with second-order volitions but there is a significant difference how these processes work in different persons. It is not necessary that one has to face all these complexities in one's volitional life; for some people the higher-order volitions arise spontaneously and not as a result of struggle and decision. As Frankfurt points out, "some people are naturally moved by kindness when they want to be kind"[35]; they do not need to exercise self-control or struggle against contrary desires. However, for some, life is not that simple. To quote Frankfurt, while "the enjoyment of freedom comes easily to some," others "must struggle to achieve it".[36] As I have mentioned, according to Frankfurt having or not having freedom is a matter of degree. Those, who have to struggle with the psychological complexities mentioned, might also have to face the loss of their freedom when overcome by them. Gary Watson also argues that freedom is a matter of degree and claims that "humans are only more or less free, typically

[33] Ibid., 20.
[34] Ibid.
[35] Ibid., 22.
[36] Ibid.

less"[37] which also suggests that a conflicted state in which freedom is diminished is more like the rule and not the exception. For some, freedom is a given but for most, it takes effort to achieve it.

Of course, this raises the question of luck. For if it true that for some freedom is simply given and some will not ever achieve it regardless of their efforts, being free is a matter of luck and it becomes unjust to blame or praise people for their actions. And even if someone becomes free due to his efforts, it still might be the case that the process is not entirely in his voluntary control. So whether freedom is given or the result of a process, it seems that it is, to some extent, a matter of luck whether one is free or not.

1.2. Watson's criticism

The most important criticism of Frankfurt's hierarchical theory is connected to the putative special status of second-order desires and volitions. Gary Watson captured this point in his paper "Free Agency" and developed an alternative view. Let us look at his criticism and see what he offers instead of Frankfurt's theory of freedom.

Watson's basic worry in connection with Frankfurt's theory is the alleged special status of second-order desires and volitions. Frankfurt's view seems to rest on the assumption that free will is somehow secured by this special status itself, but it is questionable that this is so. As we have seen, the basic difference between a person and a wanton is that while the latter is *indifferent* to what moves him to act, a person has a concern with his own motivation. But Watson argues that if indifference is the main point, it is not enough to assume the ability to form second-order desires and second-order volitions. For if one can be indifferent with regard to one's first-order desires, why could not he be indifferent with regard to one's higher-order desires as well? Let me quote Watson at length here:

[37] Watson, "Free Agency", 31.

> That the first desire is given a special status over the second is due to its having an *n*-order volition concerning it, whereas the second desire has at most an $(n - 1)$ order volition concerning it. But why does one necessarily care about one's higher order-volitions? Since second-order volitions are themselves simply desires, to add them to the context of conflict is just to increase the number of contenders; it is not to give a special place to any of those in contention. The agent may not care which of the second-order desires win out. The same possibility arises at each higher order.[38]

Watson thinks that it does not give any special status to a desire that it is of the higher order. For him, an action is free only if it is motivated by something that is *more than a mere desire*. Why is that so? As we have seen a wanton is somebody who is indifferent to what moves him to act. But Watson notes that one can be a wanton with regard to one's higher order volitions. Though Frankfurt defines a wanton as someone with no second-order volitions, it is still sensible to ask what makes an agent with these any better than a wanton. And we can ask this question because the point about wantonness is *indifference* to one's motives. Watson claims that having desires and volitions of the higher order does not necessarily imply that one is not indifferent in the same way a wanton is. One may have higher order desires in conflict and in the end one of them might win out if an agent forms a higher order volition, but one can still be indifferent which desire wins out and what higher order volitions one forms.[39] To sum this point: *indifference can go up to higher orders.*[40] Therefore, if Watson is right, hierarchy itself will not make us free.

Rather, he thinks that it will make us free only if this hierarchy is *grounded in* our evaluations.[41] Evaluations are *prior* to second-order volitions: we make a first-order

[38] Watson, "Free Agency", 28-29.
[39] Susan Wolf has a similar worry in connection with the hierarchical theory of action and she claims that an agent who is alienated from his first-order choice "can be alienated from her higher-order choices as well". See Susan Wolf: *Freedom Within Reason*, (Oxford University Press, 30). However, this matter is slightly different since she is talking about *choice*. It is really uncommon in everyday life to choose to choose something, and pushing the problem to a higher level does not seem to solve the problem of choice.
[40] Watson, in conversation.
[41] Eleonore Stump has a revised Frankfurtian account according to which the reason why second-order desires are important to freedom is not simply that they are of the second-order but "because the agent's second-order desires are the expressions of his intellect's reflection on his will". Stump, Eleonore, "Sanctification, Hardening of the Heart, and Frankfurt's Concept of Free Will", *The Journal of Philosophy* 85, No. 8 (August 1988), 408.

evalutional judgment about what to do and only *then* they generate second-order volitions.[42] In another paper, "Free Action and Free will" Watson puts this point in the following way: "when and only when higher-order volitions are so grounded are they any more than mere desires".[43] Evaluations have a special status because that can give *reasons* to oppose first-order desires.

I believe that the concept of second-order desires and volitions can be defended if we look at the two ways how one can have second-order desires and volitions. On the one hand, some people spontaneously want to be motivated by the very desires that move them all the way to action, for example some people spontaneously do something kind because they want to be kind. On the other hand, as we have seen, some others need active effort to do the same. Now, the spontaneous agent cannot be completely indifferent to what moves him to act, since by definition, he wants to be moved by kindness; he is not indifferent with regard to his motivation. But what about the case in which there is no spontaneous conformity between our first-order desires and second-order volitions? In case it is needed, second-order desires and volitions can have a special role; they make it possible for us to *actively* engage in our own agency.

Watson claims that it is possible to be totally indifferent with regard to our higher-order processes, but this assumption fails because it assumes that it is possible to be *both* active and indifferent. For me it seems that once somebody is actively doing something, at least he is less indifferent than if he was not doing it all. What would happen if someone was completely indifferent to what second-order desires and volitions he forms? In that state I

[42] In a recent paper Watson developed a different view on judgment and volition, emphasizing the role of the will in decision-making, as an ability independent from both beliefs and desires. See "The Work of the Will" in Watson, *Agency and Answerability*, 123-157, originally in Stroud, Sarah and Tappolet, Christine (eds.) *Weakness of the Will and Practical Irrationality*, (Oxford: Oxford University Press, 2003), 172-200. I discussed this view in detail in my paper "Practical Reason and the Work of the Will", *Croatian Journal of Philosophy* 19 (2007), 93-102. For a similar view, both as regards the skepticism about second-order agency and the role of the will in decision making see also Pink, Thomas, *The Psychology of Freedom* (Cambridge: Cambridge University Press, 1996), especially 33-64.

[43] Watson, "Free Action and Free Will", in Watson, *Agency and Answerability*, 167.

believe that one would simply cease to form any second-order desires and volitions. Since it is part of this state that I am not interested any more what moves me, once I reach complete indifference, I stop all reflective activity. Being indifferent in Frankfurt's sense means precisely that one does not have any higher-order concerns, consequently indifference cannot go up to the higher orders.

Watson claimed that it is possible to be indifferent to or alienated from our second-order processes. I have tried to show that it is not possible to have second-order processes and to be indifferent at the same time; and I believe that Watson himself can be defended from a similar kind of argument. As we have seen, Watson claims that our actions should rest on our own values, and even if we form second-order desires, they must be based on them. However, Velleman thinks that this suggestion does not help, for the same reason why Watson thought Frankfurt's theory does not work. He states that grounding our actions in our values does not solve the problem of action because one can be alienated from his own values, for example when "someone recoils from his own materialism or his own sense of sin."[44] But this argument does not seem to be convincing. For why would one recoil from one's values? If one does that, he has some reason to do so, and in that case his attitude to his materialism rests on some other values. But even if one could recoil from one's values without having any reason to do so, it would imply that he no longer has them. In the alienated state the agent simply does not believe in his values any more, he simply lost them. Let me quote here at length what Watson says on the subject:

> One's evaluatinal system may be said to constitute one's standpoint, the point of view from which one judges the world. The important feature of one's evaluational system is that one cannot coherently dissociate oneself from it *in its entirety*. For to dissociate oneself from the ends and principles that constitute one's evaluational system is to disclaim or repudiate them, and any ends and principles so disclaimed (self-deception aside) cease to be constitutive of one's valuational system. One can dissociate oneself from one set of ends and principles only from the standpoint of another such set that

[44] Velleman, "What Happens When Someone Acts", 134.

> one does not disclaim. In short, one cannot dissociate oneself from all normative judgments without forfeiting all standpoints and therewith one's identity as an agent.[45]

Though one can get alienated from some of one's values, one cannot get alienated from *all* of one's values. The same is true in connection with Frankfurtian higher-order reflectivity. One cannot be completely indifferent with regard to what second-order volitions one forms. Watson points out that once an agent totally dissociates himself from all of his values, they cease to be his values. But the same is true in connection with indifferent once somebody is completely indifferent to his own motivation, he ceases to form higher-order desires and volitions.

Of course, we can imagine examples that seem to show that Watson is right. For example a person, who is regularly motivated by unkindness, but who has both a higher-order desire to be kind and one to be unkind, might be completely indifferent whether he behaves kindly or unkindly with someone who has just suffered a loss. In the end, since he does not care what moves him, he makes an unkind remark while letting his second-order desire to be unkind move him. But has he really performed a second-order volition? It seems to me that he has simply let some of his desires move him without his own active participation in the process.

So it does seem to be the case that in themselves, second-order desires do not solve the problem. However, second-order volitions are different. Once one actively performs them, it robustly shows that one is *not* indifferent. But in a sense Watson still has an important point. And this is not the special status of second-order volitions, but an explanation of *why* an agent forms a particular second-order volition and not another. And at this point Watson seems to be right: there must be something on the basis of which we form our second-order volitions. In his view, it is our *valuational judgments* that we form them. If

[45] Watson, "Free Agency", 26.

he is right about this, it shows that after all, it is not our volitional faculties but our reason that matters, thereby refuting Frankfurt's Humeanism about action.

But showing that we should perform higher-order volitions on the basis on our values does not necessarily show that a Humean theory of action is false. Watson himself differentiated between *judging* something valuable and *valuing* it.[46] Having a more or less coherent valuational system based on our valuational judgments seems to involve an overly rationalistic picture of human agency. He admits in "Free Action and Free Will" that the theory presented in "Free Agency" is too rationalistic.[47] One can fail to desire that one judges to be valuable, and one can even fail to value it.

Though Watson still wishes to defend a non-Humean, or Platonic view on agency, I believe that his distinction between valuing and judging valuable shows that a rational judgment in itself is not enough and that valuing something is different from a rational judgment. In my dissertation I would like to show that Frankfurt's concept of caring can be understood as valuing. If sufficiently revised, it can answer to the criticism that our second-order desires and volitions should be based on something while making it possible to preserve the basic Humean claim according to which humans are not moved by pure reason.

In the next section I will look at another important objection to the hierarchical theory based on the problem of infinite regress.

1.3. The Problem of Regress

Another important problem of Frankfurt's classical theory is that it faces the danger of *infinite regress*. It seems that the theory is vulnerable to the infinite regress problem since if it is possible to form second-order desires and volitions with regard to a certain first-order desire, why could not it be possible to form yet higher-order desires on the third, fourth, etc.

[46] Watson, "Free Action and Free Will", 168.
[47] Ibid.

orders? Frankfurt does admit that this is possible to ascend to higher orders than the second. It might be possible or even makes sense, to try to re-evaluate our second-order attitudes on a yet higher level, for example, when one acts out of kindness because one wants to be kind (or in other word because one has a higher order desire that endorses the desire to be kind) one might still ask whether one really wants to be moved by this particular desire.

But it seems that there are certain limitations to the process of forming higher-order desires and volitions. I will argue that there are 1) *cognitive* and 2) *practical* limitations that prevent us to ascend to higher orders than maybe the third. These limitations make it increasingly difficult or impractical for us to ascend to higher orders.

First, to look at our *cognitive* limitations, it seems that on each level, it gets harder and harder to understand what it means to form a desire. For example we are able to understand the sentence that "I want to want to be kind", adding more clauses to the sentence makes it less and less possible to follow what the speaker means. We can see this easily if we follow adding clauses to the sentence: "I want to want to want to want to want…. to be kind" is not a sentence we can make sense of. To see that this is primarily a cognitive limitation we just have to look at other complex sentences containing several co-ordinations using the same verb. While we can make sense of the sentence: "I know that I know that *p*" or perhaps even "I know that I know that I know that *p*", we seem to loose track of what is being said when hearing a sentence like: "I know that I know that I know that I knows that I know that I know that *p*" simply because our cognitive limitations. There is another similarity between the verb "want" and "know". When I do not simply say that "I know that *p*" but "I know that I know *p*" my utterance might mean that I have a higher level of certainty about the truth of *p*. But after a while, adding more clauses to this sentence will confuse us until we completely lose sense of what it means to know something. The case is similar in case of the verb "to want". When I do not simply say that "I want to *a*" but "I want to want to *a*", it might imply that I

have a higher self-confidence in what I want.[48] But adding more clauses to the sentence would gradually undermine our confidence in what we want until the point we completely lose sense of what it means to want something. To conclude, we can understand sentences containing a variety of subordination and co-ordination even if they are very long, but as a result of our cognitive limitations, repeating the same type of clause even a few times make it difficult for us to understand a sentence. Thus it seems that infinite regress is not a serious problem for Frankfurt since we are simply unable to understand what it means to ascend to the sixth or the eleventh order.

But besides our cognitive limitations there are *practical* limitations as well that should prevent us from ascending to higher orders than the third. When facing a decision it would be very time-consuming to try to re-evaluate our motivation on yet higher and higher orders. Besides, not only that it would postpone arriving on a resolution, in this process we might simply get paralyzed, not knowing what desire to be moved by. When one asks oneself the question: "What kind of person do I want to be after all?" he might think about what kind of person he really wants himself to be, that is, ascend to the second-order, but going on yet higher levels would soon make him loose a sense of who he is. After a while he would not know which order expresses what he is and this would surely paralyze him. There might be people who engage in such kind of activities within the cognitive limitations that they have, but it is clear that it is very unpractical and after a point it is clearly pathological. And in case one forms a certain second-order desire why bother to question it? This raises the problem of self-trust, for someone who is always questioning his motivation on a higher order clearly does not trust himself, perhaps because he is seriously uncertain about himself. And this is a state that he should overcome by stopping the process and forming a second-order volition.

[48] The verb "to want" is slightly different than the verb "to know" though. If I say the sentence "I want to want to a" and I actually want to *a*, it might express a confidence in what I want, but the same sentence makes sense also in case I do not want to *a*, in which case it does not seem to imply self-confidence; it expresses our wish that our desires were different. This problem will be very important in connection with the problem of self-control and self-management, which I will discuss in the last chapter.

But how can we prevent ourselves from ascending to higher levels? Let me quote Frankfurt at length here to see what Frankfurt say on this problem. Talking about the process of ascending to higher and higher orders, he claims that

> It is possible, however, to terminate such series of acts without arbitrarily cutting it off. When a person identifies himself decisively with one of his first-order desires, this commitment "resounds" throughout a potentially endless array of higher orders.[49]

Here Frankfurt introduced two new important concepts: identification and commitment. Thus, he tries to solve the problem if interminable ascent by introducing new ideas into his theory. At first is seems that this move questions his original thesis according to which free will depends on our ability to make our first order-desires and second-order desires conform to each other. If identification and commitment as different abilities do the important work, it seems that the original stress on second-order volitions has lost its significance. However, as we have seen, second-order volitions are *acts of will*, and the new terms introduced refer to special types of acts of will themselves. In the next chapter I will analyze Frankfurt's attempt to answer to criticisms of his theory by making it more elaborate and sophisticated. We will see that though these concepts raise important questions as regards reflective self-evaluation, they introduce new complexities and they do not have a firm basis.

[49] Frankfurt, "Freedom of the Will", 21, italics in the original.

Chapter 2: Further elaboration of reflective self-evaluation

In the first chapter I looked at Frankfurt's classical hierarchical theory of desires and the theory of free will based on it, and tried to evaluate some arguments developed against it. As I tried to show, the most important counter-argument is that higher-order desires and volitions seem to have no basis. As Bratman points out, a second-order desire is just "one more desire in the psychic stew"[50] and the ability to form second-order desires is only the manifestation of *weak* reflectivity. He argues that we need something more, namely *strong reflectivity* which requires "the capacity to make a stand as an agent" with the help of which the agent determines where he stands with respect to a first-order desire.[51]

This is precisely what Frankfurt tried to do by stressing the importance of commitment and identification. In the second chapter I will discuss these concepts to see whether the hierarchical theory of desires can be rescued with the help of them. I will argue that though they help to make the concept of reflective self-evaluation more elaborate and sophisticated, this attempt is unsuccessful since the original problem was the lack of basis of

[50] Bratman, Michael E. *Structures of Agency*, (Oxford: Oxford University Press, 2007), 23.
[51] Ibid., 24. Bratman also tried to combine Frankfurt's hierarchical theory of desires and his planning theory in his "Hierarchy, Circularity and Double Reduction", In *Contours of Agency: Essays on Themes from Harry Frankfurt*, ed Buss, Sarah and Overton, Lee, 65-85 (Cambridge, Massachusetts: The MIT Press, 2002).

higher-order desires and introducing yet more higher-order phenomena will not help. These require some basis and explanation just as higher-order desires do, and consequently they cannot solve the problem, they can just push it one step further. What is more, the diverse higher-order phenomena that Frankfurt has introduced are very controversial themselves and make the theory vulnerable to other counter-arguments.

First, let us look at the concept of *decisive commitment* introduced first in Frankfurt's classic paper. As we have seen, Frankfurt claimed that one can stop an ascent to higher orders by making a decisive commitment and his decision "resounds to a potentially endless array of higher-orders".[52] He also states that such a decision is not arbitrary. However, Watson is not convinced that this is so. He thinks that Frankfurt's concept of decisive commitment simply means that an "interminable ascent to higher orders is not going to be permitted",[53] and argues that this is arbitrary.

Now, is "arbitrariness" always a problem? In what sense is the decision arbitrary? Sometimes, when we are faced with two equally appealing options, as for example in case of Buridan's Ass, it is rational to simply arbitrarily pick one of the options. Similarly, after having ascended to say, the third or the fourth order, we might say that further ascent would be time-consuming and impractical and commit ourselves to one of our higher-order desires. Commitment in this sense is not arbitrary: it serves reason and effective action. It is a function of decision making that one does not permit any further ascent; to use Frankfurt's etymological explanation it is not a coincidence that the verb "to decide" is related to the verb "cut off".[54] Therefore, since decision serves practical purposes, it is not altogether arbitrary to terminate the ascent at a given point. It is part of our ability to act that at some point or other, we commit ourselves to some of our desires.

[52] Frankfurt, "Freedom of the Will", 21.
[53] Watson, "Free Agency", 29.
[54] Frankfurt, Harry G. "Identification and Wholeheartedness", in Frankfurt, *The Importance of What We Care About*, 170.

Another point is that in order to make a decisive commitment, we need to trust ourselves. Self-trust is an important condition of agency, as it has been pointed out by Keith Lehrer.[55] Ascending to higher and higher orders signifies a lack of self-trust and an extreme form of self-doubt that paralyzes us and makes it impossible for us to function as agents. As we will see, *caring* is central to agency, and what we care about is something that we do not choose and cannot control directly by our decisions. This is precisely why we need self-trust: we have to regard ourselves as capable of caring about the right things.

Nevertheless, Watson is right that our decisions should be based on something, namely on a *reason*. In turn, he thinks that our reasons should be based on our value judgments. Thus, his view is Platonic, since value judgments are rational to some extent. But this is not the only way how we can have reasons. We will see later on how Frankfurt's recent views on caring were developed into a – very controversial – theory of practical reason. Especially in his book, *The Reasons of Love*, Frankfurt claimed that caring and love give us reasons to act. It is not arbitrary to commit ourselves to what we care about; rather, our ability to do this reflects our ability to trust ourselves and stand behind what is important to us. To sum up this point, though Watson is right that it would be arbitrary to simply terminate a series of ascent to higher orders, decisive commitment is not simply stopping the process at any given point but an important ability that is part of being an autonomous agent.

But before Frankfurt turned to caring and love, he first tried to defend his classical view against criticism with the help of introducing other types of higher-order phenomena, including and commitment and identification, endorsement and satisfaction. In the next section I will look at these and show that though they make the concept of reflective self-evaluation more elaborate, they have to face similar counterarguments as second-order desires and volitions.

[55] Lehrer, Keith, *Self-Trust: A Study of Reason, Knowledge and Autonomy* (Oxford: Claredon Press, 1997).

2.1. Identification and Rejection

The concept of identification is one of the most problematic parts of Frankfurt's philosophy of action. It is an important part of the arsenal that makes up *reflective self-evaluation*, which is a centrally important element in Frankfurt's thinking. As we have seen, this view stresses that we can only be persons and agents if we are able to ascend to hierarchically higher orders and reflect on our own motivation in a special way.

Frankfurt also stresses that *being active* is indispensable for agents. It is intuitively true that being an agent requires that we actively participate in the process of deliberation. Velleman argues that Frankfurt's theory is a "non-standard" philosophy of action, in the sense that it builds agency on something that the agent actively does as opposed to something that merely happens in him. In his view,[56] standard models claim that desires taken together with our beliefs cause actions, which seems to suggest that agency is reduced to a process in which some kind of events (beliefs and desires) cause other events (intentions and actions).[57] He thinks that the central problem with this view is that it seems to reduce agency to something that merely happens in the agent, making him some kind of passive bystander. However, it seems to be intuitively true that agents are *active*, whatever that may mean.

The importance of the concept of identification is that it both complements the theory of hierarchical desires and tries to account for the active nature of human agency. Let us look at first Frankfurt's discussion of this problem in one of his earlier papers, "Identification and Externality".[58] The notion of identification is quite elusive and difficult to grasp; as

[56] Velleman "What Happens When Someone Acts?"
[57] For a standard causal explanation of action see Davidson, Donald, *Essays on Actions and Events* (Oxford: Claredon Press, 1980) Davidson argues that we can understand actions through causal explanations. Frankfurt criticized this view in detail in his paper, "The Problem of Action", in *The Importance of What We Care About*, 69-79, reprinted from *American Philosophical Quarterly* 15 (1978).
[58] Frankfurt, Harry G.: "Identification and Externality", in: Frankfurt *The Importance of What We Care About*, 58-69, originally in Rorty, Amelie (ed.), *The Identity of Person*, 239-253 (Berkeley: University of California Press, 1977).

Shoemaker notes it, the "talk of freedom and identification is messy and difficult, precisely because human beings are messy and difficult".[59] But it seems that at first approximation, identification is a conscious act that the agent performs, with the help of which he renders some of his desires as truly his own while expelling some of them from it.[60] Frankfurt has the following example: "[S]uppose a person wants to compliment an acquaintance for some recent achievement, but that he also notices within himself a jealously spiteful desire to injure the man".[61] If this person identifies with his desire to make a compliment to the acquaintance, this makes the desire genuinely his own and his rejection of his jealous desire to hurt him makes it external to him. Thus, the notion of identification is not only connected to the active-passive, but to the *internal-external* distinction as well: if an agent identifies with a certain desire, it becomes internal to his identity, and if he rejects this desire, it becomes external. But Frankfurtian identification is not merely an attitude; rather, it is a *commitment* or a *decision*. Frankfurt thinks that a person clarifies his relation to his desires by this decision and adds that "[I]t may be that a decision of this kind, even if it is not as visible as in the present example, lies behind every instance of the establishment of internality or externality of passions".[62] This also implies that since the desires that we identify with become part of our identity and the desires that we reject become external to, we *create* our identity through such decisions.

Frankfurt further explores the concept of identification in his later paper, "Identification and Wholeheartedness". In this paper we can find a detailed discussion of identification as a particular type of decision:

> …it is characteristically by a decision […] that a sequence of desires or preferences of increasingly higher orders is terminated. When the decision is made without reservation, the commitment it entails is decisive. Then the person no longer holds himself apart from the desire to which he has committed himself. It is no longer

[59] Shoemaker, David W, "Caring, Identification and Agency", in *Ethics* 114, No. 1, (October 2003), 117.
[60] Ibid., 68
[61] Ibid., 66-67
[62] Ibid.

> unsettled or uncertain whether the object of that desire – that is, what he wants – is what he really wants: 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*. The pertinent desire is no longer in any way external to him. It is not a desire that he "has" merely as a subject in whose history it happens to occur, as a person may "have" an involuntary spasm that happens to occur in the history of his body. It comes to be a desire that is incorporated into him by virtue of the fact that he has it *by his own will*.[63]

All this clearly suggest that identification is a crucial element of the higher, reflective level. It manifests itself in decisions with the help of which, as it were, the agent sorts out his desires, and groups them into those which are fully his own, and those which he expels from his identity. And when he acts on a desire that belongs to his identity or "real self", he acts as an active, free, autonomous agent, as opposed to acting on external desires which involves becoming passive and subjected to desires that one regards as alien to oneself. Thus, identification is a quite sophisticated tool with the help of which the agent *constitutes* his own personal identity.[64]

Now, what is exactly the point of identification? Why should we have to make some of our desires internal, and why do we have to externalize some of them? According to Frankfurt, such decisions regarding our desires have an important function for the agent: they "replace the liberty of anarchic impulsive behavior with the autonomy of being under his

[63] Frankfurt, "Identification and Wholeheartedness", 170.

[64] For some, this concept of identification might be too sophisticated to accommodate average human agency. Frankfurt is not the only one who claims that identification is a sophisticated higher-order process. Bratman, for example, claims that when an agent identifies with a desire, it means that he decides to treat it as part of a policy that guides his actions. See: Bratman: "Identification, Decision and treating as a Reason", *Philosophical Topics*, 24, 1996, 1-18. On the other hand, Mele defines this notion in the following way: "to identify with a desire one has is, roughly, a matter of desiring to continue to have that desire and believing that one's having it is a good thing" (Mele, *Motivation and Agency*, 227). Curiously, this definition involves Frankfurt's decfinition of caring, plus a belief about the value of the given desire. Mele admits that this is a "thin notion of identification; but he does not want to include self-constitution and policies in it as Frankfurt and Bratman do, in order not to "raise the bar too high", which would not make it possible to accommodate average agency, only agency *par excellence*. Or a different definition: "an agent's identifying with a desire D, or a value V, that he has may be understood as a combination of his having D, or his having V, and his valuing his having it". Mele, Alfred R. *Autonomous Agents*, Oxford: Oxford University Press, 1995, 117. Though this definition also includes higher-order processes, it does not go as far as Frankfurt or Bratman in the level of sophistication that they require.

control".[65] But not only self-control is the primary issue here. For what is the point of being in such kind of control of our motives? Let me quote here Frankfurt at length here:

> It might be said, then, that the function of decision is to integrate the person both dynamically and statically. Dynamically insofar as it provides […] for coherence and unity of purpose over time; statically insofar as it establishes […] a reflexive or hierarchical structure by which the person's identity may be in part construed.
> In both respects, the intent is at least partly to resolve conflict or to avoid it. This is not achieved by eliminating one or more of the conflicting elements so that those remaining are harmonious, but by endorsing or identifying with certain elements which are then authoritative for the self.[66]

The distinction between dynamic and static integration is very important and I will discuss it in detail in the next part as *diachronic* and *synchronic volitional unity* of the agent which Frankfurt discusses in detail in his book *Taking Ourselves Seriously and Getting It Right*.[67] But we can already see here that identification has a very important and specific function: it provides the agent with temporal continuity and unity.

The reason why we need such unity is the same why we need second-order desires. We have to avoid wantonness, or being agents who act on any occurring desire without considering what they want. As I will discuss later in detail, wantonness has both a synchronic and a diachronic aspect. We can avoid being wantons at a particular time by creating a reflexive, hierarchical structure, that is, by forming second-order desires and volitions. But we persist through time, and we need something that prevents us from replacing this structure with another in a chaotic manner. Frankfurt admits that there might be a tendency for a given structure to persist, but in that case we would be at mercy of this tendency.[68] Therefore, we need something to actively sustain a particular hierarchical structure, and this is precisely the point of creating dynamic coherence by means of the particular type of decisions described as identification above. For when an agent identifies

[65]Frankfurt, "Identification and Wholeheartedness", 175.
[66] Ibid.
[67] Stanford: Stanford University Press, 2006. The book contains Frankfurt's *Tanner Lectures* delivered at Stanford University (April 14-16, 2004).
[68] I will come back to this claim in *chapter 5* and argue that some desires by nature tend to persist and this makes them more important than passing wishes and whims.

with a desire, it does not only mean wanting it on a higher order, it also means that one conceives oneself as a person who has that particular desire. And since agents persist through time, identification helps make a particular reflexive, hierarchical structure temporally stable.

Now, why do we exactly need such kind of temporal unity in our volitions? Somebody might simply say that he does not prefer to have one, and is happy with how his volitions change through time. But though people might disagree as to the extent of volitional continuity, it is not the case that one can be an agent and not have any persisting desires. Even the agent above expresses a positive attitude towards not having temporally stable desires, and in this case, this attitude itself is a temporally stable desire that suffices for some level of volitional continuity.

To see that this is so, let me briefly consider Galen Strawson's criticism of a type of diachronic conception of personal identity, the narrative self. In a recent paper he argues against a diachronic notion of the self, according to which human life should be lived as some kind of narrative[69] and he argues for an "Episodic View" instead. If one is episodic, "one does not figure oneself, considered as a self, as something that was there in the (further) past and will be there in the (further) future".[70] But this view can be reconciled with Frankfurt's view. For it only implies that there are narrative people and episodic people and that they find different aspects of life more important. A Strawsonian agent *does* care about something: avoiding a single narrative and living in the episodic way because he believes that the "truly happy-go-lucky, see-what-comes-along lives are among the best there are, vivid, blessed, profound".[71] Such people live their life the way they do because they think that it is the best way one can live one's life. Though these people might never settle in life in any sense and they value independence as lack of being attached to a carrier, a place or persons, this does not imply that their lives are not diachronic. They are just indifferent about some (narrative)

[69] Strawson, Galen: "Against Narrativity". In: Stawson (ed.), *The Self?*, 63-87 (Malden, MA: Blackwell, 2005).
[70] Ibid., 65.
[71] Ibid., 84.

values including carrier, personal relationships or family. The episodic individual's self is defined by a concern to always live in the present, which concern itself is diachronic.

Thus, agency requires at least some level of continuity in our desires. This is because we are temporal creatures and the process of deliberation is embedded in our temporal existence. As a consequence, decision making should be influenced both by our past and our future. Our actions are essentially connected to the future through their consequences and are tied to our past. When we make a decision to do something, we have to consider its consequences and take into account our future preferences as well. For example, when somebody is considering accepting a job, he has to think about whether he will want to do that job for a certain length of time. It is not enough to want to do something at the moment of decision but the agent has to consider whether he is such kind of person who wants to do that job. He also has to think about how this particular job fits into his long term goals and past experience. Or to take another example, when somebody is offered an alcoholic drink, he has to think about how it relates both to his immediate and distant future and his immediate and distant past. If he wants to drive for instance, this should influence his decision, but he should take into more distant consequences as for example a commitment to avoid drinking liqueur or to avoid developing an addiction. Besides, his decision should also take into account his near and distant past: if he had already several drinks, or if he has a past of alcohol addiction, this should affect his decision. These examples show that very often, especially in case of important decisions, our actions are future directed and rooted in the past. Therefore, agency is essentially diachronic, or temporally extended.

Frankfurt's term for volitional integration is *wholeheartedness*. A wholehearted agent stands behind some of his desires and commits himself to them without any conflict. Thus, wholeheartedness is not simply the feeling of "enthusiasm" or "certainty"[72] but implies the

[72] Ibid.

volitional integration of the agent. One can be wholehearted synchronically, that is, having a second-order desire without any reservation, or one can be wholehearted diachronically as well, as a result of one's commitment to preserving a given hierarchical structure without any reservation.

But something is still missing here. Though now we have a functional explanation of identification, both a factual and a normative explanation is missing. Why does an agent, as a matter of fact, identify with a particular desire and not another? And what is more important, which desire *should* the agent identify with? If it is only the functional role of identification that is important, there would be no limitations on which desires the agent could identify with. As I will show in the next part, there are certain limitations on our ability to identify with our desires, both *factual* and *normative*, but we cannot explain these without introducing the concept of *caring*. In itself, the concept of identification is insufficient, for it seems to be *prima facie* true that when one makes decisions which affect how one lives one's life in the long term, these should be based on something. In *chapter 3* I will argue that Frankfurt's concept of caring, sufficiently redefined, can perform this role. But before that, let me consider another objection to Frankfurt's concept of identification in the next section.

2.2. Identification and the Real Self

Before we move to the discussion of caring, let us look at other important objections to the concept of identification. The most important one is concerned with Frankfurt's claim according to which the agent constitutes his identity through identification. As I will show later, Frankfurt's theory of autonomy is a *real-self view*, that is, it claims that one is free and

autonomous when one's actions are determined by his real self.[73] Since Frankfurt's concept of caring makes it possible to develop a richer real-self view, I will discuss it more in detail later on. Here I will only look at the following: does identification help us to find what our real self is? That is, does my identifying with a desire make it part of my identity as Frankfurt claims? What is one's "real self" after all, and what does it mean that our actions should be determined by it?

As we have seen, identification in Frankfurt's sense involves making decisions concerning our desires, either authorizing or rejecting them. Now, this implies having a strong control over our desires. One objection is connected to this feature of identification, and claims that Frankfurt's views to imply too tight self control. As Schechtman puts it, "we frequently view overly rigid self-control as an impediment to being oneself".[74] Thus, it seems that the practice of identification, rather than helping one create one's identity, prevents one to have one. When the agent is identifying with some of his desires and is repudiating some of them, he seems to be exercising self-control to an extent that kills spontaneity. Such kind of self-control might be pathological and the result of excessive self-concern and self-doubt. Why should agents be concerned with their desires all the time? Isn't the primary issue of deliberation a certain course of action after all? In other words, isn't deliberation primarily a first-order process? If that is so, putting the emphasis on reflective self-control in Frankfurt's sense involves putting the emphasis on the wrong side. Agents who are focusing on their desires too much might not be able to concentrate on the primary issue of deliberation, namely on what they should do.

[73] Susan Wolf discusses and criticizes the real-self view in detail. Wolf, *Freedom Within Reason*; see especially Chapter 2, 23-46. Lippert-Rasmussen offers convincing arguments against real-self views of responsibility in his paper "Identification and Responsibility", *Ethical Theory and Moral Practice* 6, No. 4 (December 2003), 349-376.

[74] Schechtman, Maya: "Self-expression and self-control", in Strawson, *The Self?*, 49.

Also, an agent who is so much concerned with his desires lacks a basic sense of trust in them. I have already mentioned that the lack of self-trust is against the effective action and thus, autonomous agency. Of course, Frankfurt is right that we should not just simply act on any desire in a wanton way as they occur, but questioning them too frequently is the sign of the lack of basic self-trust. Identifying and rejecting desires can easily become a method of pathological second-guessing of what one really wants.

But besides the danger of preventing one to be oneself, the concept of identification seems to be problematic from a different point of view as well. Frankfurt seems to imply that one's identity is made up of those desires that he identifies with, and those that he rejects are excluded from it. That is, the desires the agent identifies with are *internal* and those that he repudiates are *external* to the self. This suggests that while internal desires are a constitutive part of what I am, or my real self, external desires are somehow not part of me. But this is far from being intuitively true. Many would have a very different intuition, namely that there are desires which are mine and express what kind of person I am whether or not I identify with them. For instance, Schechtman points out that often, "a person is truly herself when she fails at self-control".[75] In some cases one fails at self-control and one ends up acting on a desire that one consciously repudiates, but whether such an agent wants it or not, this desire is genuinely his own. In such cases a failure at self-control implies a failure to conceal what an agent really wants and what his real self is.

This problem is even more apparent in light of some pathological cases. For example, Vellemann discusses in detail the problem of *repression*.[76] Through this pathological process, the agent represses some of his desires, and expels them from his identity. He discusses in detail the well-known case of the "Rat Man" from Freud. As it is well known, the Rat Man's problems were rooted in the fact that he repressed his hatred for his father, which lead to the

[75] Ibid., 50.
[76]Velleman, "Identification and Identity", in *Contours of Agency: Essays on Themes from Harry Frankfurt*, ed. Buss, Sarah and Overton, Lee (Cambridge: The MIT Press, 2002), 91-124.

development compulsive behavior. Velleman claims that this is close to the practice of identification and rejection Frankfurt describes;[77] that is, the practices that he suggests are similar to pathological repression as a defense mechanism. By claiming that his hostility towards his father is not part of his identity, the Rat Man represses it and defends himself against it. This is clearly a pathological practice. It is not true that his hatred is not part of his identity; he only would like it to be the case.

Velemann claims that what Frankfurt talks about is not one's self but one's self-conception, self-understanding or self-image[78], or we can perhaps call it one's ideal self. According to Velleman's reflexive, Lockean conception of the self, "a person's past or future selves are just the past or future persons whom he can pick out with thoughts that are notionally reflexive, first-personal".[79] In this sense, a person's self is not identical with the motives he identifies with; rather, one's self is the person at which reflexivity is directed. In this sense, the Rat Man's hostile wishes against his father belong to him, whether he wants it or not.

Thus, it seems that first-order desires can be part of one's real self, even if one refuses them. To go back to the Rat Man's example, the right strategy for him is not expelling his hostility towards his father as not part of himself, but rather, admitting that unfortunately he *does* have this motive; it is part of what he is. Or as Velleman puts it: "what the Rat Man should have done was to accept his filial hostility as part of himself, to accept *himself* as ambivalent toward his father".[80] Curiously, then, it would have been better for the Rat Man to be ambivalent than wholehearted. Accepting his ambivalence would lead to an opportunity to change himself and to manage his hostility toward his father.

[77] Ibid., 102
[78] Ibid., 110
[79] Ibid., 111
[80] Ibid., 103, italics in the original

Talking about hidden, hostile feelings in general, Solomon states something similar: "[A] person who recognizes his or her own envious or resentful nature will grudgingly accept envy or resentment as his or her own, and acknowledge the larger narrative in which the unflattening emotion plays a part".[81] Recognizing such emotions is a process that involves gradually accepting them as one's own. This recognition can form the basis of a strategy to change these emotions. Notoriously, Solomon believes that we have a great degree of control over our emotions; what is more, in one of his earlier and most controversial papers, "Emotions and Choice",[82] he claimed that we choose our emotions.

Frankfurt has a detailed answer to Velleman's criticism. He refuses Velleman's claim that the Rat Man's rejecting his hostile wishes against his father is similar to repression as a pathological defensive mechanism. Repression involves concealing a desire or making in unconscious, and he goes on to say that "making a decisive stand against certain feeling does not require (or even permit) that a person misrepresent those feelings or that he conceal them from himself".[83] The difference here seems to be the following. What the rat Man performed is a pathological defensive mechanism, and it involves *not letting in* some of his desires into his consciousness. As a consequence, he is not even aware (at least in a full sense) of his having them. On the other hand, the practice of identification and rejection suggested by Frankfurt can only work for the Rat Man if he first goes through an ambivalent state. For how could he reject his hostile desires toward his father if he is not even aware of them? As we have seen, Frankfurtian identification is a sophisticated tool used to create one's identity selecting among one's desires, and it obviously requires knowing what the desires to select from are in the first place. The Rat Man can simply acknowledge that his hostile wishes against his father are the part of his psychological makeup; it belongs to, so to speak, the raw

[81] Solomon Robert C., "On the Passivity of Passions", in *Not Passion's Slave: Emotions and Choice* (Oxford: Oxford University Press, 2003), 207.
[82] Solomon, Robert C., "Emotions and Choice", in *Not Passion's Slave*, 3-24, reprinted from *Review of Metahysics* 28, No. 1 (September 1973).
[83] Frankfurt: "Reply to J. David Velleman", in *The Contours of Agency*, 126.

material he has to work with. Then he has to make a stand in the sense of Frankfurtian identification, and repressing his desire just precisely consists in *avoiding* doing so.

But even if Frankfurtian identification and rejection is not identical with repression as a pathological self-defense mechanism, we are still without an explanation of it. Though we have now a functional explanation, we are still without a real explanation. It seems that identification is quite a mysterious phenomenon without any basis. As we have already seen, an important objection to Frankfurt's views on agency is that it involves arbitrarily picking an element in one's psychology. Why does one form certain higher-order desires and volitions instead of forming others? Critics like Watson demand that one's higher order attitudes be based on something, for example, one's evaluations and judgments. Frankfurt cannot effectively answer to this criticism by referring to yet other higher order elements in one's psychology. If we want to explain higher-order desires it will not be much help to refer to yet other higher-order attitudes that themselves demand an explanation. In such a debate a Frankfurtian can always be forced by his critics who demand an explanation of a certain higher-order desire to ascend to yet higher-orders or postulate different kinds of higher-order phenomena. Of course, one obvious strategy would be to assert that there is a certain special higher-order phenomenon that *terminates* any ascent to further levels by simply not permitting it. As we have seen, Watson objected that terminating the ascent at a specific point is arbitrary, but one might simply claim that in a sense, the will itself is an arbitrary phenomenon; furthermore, its freedom lies in its very arbitrary nature. Sartre's existentialism has a similar point about freedom: it is not the case that man has an essence that determines his actions but we are free to define ourselves in an apparently arbitrary way.[84]

[84] "What is at the very heart and centre of existentialism, is the absolute character of the free commitment, by which every man realizes himself" (Sartre, Jean Paul: *Existentialism and Humanism*. London: Methuen, 1955), 47. Or, he declares that "[M]an makes himself; he is not found readymade; he makes himself by the choice of his morality" (Sartre, 50). However, I will argue later on that our second-order commitments are not arbitrary in the sense Sartre implies, but are based on what we care about, which, as I will try to show, is a first-order matter.

Frankfurt treats concept of identification as involving crucial decisions that shape one's practical identity and thus guide one's actions. His notion of identification, then, is connected to a form of evaluation, as he calls it, *endorsement*, which involves evaluating one's motives and endorsing some of them as legitimate parts of one's identity and rejecting some others as alien to it. But what does identification in the sense of endorsement depend on? Why does one value some of his motives and not others? Naturally, one would think that one values his motives because one has a good *reason* to do so. But Frankfurt rejects rationalistic interpretations of his concepts.[85] It is not the case that one identifies with a desire because he has an independent, justifying reason to do so. It is not reason that leads to identification, but the other way round: the fact that I identify with a desire gives me a reason to act in a certain way. In his reply to Moran's criticism, he admits that "identification and reasons are essentially related" but it is not true that identification depends on reasons, rather, "it is identification that indispensably constitutes the source and the ground of reasons".[86]

The same is true about value: there is no objective value that would precede our evaluations, rather, the other way round: when we endorse or value positively one of our desires we *create* value. Frankfurt developed this very controversial theory is in detail in his book, *The Reasons of Love*, and I will discuss it later. The basic tenet of this theory in its rawest form seems to be quite unacceptable. It claims that we do not desire or want something because it is valuable, but the exact opposite: something has value for us because we desire or want it. In this crude form this view is obviously mistaken. If I want, say, to jump out of the window, it does not make this activity valuable. However, we will see that desiring or wanting something can mean something more complex and sophisticated, and this

[85] See for example: Scanlon: "Reasons and Passions" and Moran: "Frankfurt on Identification", in *Contours of Agency*, 165-184 and 189-218. He also contrasts his view with the rationalism developed by Joseph Raz and refuses Raz's view that wholeheartedness is based on rationality. See Frankfurt, "Disengaging Reason", in *Reason and Value: Themes from The Moral Philosophy of Joseph Raz*, ed. Wallace, Pettit, Scheffler and Smith, 119-129. Oxford: Claredon Press, 2004.

[86] Frankfurt "Reply to Richard Moran" in *Contours of Agency*, 219.

can make this theory more appealing (though still very controversial). In the Frankfurtian framework, caring and love are modes of desiring or wanting which are capable of guiding us, and creating real value in our lives. But before I move on to these, I will look at Frankfurt's concept of *satisfaction* which is also a feature of a hierarchical system of desires, but very different from the others.

2.3 Satisfaction

As I have tried to show above, the concept of identification is problematic because it does not seem to have a basis. But Frankfurt introduced yet one more concept that might help defend his views. This is the concept of satisfaction, which, in a sense, strikingly different from the earlier ones and in a sense, is an improvement, but as I will show, it has to face other counterarguments.

Frankfurt is not the only one who makes use of the concept of satisfaction. It also has a role in Bratman's theory of action who points out that its importance is connected to *stability* which is required for *temporally extended agency*.[87] Thus, its role is similar to that of identification: it helps us create a temporally stable volitional coherence. For example, my satisfaction with my desire to pursue a career in philosophy helps me make my actions more coherent and organized through time. My satisfaction with this desire contributes to its enduring presence and to its action guiding role.

Let us first look at what exactly Frankfurt means by satisfaction. He discusses this concept in detail in his paper "The Faintest Passion" where he defines it as "a state of the entire psychic system – a state constituted just by the absence of any tendency or inclination to alter its condition".[88] It is important to note that satisfaction is not a further higher-order

[87]Bratman, *Structures of Agency*, 99.
[88] Frankfurt, Harry G. "The Faintest Passion", in Frankfurt, *Necessity, Volition and Love* (Cambridge, Cambridge university Press, 1999), 95-107.

desire or attitude; that is, being satisfied does not mean that we have yet one more higher-order desire directed at some of our desires. Rather, when we are satisfied with a desire it simply refers to the absence of any contrary attitudes to it. As Bratman puts it, Frankfurt's concept of satisfaction refers to a "non-occurrence".[89] Let me quote here Frankfurt at length:

> Now what does it mean to say of a person that he is satisfied with his psychic condition, or with some element or aspect of it? It does not mean that he considers it the best condition available to him. Some people may be so demanding that they are never willing to settle for anything less than that. But as a rule, satisfaction is not conditioned by an uncompromising ambition to maximize. People often settle gladly for less than what they think it would be possible for them to get. From the fact that someone is satisfied with his condition, then, it does not follow that no alteration of it would be acceptable to him. It goes almost without saying, of course, that he would be satisfied with an improved condition. However, he might also be satisfied even with a condition inferior to the one he is in. What satisfaction does entail is an absence of restlessness or resistance.[90]

In a sense, this is an important improvement to the concepts of commitment and identification. As we have seen, these are problematic because an attempt to find a basis for higher-order desires and volitions with the help of introducing other higher-order attitudes that themselves demand an explanation is doomed to failure. Therefore, since satisfaction is not a further higher order desire, it is a good candidate to solve the problem. We could say for example that someone identifies with a desire and commits himself to it because he is satisfied with it, or in other words he lacks any negative attitudes to it.[91]

But this concept raises an important problem. As Bratman notes, satisfaction with a desire in this sense "may be grounded in depression, and in such cases satisfaction with a desire does not seem to guarantee agential authority".[92] We might also add that in light of his stress on the active nature of deliberation, Frankfurt's characterization of satisfaction is

[89] Bratman: "Identification, Decision and Treating as a Reason", in *Faces of Intention* (Cambridge: Cambridge University Press, 1999), 194.
[90] Frankfurt, "The Faintest Passion", 103
[91] Frankfurt notes that satisfaction does not necessarily entail commitment (ibid), but it seems that since being satisfied with a desire entails treating it as the best available option, it still contributes to commitment through excluding contrary desires.
[92] Bratman, "Identification, Decision and Treating as a Reason", 204. Velleman made a similar point and claimed that if an agent's satisfaction with a desire is due to depression or *ennui*, it cannot be "an expression of his own will". Velleman, "Introduction", 13.

curious since it seems that it could refer to some sort of *passivity* with regard to what moves one to action. An agent might know that he could be much better off and that he could improve his situation but does not do anything to reach that goal, because he is satisfied with the way things are. This state is not necessarily negative; acceptance can be justified if the given situation, so to speak, it is "good enough" for the agent. But it can easily be the result of resignation or severe depression as well. In that case it does not express what the agent really wants, or as Bratman puts it, his authority as an agent, rather, it might be passivity as a result of his depressed state.

Satisfaction therefore, as the sheer non-occurrence of negative attitudes toward a desire is insufficient to ground agential authority. As we have seen, in itself it can be the result of depression and resignation, so the real question is not that someone does not have any negative attitudes toward a desire but *why* he does not have any. It seems then, that Frankfurt faces the familiar objection we encountered earlier: he is trying to explain agency with a certain mental phenomenon that itself needs an explanation. When satisfaction rests on a certain basis, that is, when one is satisfied with a desire *because of* a reason (in the broadest sense of the word), it can be a sign of agential authority, but in that case, authority rests on the reason itself and not satisfaction after all.

Also, why do we even have to say that satisfaction is merely a *negative* state?[93] I think this view involves an unwarranted pessimism. There are striking similarities between Frankfurt's concept of satisfaction and the concept of satisfaction involved in Schopenhauer's classic theory of pessimism. Of course, Schopenhauer and Frankfurt talk about something very different: while the former talks about *satisfying* our desires, the latter emphasizes *being*

[93] Note though, that in a sense, the term "negative concept" is rather vague. For it is possible to define any concept as "negative" by defining it as the absence of something. But defining for example freedom as the absence of coercion does not make it a negative concept. However, Bratman's point that Frankfurt's concept of satisfaction refers to a "non-occurrence" touches upon a real problem for Frankfurt: when we are satisfied with our own motives, for example, we do not mean that any contrary desire is absent. Rather, there is a positive sense of approval which differentiates it from depression.

satisfied with having them, which obviously could make no sense at all for Schopenhauer. But though these are very different theories, they both have a negative concept of satisfaction. Schopenhauer thinks that satisfaction does not refer to the presence of something positive, only to the absence of desire through its elimination:

> All satisfaction, or what is commonly called happiness, is really and essentially always *negative* only, and never positive … [T]he satisfaction or gratification can never be more than deliverance from pain, from a want … [N]othing can ever be gained but deliverance from some suffering or desire; consequently, we are only in the same position as we were before this suffering or desire appeared.[94]

Desire causes tension and pain, and satisfying it is nothing but the elimination of this pain; it does not involve any positive element.[95] A pessimist might say that happiness is nothing but the absence of pain, meaning that it does not involve any positive state only the absence of negative states. But this dark view is unwarranted. Why could we not define happiness and satisfaction as the presence of a positive state? As Simmel pointed out, Schopenhauer "should not […] have overlooked the positive moment of happiness which differentiates it as a psychological fact from sleep and death, the two other states that end suffering".[96] Satisfaction ends suffering, but contains positive elements as well. We can easily see that this is a coherent option if we realize that one can be happy and yet experience pain and that one can be without pain and yet be unhappy. Happiness is constituted by the presence of something positive and thus is not taken away by the presence of something negative; and the absence of negative states is not sufficient for happiness either. This suggests that satisfaction and happiness cannot be defined by a simple pain-pleasure calculus, as pointed out by

[94] Schopenhauer, *The World as Will and Representation*, (New York: Dover Publications, 1969), Vol. 1, 319.
[95] Compare with Locke's concept of uneasiness: "That desire is a state of uneasiness, everyone who reflects on himself will quickly find", Locke, John, *An Essay Concerning Human Understanding*, (Oxford: Claredon Press, 1894) Vol. 1, 333. Thus, when we act we eliminate the most pressing uneasiness. Locke's concept of happiness is quite different from that of Schopenhauer, though. For him the elimination of uneasiness (the satisfaction of a desire) is not yet happiness, only the "first step" to happiness: uneasiness "is inconsistent with happiness, spoiling the relish of even of those good things which we have" (178). For a recent discussion of Locke's concept of uneasiness see also Yaffe: *Liberty worth the Name: Locke on Free Agency* (Princeton University Press, 2000).
[96] Simmel, Georg: *Schopenhauer and Nietzsche*, (Amherst: University of Massachusetts Press, 1986), 64.

Janaway, who thinks that Schopenhauer's pessimism rests on an unfounded and simple form of hedonism.[97]

Strikingly, regardless of their different treatment of the concept of satisfaction, the two theories have a logical similarity: they both try to define satisfaction in negative terms, or as the absence of any negative state. And I think that in the same way as Schopenhauer's pessimism is unwarranted, Frankfurt's negative concept of satisfaction is unfounded as well. Why do we have to say that satisfaction with a desire is merely a negative state, that is, it is constituted by the absence of any negative attitudes toward it? Why couldn't we say that being satisfied with a desire means that we have a positive attitude toward it? If we defined satisfaction this way, the passive state of the depressed could not qualify as satisfaction. Rather, satisfaction would be a positive attitude toward a desire. But this would expose the theory to the same difficulty that we encountered earlier: we cannot effectively explain a higher-order attitude by referring to another higher-order attitude that itself demands an explanation. For why is an agent satisfied with a certain desire in this positive sense and not with another?

As I pointed out above, we need to have some kind of *reason* to be satisfied with a desire. Let us look at Bratman's solution to the problem. He also emphasizes that satisfaction with a desire is a crucial element of temporally extended agency and his extended concept of identification connects it with treating as a reason:

> … a person is, in an extended sense, identified with a desire if (*i*) she treats it as reason-giving, (*ii*) she does not treat it as external, and (*iii*) she would decide to continue to treat it as reason-giving, be satisfied with that decision, and continue to treat it as reason-giving if she were to reflect on the matter.[98]

[97] Janaway, Christopher: "Schopenhauer's Pessimism". In: Janaway, Christiopher (ed.): *The Cambridge Companion to Schopenhauer*, (Cambridge University Press, 1999), 318-344. For further arguments against Schopenhauer's pessimism see also: Nussbaum, Martha C.: "Nietzsche, Schopenhauer and Dionysus". Janaway, 344-375.

[98] Bratman, "Identification, Decision and treating as a Reason", 204.

Here Bratman argues that satisfaction is the element needed for continuity and temporally extended agency. But one is not simply satisfied with a desire but with treating it as reason-giving and continuing to treat it that way. That is, satisfaction is only an element of a more complex phenomenon. Parts of this phenomenon are the desire itself, our decision to treat it as a reason, being satisfied with that decision and also with continuing treating it as a reason upon reflection.

The way how Bratman puts the concept of satisfaction into a larger context suggest that it is mistaken to try to explain human agency by stressing one single element in our psychology. Frankfurt, as we have seen, emphasized different elements in his different papers. At first he emphasized our capacity to form second-order desires and volitions, but after he encountered objections to his theory, he tried to rescue it by emphasizing different higher-order phenomena, including identification and endorsement and later on, satisfaction. But if we try to explain human agency by stressing the importance of one of these, we will always face the same objection, namely that in themselves, these higher-order attitudes all demand an explanation, so they cannot be used to explain the others. It is a virtue of Bratman's theory of action that it tries to construct a complex, unifying theory in which the different elements are connected together.

In the second part of my thesis I will look at a very important development of Frankfurt's thinking, namely his concept of *caring*, and an important version of it, love. I will show that with these concepts, Frankfurt turns to a radically different direction.[99] While earlier he emphasized second-order acts of will that we perform voluntarily, his concept of caring is something quite different. What we care about has an important bearing on our actions but we cannot directly control what we care about. On the contrary, caring and love

[99] A similar interpretation of the shift in Frankfurt's thinking was developed by Stefan E. Cuypers who describes two different Frankfurtian views, Frankfurt-1 and Frankfurt-2. Frankfurt-1 emphasizes voluntarism and Frankfurt-2 nonvoluntarism. The first concept leads to a volunatristic concept of autonomy and the second to a nonvoluntaristic one. I will come back to Cuypers' views in *Chapter* 7. See Cuypers, "Autonomy beyond voluntarism."

impose *volitional necessities* on us and yet do not take away our freedom. Rather, since these constraints are internal to the will they make autonomous agency possible.

Part II: Frankfurt's Theory of Caring and Love

Chapter 3: Frankfurt's Concept of Caring

The concept of *caring* is centrally important in Frankfurt's later thinking. He first discussed it in his paper "The importance of What We Care About", and after he developed it in detail it became the basis of his view on human action. For him, caring is the most basic guiding principle of our lives. He claims that we have to be able to wholeheartedly care about things and commit ourselves to what we care about. Besides, he thinks that *love* is the most important mode of caring and thus the most important guiding principle of action. These are very controversial views and Frankfurt's theory of caring and love and their role in agency has been widely discussed.

We will see that this concept can be interpreted as a significant shift in Frankfurt's thinking. While earlier he emphasized the importance of second-order desires and volitions that we are capable to form voluntarily, the concept of caring involves *volitional necessities*, which we cannot directly control and yet have an important role in agency: they are part of our volitional essence. But before I discuss these issues in detail, in the first section I will differentiate Frankfurt's concept of caring from other contemporary theories that stress the significance of the ability to care.

3.1. Frankfurtian Caring and other Theories of Caring

The concept of caring has a central role in several influential contemporary theories of ethics, including feminist *ethics of care* and Michael Slote's *virtue ethics*. Therefore, it is necessary to differentiate these from Frankfurt's concept of caring before discussing it. Ethics of care is an important normative ethical theory that was developed by mostly feminist thinkers, among others, Anette C. Baier, Carol Gilligan and Nel Noddings. Similarly, Michael Slote's, virtue ethics is based on a normative concept of caring. Both are contrasted with the two most traditional ethical theories: utilitarianism and Kantian ethics and can be labeled as *moral sentimentalism*.[100]

It is important to stress that the concept of caring in feminist ethics and virtue ethics is not what Frankfurt has in mind. The main difference between the two concepts is that while both of the former belong to the tradition of *moral sentimentalism* and thus treats caring as a sentiment,[101] for Frankfurt it is a more complex phenomenon which includes affective, cognitive and volitional elements, with the emphasis on the last of these. Another difference is that while for moral sentimentalism empathy and caring are the basis of a concern for others, and thus for morality as such, for Frankfurt, caring is a much wider phenomenon, and it does not by any means imply being moral. Morality is just one thing to care about and consequently agents as Hitler can be characterized in terms of Frankfurtian caring and this approach to the concept would obviously be rejected by the ethics of care or virtue ethics.

But regardless of these differences, Frankfurt's concept of caring and that of virtue ethics and ethics of care has some important features in common as well. For one thing, in a sense they all belong to the Humean tradition, that is, they try to base their views on

[100] Another sentimentalist opposing Kantianism who thinks that care is the basis of ethics is Solomon. See Solomon, Robert C., *In Defense of Sentimentality* (Oxford: Oxford University Press, 2004), especially chapter 3: "Care and Compassion: Moral Sentiment Theory Revisited", 43-75.

[101] See for example:Noddings, Nel: "Caring". In: Held, Virginia (ed.): *Justice and Care. Essential Readings in feminist Ethics* (Boluder, Colo: Westview Press, 1995), 7-31 and Slote, Michael: "Moral Sentimentalism", *Ethical Theory and Moral Practice* 7, No. 1 (March 2004), 3-14.

something non-rational. They all think that the importance of rationality is exaggerated and try to base their views on a different foundation.

Another similarity is that for all of these theories, caring has a *normative* dimension. For virtue ethics, emphatic concern and caring is the basis of *moral virtue*.[102] The feminist thinker, Noddings claims that moral statements are "derived not from facts or principles but from the caring attitude".[103] Thus, Slote's virtue ethics and ethics of care are both opposed to Kantian ethics. Frankfurt shares this view and also refuses Kantian ethics very emphatically. For him, what we most care about tells us what we should do, but as I have mentioned above, his concept of caring is much wider, and consequently is not focused on being moral. He would say that though one can care about being moral, it is only one possible object of caring: he believes that "the importance of morality in directing our lives tends to be exaggerated".[104]

One more similarity is connected to the concept of *identification*. We have seen in the previous chapter that the concept of identification is centrally important for Frankfurt; however, he uses this term in a different sense in connection with love, which he regards as the most important mode of caring. In his book discussing love he claims that "a lover *identifies himself* with what he loves".[105] Slote stresses the importance of identification as well and believes something similar. He points out that empathy (and thus empathic caring) involves "a kind of metaphorical idea of identification with a feeling of the same thing."[106]

Another important similarity is that caring and *autonomy* are related both in different versions of moral sentimentalism and in Frankfurt's views. One can interpret the latter's concept of caring as providing a basis for autonomy, and likewise the feminist thinker, Grace Clement argues that personal autonomy and ethics of care can be reconciled with each

[102] See Slote, 7.
[103] Noddings, 22.
[104] Frankfurt: *The Reasons of Love* (Princeton, Princeton University Press, 2004), 6.
[105] Ibid., 61 (Italics in the original).
[106] Slote, 6.

other.[107] Both of these views have to face a similar problem in this project. However different their definition of caring is, both of them implies that caring somehow limits or restricts one's range of choices. Nevertheless, the both argue that these limitations do not necessarily mean the loss of freedom or personal autonomy, on the contrary: they somehow contribute to it.

Rational care theory of welfare developed by Stephen Darwall is another important contemporary ethical theory.[108] This view links personal welfare with the concept of caring by claiming that caring about someone means promoting someone's welfare or in other words promoting what is good for someone. Now, Darwall claims that "something is for someone's good if it is what that person would want for herself, as she actually is, insofar as she is fully knowledgeable and experienced *and* unreservedly concerned for herself".[109] I will come back to this claim later while discussing the subjectivism involved in Frankfurt's theory of love. Darwall argues that the ability to care involves being able to recognize and promote what is *really* good for the individual cared for. This applies for the ability to care about ourselves: we have to be able to promote what is objectively good for us. Since this theory brings in the objectively good into the problem of caring, I will come back to it while discussing Frankfurt's subjectivism in connection with his theory on love.

Finally, Bennett W. Helm developed a theory of caring and love which emphasizes the social nature of persons. Caring and love links us to others and we cannot interpret these concepts in the way we are traditionally interpreted, focusing only on the psychology of an agent. As Helm puts it, we have to refuse the *individualist conception of autonomy* and the *individualist conception of persons.*[110] My view has a lot in common with what Helm has to say and thus I will discuss some of his points on the way.

[107] Clement, Grace: *Care, Autonomy and Justice* (Boulder, Colo: Westview Press, 1996).
[108] Darwall, Stephen: *Welfare and Rational Care* (Princeton: Princeton University Press, 2002).
[109] Darwall, 31.
[110] Helm, Bennet W., *Frienship and the Self: Intimacy, Identification and the Social Nature of Persons* (Oxford: Oxford University Press, 2010).

In the next section I will move on to the definition of Frankfurtian caring and its significance.

3.2. Frankfurtian Caring and its Significance

We saw in *chapter 2* that an attempt to defend the hierarchical theory by making second-order phenomena more elaborate and sophisticated is doomed to failure because it faces the same counterarguments as the original theory of higher order desires and volitions. Endorsement and identification are themselves higher-order phenomena that need some explanation and consequently introducing them just pushes the original question one step further. And this was the following: what is the basis of higher-order desires and volitions? This question can be broken down into 1) *a factual* and 2) *a normative* question. The former concerns the origin of our higher-order desires and volitions: why do we form a certain higher-order desire and volition and not a different one? The latter focuses on the normative dimension of this process: what kind of higher-order desires and volitions should we form?

As we have seen, endorsement and identification are both part of the ability of *reflective self-evaluation* and at first sight can give an answer to both of these questions. The answer to the factual question is that when we form a certain higher-order desire and volition it is because we endorse or value it, and because we identify with it. And concerning the normative question, Frankfurt would say that we should put those desires effective in action by forming higher-order volitions which we endorse or identify with. Thus, endorsement and identification explain the process of forming higher-order desires and volitions more in detail. Nevertheless, it cannot solve the original problem, as the same question arises here: why do we endorse some of our desires and not others? Or to take the normative problem: which of our desires should we endorse and identify with? That is, though these help make the concept of reflective self-evaluation more elaborate, they cannot solve the problem, they just push it

one step further. We cannot effectively give the basis of a phenomenon by referring to other phenomena that themselves demand a basis.

The significance of the concept of caring, sufficiently redefined, is that it can be used to provide the basis of different higher-order phenomena. I will argue that it can be interpreted as an answer to Watson's criticism according to which a certain higher-order desire or volition must be grounded in something. Of course, because of his Platonism he would not accept caring as the ground for higher-order desires but I will argue it can fulfill that role under Humean assumptions.

Let us first look at how Frankfurt defines the concept of caring. What does it exactly mean to care about something according to him? He explains what caring about something is by contrasting it both with desiring something and finding it valuable.[111] We might desire something but that does not imply that we care about it. I might desire an ice-cream, but it surely does not entail that I care about eating ice-cream because it might be that this kind of activity is of no importance to my life whatsoever. It is also possible that though I recognize that something is intrinsically valuable, I do not find that thing attractive at all; it might even be totally indifferent to me. For example, I might judge that a healthy life is intrinsically valuable, but nevertheless go on smoking, eating unhealthy food and drinking all the same. In this case, though I judge that a healthy life is intrinsically valuable, I am not moved by it at all, because I personally do not care about healthy life, or in other words, I do not find it important to me at all. Thus, Frankfurt treats caring about something as equivalent to *finding it important* to us,[112] and claims that it is ultimately different from both desiring something and finding it valuable.

Frankfurt's point that caring has to be contrasted with both desiring something and finding it intrinsically valuable is quite interesting because one might think that caring

[111] See for example Frankfurt, "On Caring", in *Necessity, Volition and Love*, 158 or *The Reasons of Love*, 12.
[112] Frankfurt, "On Caring", 155.

involves *both* having the appropriate desire and the judgment that the object cared for has some kind of value. Such view could say that there has to be a proper mixture or mesh of the rational and the non-rational elements of caring, in this case we need a mixture of desires and value judgments. If one wanted to pursue this view, one would have to argue the following way. On the one hand, desire seems to refer to some brute driving force in us. At least this is the case in connection with desire for food, sex or survival. These are mere "unintelligent" or non-rational driving forces that are rooted in our biological nature. Other types of desires like desire for fame, money and the like might be similar, non-rational driving forces. On the other hand, finding something valuable seems to be something more intelligent than mere desire. When I say that something has value I make a rational *judgment* about that thing, namely that it has a certain objective value.

One might argue that caring should imply both desiring and value judgment. I cannot care about something without desiring it, but desire does not suffice either because I have to find the object of caring valuable, too. Desiring refers only to a non-rational driving force in me; it does not tell anything about what the thing I desire is like. I might desire something that I find as having no objective value; furthermore, I might judge it very negatively, and consequently not care about it. And it is also obviously true that I might find something intrinsically valuable without desiring it, and consequently, fail to care about it. The discrepancy between desire and value judgment comes from the fact that desire *only* refers to something *in me* (the driving force), while finding something intrinsically valuable *only* refers to something *outside of me* (the object and its characteristics). The happy case, of course, is when I *both* find something intrinsically valuable and desire it. Is this caring? It is certainly not for Frankfurt. His concept of caring, though involves both some kind of valuing and desiring, it does so in a very different way I described above. First, for him, caring does not involve any judgment about the value of the object cared for. As we will see, the concept

of value that is important for Frankfurt's philosophy of agency is a *subjective* one. The value of the object that we care about is derived from our caring about it, and does not imply objective value. Second, as we have seen, his concept of desire is different, since it is not merely a brute force: through second-order desires, we are able to evaluate ourselves reflectively.

One of the most significant aspects of caring is that it has a *temporal dimension*.[113] That is, it is similar to satisfaction inasmuch it is directed at the temporally extended presence of a given desire. According to Frankfurt, being satisfied with a desire means that we do not have any negative higher-order attitudes toward its presence and continuation, and as we have seen, for this reason satisfaction has an important role not only in Frankfurt's but in Bratman's theory of action as well. Now, caring also has a temporal dimension. As Frankfurt puts it, "[T]he outlook of a person who cares about something is inherently prospective: that is, he necessarily considers himself as having a future".[114] This is also something that differentiates caring from desires and beliefs; desiring or believing something does not, at least necessarily, entail the continuation of the given desire and belief. It is part of the inherent nature of caring that it persists for a given period of time.

Caring is also different from *acts of will*, the significance of which are exaggerated according Frankfurt's recent views.[115] These include decisions, choices and the like, which one consciously, voluntarily performs. I will argue later that as a result of this contrast between caring and acts of will, it is significantly different from all the higher-order attitudes I discussed in the previous chapters. While those are performed consciously and voluntarily, caring not necessarily is. Acts of will help us form intentions but they do not entail caring,

[113] As we have seen, Bratman also thinks that temporality is an important part of agency. For a recent detailed discussion of why diachronic agency is important see also Ferrero, Luca, "What Good is a Diachronic Will?", *Philosophical Studies* 144, Vol. 3, 403-430.
[114] Frankfurt, "The Importance of What We Care About", 83.
[115] Ibid., 84.

since one can intend something without caring about it.[116] It is a very important aspect of Frankfurt's thinking that what we care about is more central to our *will* than our intentions, or as he puts it, what someone cares about expresses "what his will truly *is*".[117]

As I mentioned, this is a significant shift in Frankfurt's thinking. We have seen earlier that Frankfurt himself emphasized the importance of decisive commitment which he understood as a *decision*. With the introduction of the concept of caring he moved to the opposite direction and claimed that the most central part of the will is something that we cannot directly control and voluntarily adopt by a decision.

Now, if caring is so different from acts of will, what is exactly the mechanism through which it works? When Frankfurt first discussed this concept in detail, in his "The Importance of What We Care About", he did not have a detailed description of how exactly the process of caring works. There he only said that caring is "constituted by a complex set of cognitive, affective and volitional dispositions and states".[118] However, later on he seemed to suggest that the former two of these elements are not necessary for caring.[119] It is not necessary to have any emotions or beliefs to care about something. On the other hand, caring necessarily has a volitional element. In Frankfurt's terminology this means that caring is connected to what we desire or want. His emphasis on the volitional element is hardly surprising and is in vein of his Humean theory motivation. Thus, caring about something is essentially a complex way of desiring or wanting it.

In a later paper, "On Caring", Frankfurt describes in detail how caring as a complex desire works.[120] In his example someone is about to attend a concert but a close friend of his asks him an important favor. Now, suppose that he cares about going to the concert. In that

[116] I discussed the concept of intention and its relation to Frankfurt's views on agency in my paper "Intentions and Agency", in Filozofia 64 (2009), 739-747.
[117] Frankfurt, "The Importance of What We Care About", 84 (italics in the original).
[118] Ibid.,85.
[119] Frankfurt, personal communication.
[120] Frankfurt, Harry G, "On Caring".

case, he *continues* to desire to go to the concert. This is connected to the temporally extended nature of caring, but it is important to stress that this continuation of his desire is not simply due to "its own inherent momentum" but is the result of the agent's own activity.[121] But how does one actively sustain a desire? Frankfurt thinks that this works through forming a higher-order desire, directed at the continuation of a given first-order desire. In the example, if the agent really cares about the concert, it means that he has a first-order desire to attend the concert plus a second-order desire that is directed at the continuation of this first-order desire. In other words, he wants to keep or sustain this desire. But this is not all that caring involves. Frankfurt also adds to the recipe of how to care his concept of identification. The full description of the music lover's psychology is thus the following: "[H]is caring about the concert would essentially consist in his having and identifying with a higher-order desire of this kind".[122]

In the next section I will look at how caring as a special kind of complex higher-order phenomenon changes Frankfurt's classic hierarchical theory of desires and his concept of reflective self-evaluation.

3.3. Caring and Reflective Self-Evaluation

In this section I will look at the relationship between Frankfurt's theory of reflective self-evaluation and his later theory of caring. How are these two related to each other? As we have seen, earlier he claimed that reflective self-evaluation is based on our ability to form higher-order desires and volitions. But caring is another important ability so we need to look at the relationship between these two. Does the ability to care contribute to our ability to form

[121] Frankfurt, "On Caring", 160. Ferrero also argues that "[O]ur status as temporally integrated agents is not to be taken for granted; we are not born to it and, once acquired, it must be sustained". Ferrero, 430.
[122] Ibid, 161.

higher-order desires and perform higher-order volitions or are these independent from each other?

Frankfurt's definition of caring as identification with a higher-order desire to sustain a given first-order desire clearly suggests that caring is itself an activity that belongs to our ability to reflectively evaluate ourselves. Thus, he seems to think that it is just one more element in the arsenal that makes up reflective self-evaluation. We have seen in the earlier chapters that since the introduction of higher-order desires and volitions, Frankfurt has been working on making his theory on reflectivity more and more sophisticated and detailed. He introduced and discussed in detail further reflective abilities, including decisive commitment, endorsement, identification and satisfaction. Now, it seems that caring is basically one more type of higher-order desire. At first sight, then, it seems that caring is just one more extension of the list of abilities that make up reflective self-evaluation.

But in what follows, I will try to show that Frankfurtian caring is not just simply a part of reflective self-evaluation. It has a special significance and the ability to care is required for the ability to form higher-order desires and volitions. Though we can conceptually separate these two abilities, in reality this separation is merely conceptual because the two are very closely tied together. Though we can imagine agents who do not care about anything and yet can form higher order desires, such creatures are merely conceptual possibilities because normally this ability is dependent upon caring. We are able to form higher-order desires because we are able to care.

First, let us recall Frankfurt's classic theory of the hierarchical theory of motivation. The significance of this theory was that in his early writings, Frankfurt tried to solve the problem of free will by developing this theory. By placing our desires in a hierarchy, he managed to explain how we can be persons and have freedom of the will using only the concept of *desire*. He claimed that if an agent's will or effective desire depends on a second

order volition he has free will. If somebody is moved to act by a first order desire while forming a second-order volition that makes this first-order desire his will, he has free will, but if he is moved to act by his first-order desire while his highest-order desire is contrary to his first-order desire he has not. In other words, free will means that an agent is free to want what he wants to want just as freedom of action means that one is free to do what one wants to do.

It seems that with this solution Frankfurt laid to rest the problem of freedom of the will, at least for himself, and started to deal with different topics, namely caring and most lately, love, which he understands to be the most important mode of caring. But in my view, his theory of caring seem to grow out directly from his hierarchical account of desires and thus these are closely connected. Caring and love are reflexive with regard to one's higher-order desires just as higher-order desires are reflexive with regard to one's first order desires. As we have seen, Frankfurtian caring and love are directed at sustaining some of our first-order desires through forming yet other higher-order desires. So, Frankfurt's theory about the importance of caring grows out from his hierarchical theory of desires. Caring has an important role in the cross-temporal organization of one's activities and as a result it leads to the reinterpretation of both the hierarchical theory of desires and the concept of reflective self-evaluation itself.

Frankfurt's agent who cares seems to be more developed than the agent who forms second-order desires but does not care about anything.[123] At first sight, it seems that there can be creatures that have higher order-desires but do not care about anything. If that is true, caring is not a condition of having higher-order desires. So what is the difference between the caring agent and the one who has merely second-order desires? Let me quote Frankfurt at length here to answer this question. He claims that even if we do not care about anything,

[123] In an exciting paper Bratman tells a similar story about a series of more and more complex and developed agents based on Gricean creature construction. See Bratman, Michael E., "Valuing and the Will", in *The Structures of Agency*, 47-67, reprinted from *Philosophical Perspectives: Action and Freedom* 14 (2000), 249-65.

> [W]e might …still want to have certain desires and to be motivated by them in what we do; and we might want not to have certain others and want not to be moved by them to act. In other words, our capacity for 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. Desires and volitions of various hierarchical orders would come and go; and sometimes they might last for a while. But in the design and contrivance of their succession we ourselves would be playing no concerned or defining role.[124]

The caring agent is more developed than the non-caring one because as he is able to shape and maintain his volitional identity over time and consequently he can have long-term goals and projects. Quite intuitively, it is preferable to be a creature that can do that.

But is caring merely an additional ability to forming higher-order desires and volitions? It seems that even though somebody who does not care about anything can have second-order desires and a coherent hierarchical structure of desires, there is something wrong with him: his volitions do not have a temporal dimension; he does not seem to be able to organize his actions cross-temporally.

Thus, caring is not simply one more tool to evaluate ourselves reflectively. Self-evaluation seems to be quite defective without a temporal dimension. Frankfurean caring then is the most primary form of self-evaluation, upon which other forms of self-evaluation such as the ability to form higher-order desires and volitions, endorsement, commitment and identification depend. Because of its temporal dimension, it makes self-governance possible for the agent. Thus, caring is the basis of autonomy of a certain kind. Being autonomous in this sense requires or implies Frankfurten free will, but not the other way round. One can have free will without organizing one's agency through caring but without the latter something is lost.

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

In the next section I will discuss how caring is related to reflective self evaluation and Frankfurtian free will. I will also discuss the theory of autonomy based on this concept.

3.4. Caring, Free Will and Autonomy

In this section I will look at the relationship between Frankfurt's theory of free will based on his hierarchical theory of desires and his later theory about caring.[125] My questions are the following: does the problem of caring has any relevance to the problem of free will? Does caring contribute to our freedom or we can interpret Frankfurt's theory of free will without the help of the concept of caring? What is the relationship between the psychological structure that grounds free will and the one that grounds caring and thus autonomy?

I will argue on two lines. First, I will try to show that caring contributes to our freedom by showing – in agreement with Frankfurt – that freedom is a matter of degree and caring makes it possible for us to enjoy it to a greater degree. For this we have to compare caring and non-caring free agents. The comparison will show that a caring free agent has a greater degree of freedom than the non-caring one. Caring increases our freedom by making us free with regard to more aspects of human life than we would be without it.

The second line of my argumentation goes like this: though we can conceptually separate Frankfurtian free will and caring, in reality this separation is merely conceptual because the two are very closely tied together. Though agents who do not care about anything have some amount of freedom (as showed less than those that do), such creatures are merely conceptual possibilities because free will is normally dependent upon caring. I will try to show that normally – though not necessarily – our freedom is grounded in our wholehearted caring. We are able to develop the psychological structure that grounds free will because we

[125] I will try to redefine the concept of caring in the last chapter, arguing that caring is primarily constituted by first-order desires with a temporal dimension and that the reflective level of caring is of secondary importance. In this chapter I will stick to Frankfurt's concept of caring, which requires and involves second-order desires of a specific kind. For a detailed argument for the possibility of caring without reflectivity see also Jaworska, "Caring and Full Moral Standing".

are able to care. On the one hand, once we developed wholehearted caring, the psychological conditions of free will are automatically satisfied. On the other hand, without caring the satisfaction of these conditions would a quite improbable.

If one is moved by a first-order desire because one has made a higher-order volition endorsing it, one is able to create a volitional coherence within oneself since endorsement implies that the agent is not opposed to his effective desire. Frankfurt calls the state of volitional coherence in which there is no internal conflict between our desires *wholeheartedness*. A wholehearted agent has free will because "[I]f there is no division within a person's will, it follows that the will he has is the will he wants."[126] As we have seen in *chapter* 1, free will for Frankfurt means that an agent is free to want what he wants to want just as freedom of action means that one is free to do what one wants to do. In wholeheartedness one is free to want what he wants to want because his will is not opposed by a contrary desire. Frankfurt thinks that limited creatures like us can have free will this way, that is, if wholeheartedness or volitional unity comes about in us. We cannot always make it happen that we have the will we want to have, because we do not have immediate voluntary control over our volitions and consequently we cannot make ourselves wholehearted voluntarily.[127]

Frankfurt's ideal agent who wholeheartedly cares and loves seems to transcend mere freedom at first sight. He does have Frankfurtian free will, of course, because the condition of having it is satisfied: he is wholehearted, and consequently he has the will he wants to have. But it seems to follow from Frankfurt's claims that it is not needed to care about anything in order to be free; caring is not a condition of having freedom of the will. So what is the

[126] Frankfurt, "The Faintest Passion", 101. To some extent, then, we are at mercy of the circumstances whether wholeheartedness comes about in us. Consequently whether we have freedom of the will or not is a matter of luck, which, according to Frankfurt, is a quite realistic conclusion.
[127] Ibid., 100.

difference between the wholeheartedness of the agent who cares about and loves things and the wholeheartedness of the agent who has merely free will?

We have seen in the previous section that according to Frankfurt, somebody who does not care about anything may still be able to perform higher-order volitions and consequently freedom of the will is not threatened by the lack of caring. I claimed that though in principle this is true, agents who have this ability and yet do not have the ability to care are merely conceptually possible, and there is reason to think that in the default case our ability to form higher-order desires and volitions is dependent upon the ability to care. In what follows I develop this claim more in detail by examining a non-caring agent who has Frankfurtian free will to demonstrate this claim.

Let me will compare a caring and a non-caring free agent. Superficially, an agent who does not care about anything, or even stronger, who lacks the ability to care still can form higher-order desires and volitions, and thus can create the psychological structure required for free will. But does it mean that such an agent can enjoy the same kind of freedom the one who cares can? It seems to me that the agent who cares is free with regard to more aspects of human life than somebody who does not care about anything.

Now, how does the caring agent's freedom transcend the freedom of the non-caring one? He transcends it because as he is able to shape and maintain his volitional identity over time and consequently he can be free with regard to his long-term goals and projects. Quite intuitively, a creature with this ability can enjoy a much greater amount of freedom than one without it. Without an ability to organize one's volitional identity temporally one's freedom would be compromised to a significant degree. An agent without this ability could possibly satisfy the condition of free will because he could make higher-order volitions. However, his freedom would have no temporal dimension and he would have no attitudes toward the continuation of his volitions.

Some might object that having or not having free will is an all or nothing affair; one either has it or does not have it. But as I have already noted in *Chapter 1*, some philosophers claim that having free will is a matter of degree. For instance, in Watson's framework, we can have various degrees of freedom depending upon the degree of overlapping between our motivational and valuational systems. Frankfurt also thinks that free will is a matter of degree and objects to Descartes' idea according to which the will is simple and indivisible and that it is perfectly free. In his paper "Concerning the Freedom and the Limits of the Will" he argues against Albritton who tries to defend Descartes' view. He compares the freedom to do something with the power to do it and states his case in the following way:

> But it seems to me that freedom is necessarily susceptible to variations in degree, and that the same is true of power. Whenever it makes sense to describe something as enjoying a certain freedom or a certain power, it also makes sense to ask how much of that freedom or power it enjoys; and if something has a characteristic about which the question of its amount or extent cannot reasonably arise, than that characteristic is properly understood neither as a kind of freedom nor as a kind of power. Freedom and power are essentially quantitative and open to comparisons of measure. One may have either more or less of them.[128]

Now, how are the power and the freedom of the will susceptible to variations of degree? The power of my will is limited simply because I cannot bring myself to want anything whatsoever. There are lots of thing that we simply cannot want and probably this is the reason why "[N]o one […] claims an immediate awareness or sense that the will has unlimited power".[129] Our power to want things is limited by our own volitional identity that makes it impossible for us to want certain things. Frankfurt calls this feature of our wills *volitional necessity* and claims that this limitation does not impair one's will because "it is grounded in the person's own nature".[130] Also, this kind of volitional necessity is a limitation of the power of the will that does not take away our freedom since it goes together with

[128] Frankfurt, Harry G., "Concerning the Freedom and the limits of the Will", in Frankfurt, *Necessity, Volition and Love*, 76.
[129] Ibid., 75.
[130] Ibid., 81.

wholeheartedness. Indeed, it implies wholeheartedness in the strongest possible sense because it makes an internal conflict impossible.[131] To take the famous Luther example,[132] Luther simply cannot want to renounce his views when he says "Here I stand I can do no other" because he wholeheartedly wants to go on with his project of reforming Christianity. We just do not have an absolute power of will that makes it possible for us to want anything.

The power of our will can be limited by volitional necessities. However, Frankfurt thinks that this does not imply the loss of its freedom. Agents like Luther do not have the power to want otherwise than they do.[133] However, somebody who is able to want a few things only might enjoy a bigger amount of freedom and be more autonomous than the one who is able to want much more things. Being able to want anything does not make us free; rather, such ability would make our wills impaired.

According to Frankfurt the more wholehearted we are, the more freedom we can enjoy. Now, this is why the agent who does not care about anything is less free than the one who does. He does not have any preferences about his volitional continuity and as a result, in each moment he has the power to want more of the alternatives than someone who cares. He might enjoy a greater degree of power of wanting anything whatsoever because his will is not limited by his long term commitments and goals. However, this does not make his will any more free. Rather, his will is impaired because he it does not determine his long-term goals and projects.

[131] Of course, one can question that question is necessarily bad and that wholeheartedness is what we should all reach. Kalis for example argues that ""internal conflict might be important for us not only as a step on the road to more wholeheartedness, but that it might be *in itself* important for being an authentic self". Kalis, Annamarie, *Failures of Agency* (PhD dissertation, Utrecht University, 2009), 177.

[132] Dennett introduced this famous example to the free will debate. Dennett, Daniel C, "I Could Not Have Done Otherwise – So What?" in *The Journal of Philosophy* 81, No. 10 (October 1984), 553-565.

[133] Of course, one might argue that Luther did have the ability to commit himself to a different life project in his past. Kane argues for example that he could have performed a self-forming act that could lead to his resoluteness, in which case his free will (and responsibility for his act) still depends on his (past) ability to do otherwise. See Kane, Robert, *The Significance of Free Will* (New York: Oxford University Press, 1996). He also argues that during the inner turmoil he went through in his youth "Luther was gradually building and shaping the character and motives that issued in his act" (Kane, 39). See also Erikson's biography, *Young Man Luther: A Study in Psychoanalysis and History* (New York: Norton, 1958) for details of Luther's identity crisis in his youth.

We can talk about degree in connection with the will in two different senses. First, it seems that it is not the case that we can equally want anything; it is not the case that, faced with a choice, we have the power to want any of the epistemic possibilities. Quite often, if not always, there will be alternatives that we just cannot bring ourselves to want. And because of this, we have only a certain degree of power of will to choose among the alternatives at hand. The more of the alternatives we can bring ourselves to want, the greater degree of power of will we can enjoy. However, the degree that we might have of this kind of power does not coincide with the degree of our freedom. If the will is limited to one certain choice it does not mean the loss of freedom, provided that the will is limited by itself. The actual Luther enjoys a greater degree of freedom than an imaginary one who stands there, hesitating about what to do. The latter has a greater degree of power of will; he *can* convert back to Catholicism or he *can* even become a Satanist or he *can* forget about religion and become a lawyer, etc. The actual Luther cannot want all these, but he has a greater degree of free will because of his wholeheartedness.

Luther enjoys a greater degree of freedom and autonomy also because his decision is stable, or in other words because it has a temporal dimension. He wholeheartedly refuses to renounce his views, because he keeps in perspective his long-term goal of reforming Christianity. His choice expresses what he cares about most in life. His case is a very good example to show that caring does contribute, or even, ground, our freedom. Just imagine a Luther that says wholeheartedly: "Here I stand, I can do no other" and a minute later he says wholeheartedly: "Ok, I'll become the advocate of the Pope now".[134]

[134] Sudden conversions of this might occur, like in the well-known case of St. Paul. However, I do not think that such changes occur out of the blue and because of this, the caring-based theory of free will can accommodate them. St. Paul might have realized on that occasion that he does not care about destroying Christianity that much, but he wants to solve something with his relentless persecution of them. He might have realized that Christianity is the best answer to solve his problems, and consequently he had some volitional continuity. For a detailed discussion of conversion see James, William, *The Varieties of Religious Experience: A Study in Human Nature* (New York: Simon & Schuster, 2004).

So it seems that caring contributes to our freedom. But are there at all free agents who do not care about anything? I will try to show that there are no such agents; and that consequently we cannot separate free will and caring from each other. On the contrary, our free will is grounded in our wholehearted caring.

It seems that even though somebody who does not care about anything can have free will, he can have it always only at a given moment. At each given moment he might will freely, and even, it might be the case that he wants the same thing in each moment. But it equally might be the case that he wants contradictory things in different moments and yet want them wholeheartedly. To go back to the example in the previous section, imagine an agent who might want to get married wholeheartedly at time t1 and want wholeheartedly to remain a promiscuous bachelor at time t2, and he might want to become a catholic priest at time t3, and these moments of time might follow each other in rapid succession. Though someone like this is wholehearted in each moment and consequently his freedom of the will is not threatened at all, he seems to have a problem. His problem is that there is no connection between his wholehearted volitional states. Of course, it might be the case that he just happens to want the same thing in each moment and have volitional continuity. But the connection between his volitional states at certain moments of time would depend merely on the spontaneous continuation of these states. His problem is that his volitional states can just as well change spontaneously, and even if he can remain wholehearted and free, he would be at mercy of where his spontaneous processes happen to lead him. Sooner or later such an agent would become, if not chaotic, at least unpredictable and unreliable, both for himself and for other people.

What is the exact difference between the wholehearted agent who cares and the one who does not? Each of them has his own volitional coherence, but while the one who has

only freed will has merely *synchronic* volitional coherence, the one who cares about things has it *diachronically*. Frankfurt explains this in the following way:

> Willing freely means that the self is at the time is harmoniously integrated. There is, within it, a synchronic coherence. Caring about something implies a diachronic coherence, which integrates the self across time. Like free will, then, caring has an important structural bearing upon the character of our lives. By our caring, we maintain various thematic continuities in our volitions. We engage ourselves in guiding the course of our desires. If we cared about nothing, we would play no active role in designing the successive configurations of our will.[135]

How exactly does caring achieve creating a diachronic volitional unity? The answer is simple: when we care about something, not only do we endorse a desire by a higher-order volition but additionally, we also desire to have this desire in the future. While a synchronically wholehearted agent is satisfied with a desire at a certain moment only, his diachronic counterpart is satisfied with it for the future, too. The main difference is that the wholehearted agent who cares has some reflexive temporal attitudes toward his volitional unity or wholeheartedness, while the one who does not care about anything does not have any attitudes like these.

Now, does not the temporal dimension of volitional unity figure in the explanation of the free will? Is it really true that an agent who does not have any attitudes toward the continuity of his volitional unity is able to will just as freely as his counterpart who does? Could there be such agents at all? In order to examine this problem, we have to look at the differences between these agents and whether these differences affect our intuitions about free will.

For this we have to look at all three main types of wholehearted agents: 1) agents without volitional continuity, 2a) agents with merely spontaneous volitional continuity but

[135] Frankfurt, *Taking Ourselves Seriously and Getting it Right*, 180.

without caring[136] and 2b) agents with volitional continuity and coherence that is the result of, or is maintained by, any modes of caring. It is important to stress that for this agent, caring results in *both* continuity and coherence. Caring itself does not make us coherent since we often care about contradictory things.

Thus, the agent who has Frankfurtian free will but does not care about anything has two versions: 1) and 2a).

We can call agent 1) a *chaotic agent*. Just like the other two, he has a wholehearted volitional unity at each moment of his life, and consequently he has freedom of the will during his whole lifetime. But he lacks continuity of his states, so it might well be the case that he has completely different volitions in each moment.

Agent 2a), whom we might call *spontaneously continuous agent* is luckier inasmuch as his volitions tend to prevail through time and consequently he wants wholeheartedly things not only in singular moments of time but over some periods of time. Probably many will share the intuition according to which this agent is luckier than the previous one. Just imagine: who would like to be the chaotic agent, that is, an agent with constant, unpredictable changes of his volitions?

But how can we be free without actively doing anything about our long-term goals? And after all, are the non-caring free agents realistic at all? If they are not, there are no free creatures like the chaotic or the spontaneously continuous agent. Let us look at the chaotic agent first. His mind seems to be completely disordered; and such a chaotic person could hardly be wholehearted, or at least it is quite implausible that such a person would develop this state. There cannot be creatures that are coherent in each moment but chaotic if looked at for a period of time. Chaos and disorder would prevail in a mind like that no matter how we look into it. This should be the case unless his mental states in each moment are totally

[136] On the revised concept of caring I will try to develop in the last chapter one can care about something without any reflective activity, though of course, caring usually involves reflectivity. Thus, on the revised view a certain spontaneous continuity of desires can be understood as caring.

disconnected from his mental states in all other moments. Without this disconnectedness his mental state at a certain moment would have at least some connection with his mental states in other moments and since these are incoherent, it is difficult imagine what sort of wholeheartedness would his state be. But if his mental states are really disconnected, we cannot talk about one person at all; a different person would exist in each moment. Each of them would have free will, but they would exist only momentarily.

The spontaneously continuous agent is more realistic since there tends to be some volitional continuity even in people who do not seem to care about anything. As I tried to show earlier in connection with Strawson's argument against narrative identity, even if some people might find it quite unimportant what they will want to do in the future, there will be some continuity in their volitions. That is, there is some degree of spontaneous volitional continuity even in creatures that do not care about anything. But just as an agent who spontaneously has no volitional continuity whatsoever is not realistic, an agent who happens to have wholehearted volitional unity spontaneously over longer periods of time seems to be also improbable to exist. Without caring about having the same volitional unity in the future an agent like this would be at mercy of his spontaneous processes and why would these point towards long-term volitional coherence? After all, they could equally lead to spontaneous changes in his volitions. But even if there are some people with a firm and spontaneously continuous volitional unity, it is hardly possible that they never reflect about it or never notice it and give it at least some consent. They would at least be satisfied that things are going the way they are for them, or otherwise, their being dissatisfied with the situation would disrupt their volitional unity. Being satisfied that one is naturally and spontaneously wholehearted for a longer period of time goes together with the lack of caring about changing the situation. Or, as Frankfurt puts it, "[S]atisfaction is a state of the entire psychic system – a state constituted

just by the absence of any tendency or inclination to alter its condition".[137] And if one's attitude toward oneself is that of satisfaction or acceptance, this already entails that one cares at least minimally about the continuity of one's volitional states, since, taken toward one's volitional continuity over time, these attitudes have a temporal dimension. And since agents who do not care about anything do not have any attitudes toward themselves of a temporal sort, those who accept their volitional continuity cannot belong to this type. The spontaneous agent could only lack having temporal attitudes toward himself if his mental states in each moment of time would be disconnected from his mental states in all the other moments, which would lead to the existence of a large number of momentary persons just as we saw it in the case of a version of the chaotic agent.

So it seems that neither of the non-caring free agents is realistic. On the one hand, the chaotic agent who is wholehearted in each moment is clearly not possible. On the other hand, agents with continuous volitional unity have at least some temporal attitudes toward their volitions and these constitute a minimal amount of caring in the Frankfurtian sense.[138] If these conclusions are true, it follows that there is nobody who cares about nothing and yet has free will. Of course, this does not mean that anybody who cares about something is free. In that case, since creatures that do not care about anything are implausible to exist, all of us would be free. The Frankfurtian condition of freedom is coherence, and consequently caring grounds freedom only inasmuch it is able to create coherence. We might care about contradictory things, and in that case we are not free. In other words, we have to care wholeheartedly in order to be free.

In order to understand how wholehearted caring and free will is related, we have to look at how synchronic coherence is 1) achieved normally and what it is 2) grounded in. Quite intuitively, it seems that whether we have synchronic volitional unity depends upon

[137] Frankfurt, "The Faintest Passion", 104.
[138] On my revised concept of caring such agents could also satisfy the conditions of caring, but for a different reason: I will argue that a certain type of spontaneous volitional continuity is sufficient for caring.

whether we have diachronic volitional unity. Or at least normally it is the case that we have volitional unity continuously for a period of time, and it just straightforwardly follows that at any *tn* moment of this period we have a synchronic volitional unity. Since our diachronic state is that of volitional coherence and unity, our state in any moments of it is also one. That is, the volitional coherence that we might have at a given moment is nothing but our diachronic unity in that certain moment.

How does this consideration help us to answer the questions about the origin and ground of volitional unity? The answer is simple: if synchronic volitional unity is part of a diachronic one as the above suggest, we can achieve it through creating a diachronic volitional unity. And if having free will depends on having synchronic volitional unity as Frankfurt claims, it further depends on creating a diachronic unity, and as a result, it is achieved through caring. That is, free will is normally dependent upon wholehearted caring.

Therefore, Frankfurian free will is achieved through, and grounded in wholehearted caring. At first sight this does not question Frankfurt's definition of free will; that remains still identical with synchronic volitional unity, as grounded in caring. But there is another way of looking at it. The above discussion about having temporal attitudes toward ourselves reveals that if there is volitional unity in us, it normally has a temporal dimension; consequently when we talk about synchronic volitional unity we talk about a diachronic unity minus its temporal duration. We are being wholehearted about having certain motives because we are wholehearted about having them over time. Diachronic volitional coherence or wholehearted caring is primary and having free will thus depends on caring about certain things wholeheartedly. Free will is essentially identical with wholehearted caring. The two are only conceptually different aspects of the human mind: one is volitional unity as looked at a certain moment and the other is the same coherence with a temporal duration.

Having synchronic volitional coherence means only that at a given moment, one is able to will freely, and this provides the agent only with a momentary free will. Now, what is wrong with being such a momentary agent? The problem is connected to personal identity. For Frankfurt, the core of one's identity is what one cares about. Since the momentary agent does not care about anything, he has no personal identity. But how can such an agent have free will? Or indeed, how can he have a will at all? If we define "will" as the agent's effective desire, a momentary agent can have a will, and consequently can have free will. In his recent book, *The Reasons of Love*, Frankfurt claimed that even though an agent who does not care about anything can have a will, his will is not "genuinely his own":

> If there were someone who literally cared about absolutely nothing, then nothing would be important to him. He would be uninvolved in his own life: unconcerned with the coherence and continuity of his desires, neglectful of his volitional identity, and in this respect *indifferent to himself*. Nothing that he did or felt, and nothing that happened, would matter to him. […] Of course, he might still have various desires, and some of those desires might be stronger than others; but he would have no interest in what, from one moment to the next, his desires and preferences would be. Even if it could be meaningfully said of such a person that he had a will, it could hardly be said of him that his will was genuinely his own.[139]

This suggests that agent's will cannot be genuinely his own unless he cares about at least something, or in other words, unless he has a volitional identity over time. Why is that so? The answer is simple: since personal identity is essentially about our own temporal duration we need to have temporal attitudes toward ourselves. If we have such attitudes we are able to make our will essentially our own or in other words we can make it a part of our identity. Our identity is created by our concern about our volitions.

It is important to note that Frankfurt uses the term "indifference" in the text above. The agent who does not care about anything is indifferent to himself. But then, he is just a *wanton* in this sense. Earlier Frankfurt defined wantonness as indifference towards one's

[139] Frankfurt, *The Reasons of Love*, 22-23 (italics added). Note that while earlier Frankfurt claimed that an agent who cares about nothing can preserve his ability to form higher order desires, here he claims only that some of his desires might be stronger than others.

desires. He said that an agent who always acts on his desires without forming higher-order desires concerning them is indifferent to his desires, or in other words, he is a wanton. An agent who lacks caring is also indifferent, and hence, a wanton, though in another sense: he is indifferent concerning the continuity of his desires. He might have higher-order desires but he is indifferent to what desires he will have in the next moment. He is not a wanton synchronically, but he is a wanton *diachronically*.

Now, can a wanton of any type have free will? After all, it was wantonness that Frankfurt thought to be the opposite of freedom. Given the wanton's indifference toward his volitional continuity it is hardly possible that he can have free will. Since he has no will at all concerning his volitional survival he cannot have free will regarding his volitional survival. Indifference means that one's volitional capacities are impaired; and this impairment results in the loss of freedom. For temporal creatures like us wholehearted caring is essential for having free will. Though it is possible to separate the two conceptually, it is not the case that freedom is realized through having volitional unity at the moment of wanting something. Since this unity is normally achieved through wholehearted caring, it is this state that makes us free. Furthermore, it seems that wholehearted caring is the primary phenomenon, and the volitional coherence that grounds free will is just the manifestation of it at the moment of willing.

It is an important feature of Frankfurt's concept of caring is that it is a *volitional* concept. What an agent cares about belongs to his volitional identity. This is how Frankfurt tries to establish the claim according to which we act autonomously when we act the way caring dictates to us. Our will is not determined by anything but itself. The volitional nature of caring makes it a better concept than value judgments which are merely intellectual and not necessarily volitional. It is also different from raw or first-order desires which do not tell what the agent really wants most. Second, it gives an account of *personhood*. The ability to

care about things presupposes having a level of psychic complexity that makes it possible for us "to have thoughts, desires, attitudes that are about our own attitudes, desires and thoughts".[140] For Frankfurt, without this reflexivity we would not care about anything, and we would be like animals, or we would be similar to the creature Frankfurt called a 'wanton' in his earlier writings. In the last chapter I will develop a different view of caring and personhood that does not stress the importance of reflectivity for caring.

To conclude this section, Frankfurt's concept of caring is successful inasmuch as it grounds both personhood and autonomy. But it is also very controversial because caring as Frankfurt describes seems to be an explanatory ultimate. There is nothing beyond it that could rationalize or justify it. Our caring is rooted partly in our biological nature and partly in our personal history. This gives rise to an important objection. Values become subjective in a way that, to take extreme examples, Hitler's and Mother Theresa's lives can be equally good, happy and fulfilling. In the next chapter I will look at Frankfurt's concept of love which he understands as the most important mode of caring, thus it is important to discuss it in order to address the objection about the subjectivism involved in caring.

[140] Frankfurt, "Freedom of the Will", 17.

Chapter 4: Frankfurt's Concept of Love

In the previous chapter I discussed Frankfurt's concept of caring and its significance for agency in detail. I tried to show that caring, has an action guiding role and creates a temporal continuity in our agency. As a result, caring can serve as the basis of a theory of *autonomy*. Recently, Frankfurt developed his views on caring and agency further by discussing the problem of love in detail. In his book *The Reasons of Love,* he developed a new and genuine concept of love. As he understands it, love is primarily a mode of caring, which, as we have seen, is in turn a complex mode of desiring or wanting. Thus, love is primarily neither cognitive nor affective but *volitional*.[141] That is, the essence of love is that the lover's will is determined in relation to the beloved in a certain way. When we say that an agent loves a certain individual it means that he cares for her in a particular way: he has a disinterested concern for her. Most recently, in his book, *The Reasons of Love*, Frankfurt developed a very controversial view according to which love has a centrally important role in practical reason.

[141] Frankfurt, *The Reasons of Love*, 42.

In this chapter I will discuss Frankfurt's concept of love and its significance for agency. My primary concern is not the concept of love itself, so the main goal of the chapter is not giving a correct definition of the concept of love. Rather, I am going to examine how love influences our agency and practical reasoning. I will also discuss and criticize Frankfurt's subjectivism about values. Especially because of his subjectivism I find Frankfurt's views on love and agency unsatisfactory and I will criticize him in detail. In the first section I will focus on how Frankfurt's controversial concept of love is related to the problem of practical reason.

4.1. Love and Practical Reason

At first, Frankfurt's claim about the importance of love for practical reason might be surprising. But when love is present, it does have a central importance for our decisions and actions. And intuitively, love should play a more important part in our actions. As Haji and Cuypers argued recently, "[O]ur world would be far better if love and care were emphasized in our dealings with others", and as a result, "the sort of normative agency associated with love should be of singular importance in the normative duty of our children".[142] One might argue that acting out of duty or prudence is equally, if not more, important. But Frankfurt argues that the role of morality and rationality in our actions is exaggerated. Whether or not this controversial claim is correct, it seems to be true that the normativity involved in caring and love is not discussed enough and that Frankfurt's contribution to this subject is of seminal importance.

Frankfurt's book on love begins with a chapter focusing on the following question: "How should we live?" Thus, the focus of Frankfurt's discussion is *practical reasoning*, which refers to "deliberation in which people endeavor to decide what to do, or in which they

[142] Haji, Ishtiyaque and Cuypers, Stefaan E., "Moral Responsibility, Love and Authenticity", *Journal of Social Philosophy* 36, No. 1 (Spring 2005), 124.

undertake to evaluate what has been done".[143] But though Frankfurt talks about *normativity*, his focus is not morality or prudence. Notoriously, he believes that morality is not as important as philosophers often think. In his view, it is much more important to live a meaningful life than morality as such. Also, though Frankfurt attempts to develop a theory of practical reason, his view is not primarily a theory of rationality. In a sense, we could say that it is not a theory of the "head" but that of the "heart" instead. What is best for us to do in a certain situation is not decided by cool-headed reasoning but what our hearts dictate us to do. Thus, one might label this view "romantic" in the sense that it claims that we should follow our heart rather then our reason.[144] But more obviously, Frankfurt's theory is Humean as regards the *instrumental* role of reason. Reason itself does not tell us what to do; rather, it helps us achieve what we already want.

Frankfurt's theory of practical reasoning is normative in a way that differs from traditional normative theories. Its task is not to answer the normative question by telling what we should do or how we should live, but instead, it asks us to answer first the *factual* question about what it is that we *already* care about in life. An individual cannot answer the normative question about how he should live without first trying to understand what he *already* finds important in life. So, in an ideal case Frankfurtian practical reason is based on self-knowledge. After one gains insight into what one really cares about, one has the following task: one has "to be decisively and robustly confident in caring about it".[145]

What is exactly the normative claim of Frankfurt's theory of practical reason, then? It is simply that once you know what you care about, you have to commit yourself consciously to it. It is a quite important and controversial feature of Frankfurt's thinking about practical reason that he does not say the following: once you know what you care about, try to answer

[143] Ibid., 6.
[144] Note however that as I will discuss later, romantic love is not a good instance of love for Frankfurt. Also, for him, love need not be emotional.
[145] Ibid., 28.

the question whether it is really *worth* caring about that thing. There is no "pan-rational" objective stance that would tell us what we should find worthy of caring about:

> "In order to carry out a rational evaluation of some way of living, a person must first know what evaluative criteria to employ and how to employ them. He needs to know what considerations count in favor of choosing to live in one way rather than in another, what considerations count against, and the relative weights of each."[146]

> "The trouble here is a rather obvious sort of circularity. In order for a person to be able even to conceive and to initiate an inquiry into how to live, he must already have settled upon the judgments at which the inquiry aims. Identifying the question of how one should live – that is, understanding just what question it is and just how to go about answering it – requires that one specify the criteria that are to be employed in various ways of living."[147]

Frankfurt puts the problem in several other ways but these quotations show the problem with views that claim that we can have a rational answer to the question "How should I live?" We cannot try to answer this question unless we already have some evaluative criteria that tell us what to regard the most important in life. Frankfurt lists some of the usual candidates for these evaluative criteria including personal satisfaction, pleasure, creativity and morality. Thus, if I answer to the question "how should I live?" by saying that "a creative life is the best", it shows that I already have criteria that picks creativity as the answer, so in some sense it is creativity that I already care for the most.

For a Frankfurtian, then, the answer to the question about how we should live must be based on the factual answer to the question about what we already care about. Now, what we care about can be devided into two groups: (1) concerns that have a *biological basis* and (2) concerns that are rooted in our *personal history*. Concerns as caring about our survival and our children have an instinctive, biological basis. Besides, we care about several other things as a result of an interplay between several factors including our upbringing, environment and experiences or shortly, our personal history.[148]

[146] Ibid., 24.
[147] Ibid., 24-25.
[148] Ibid., 47.

On the one hand, Frankfurt's view involves some kind of *naturalism*; he even goes as far as treating parental love, which obviously has a biological basis as the example of ideal love. On the other hand, this view brings in *contingency* as a major factor in our lives. This contingency is rooted in the fact that we have only a limited degree of control over our volitional lives. This view involves that we should accept the contingent features of our lives. That this does not threaten our freedom our autonomy is of course controversial in itself. And of course, we should not always simply accept things the way they are; some things need to be changed. Of course the picture is further complicated by the fact that it is not easy to determine what it is that we can and cannot change. Since my solutions to these problems involve diverging significantly from those of Frankfurt I will discuss them in detail in part III where I will try to develop a revised version of the hierarchical theory.

4.2. Love and Emotion

Before I start discussing in detail Frankfurt's concept of love, let me look at one of its important features. As I have mentioned earlier, he thinks that love is primarily a mode of caring, which is in turn a complex mode of desiring or wanting. Thus, love is primarily a *volitional* phenomenon and not an affective state. Its essence is that the lover's will is determined in relation to the beloved. As a result, Frankfurt's concept of love is not emotional: though it can, and usually it does, love does not necessarily involve strong emotions. Rather, the relevant configurations of the will are sufficient.

The claim that love need not involve affection might be surprising for many at first sight. Indeed, it is quite common to say that love is an emotion or at least that it usually involves emotions. However, there are some good reasons why Frankfurt avoids emotions in his definition of love. First, one reason might be that it is strategically important for him to do so. According to some prominent theories, emotions have some kind of cognitive content in a

sense that desires might not. Having desires implies only having a positive or negative attitude for something but apart from this, emotions might involve implicitly a positive or negative *judgment* about that thing.[149] Further, these judgments might precede or constitute the emotion itself. For example, I might judge, or at least perceive that a particular dog is dangerous and feel fear thereby. However, it is crucial for Frankfurt that judgments are not the cause, or even constitutive of, but the result of love: the lover judges his beloved valuable but the particular value attached to love is not what brings it about but rather, it is only the result of it.

I will not discuss in detail whether or not emotions involve any value judgment, but later on I will discus the problem of value and love in detail. And apart from this problem I think there is a more important reason why Frankfurt is reluctant to define love in terms of emotions. Some authorities of the subject think that putting too much emphasis on emotion while trying to define love is due to a misunderstanding of its nature. According to this misunderstanding love is simply a positive emotion. Criticizing this misconception, Erich Fromm points out that love is not just simply a pleasant emotion that one feels, but a much more difficult and demanding thing, and for this reason he calls it an *art*.[150] If love was merely a pleasant emotion, it would be an easy thing to love; it would be simply sufficient to feel the emotions constitutive of love. Love however, is much more difficult than that and requires the person's effort. Also, besides being a difficult endeavor, love involves that the lover makes himself vulnerable to loss. As a consequence, it requires certain abilities and maturity to face and accept loss and this is one more reason why it is not simply an emotion.

[149] For some eminent cognitivist views on emotions, see for example de Sousa , Ronald. *The Rationality of Emotion* (Cambridge, Mass: The MIT Press, 1987), Nussbaum, Martha C, *The Upheavals of Thought: The Intelligence of Emotions* (Cambridge: Cambridge University Press, 2001) or Solomon, "Emotions and Choice", in *Not Passion's Slave: Emotions and Choice* (Oxford: Oxford University Press, 2003, 3-24, reprinted from *Review of Metaphysics* 28, No. 1 (September 1973). Since this paper by Solomon, cognitivism about emotions became mainstream. However, Solomon himself stresses that we should not overintellectualize emotions. For another important discussion of the role of emotions in agency see also Damasio, Antonio R., *Descartes' Error: Emotion, Reason and the Human Brain* (New York: Avon Books, 1994).

[150] Erich Fromm, *The Art of Loving* (New York: Harper, 1956).

Finally, Frankfurt avoids emotions in the definition of love to differentiate it from *romantic* love. The latter obviously involves strong emotions and because of the nature of these emotions Frankfurt thinks that it is not the best example of love as a disinterested concern for the beloved. Let me quote Frankfurt at length here:

> It is important to avoid confusing love […] with infatuation, lust, obsession, possessiveness, and dependency in their various forms. In particular, relationships that are primarily romantic or sexual do not provide very authentic or illuminating paradigms of love as I am construing it. Relationships of those kinds typically include a number of vividly distracting elements, which do not belong to the essential nature of love as a mode of disinterested concern, but that are so confusing that they make it nearly impossible for anyone to be clear about just what is going on. Among relationships between humans, the love of parents for their infants or small children is the species of caring that comes closest to offering recognizably pure instances of love.[151]

In Frankfurt's view the vivid emotions that seem to be constitutive of romantic love cannot be part of genuine love. He also suggests that these intense emotions are the essence of romantic love and therefore it need not involve a disinterested concern at all. In contrast, though genuine love can be, and typically is, accompanied by emotions, these are not a part of its essence. So it seems that for Frankfurt genuine love and romantic love do not belong to the same category.

In my view, however, Frankfurt's view on romantic love is unduly stern. As we have seen above, he seems to suggest that romantic love, given its turbulent and distracting nature, is not even an instance of love. However, I am not convinced that this is so. In fact, romantic love has a lot in common with Frankfurtian love. Most importantly, romantic love is more complex that Frankfurt's text above suggest. It is not the case that it merely consists of intense emotions; and as for example Green argues, instead of defining romantic love as emotion, we should treat it as a complex set of desires.[152] If this is true, it follows that romantic love has a similar nature to that of Frankfurtian love. That is, they are both conative

[151] Frankfurt, *The Reasons of Love*, 43.
[152] Green, O. H., "Is Love an Emotion?" in *Love Analyzed*, ed. Lamb, Roger E., (Boulder, Colo: Westview Press, 1997) 209-225.

or as Frankfurt puts it, volitional. In other words, they are constituted by certain desires. Conceived this way, the primary difference between romantic love and other forms of love is that due to its erotic nature, the former involves desires that differentiate it from other kinds of love. However, this does not imply that it does not have a common essence with love in general. True, romantic love can give rise to certain excesses and even pathologies as for example jealousy and possessiveness but these can appear even in cases of what Frankfurt thinks to be the most exemplary form of love, namely parental love. Parental love can involve some of the negative features that Frankfurt listed above and lack a real disinterested concern. Besides, I believe that romantic love has the benign features of love inasmuch as it usually involves a disinterested concern for the beloved.[153] Of course, one can come up with destructive, pathological examples of romantic love but again, this does not prove its destructive nature. In the ideal case neither romantic love nor parental love involves destructive tendencies and both involve a disinterested concern.

To sum up this point, love has lots of variations and they are all partly constituted by some special desires due to the nature of the given relationship that they involve. Romantic love, parental love, friendship and other types of love all have certain desires as their distinguishing features. Nevertheless, they have one important feature in common: a disinterested concern for the beloved. As a consequence, I will not exclude romantic love from my discussion as Frankfurt does, and I will cite some authors who focus primarily on romantic love.

In the next section I will discuss Frankfurt's definition of love as one of the most important modes of caring in detail.

[153] See for example Green, 216 or Soble, Alan: "Union, Autonomy, and Concern", in Lamb, 65-93.

4.3. The Definition of Love

We have seen in the previous two sections that according to Frankfurt, cognitive and affective elements are not necessarily constitutive of love but rather, they only accompany it or are the results of it. In his view, love has these three core features: 1) it is a *disinterested* concern for the beloved, 2) it is *personal* and 3) we have *no voluntary control* over it. Let us look at these in detail.

The first feature defines how love constrains our will: loving has to do with "a configuration of the will that consists in a practical concern for what is good for the beloved".[154] If I love an individual I have certain dispositions to act, namely I want to benefit my beloved and I want my beloved not be harmed. Now, it is quite crucial that these concerns for the beloved are *disinterested*. I do not have them to promote some other goal, but the flourishing of the beloved is desired only for its own sake.

The second feature of love is that it is strictly *personal*. The concerns described above are concerns for my beloved and are concerns *only* for her. The beloved cannot be replaced and substituted by another individual even if that individual were very similar to the beloved. This feature of love is explained by the fact that "[T]he significance to the lover of what he loves is not that his beloved is an instance or an exemplar".[155] The beloved is not an example of a class of individuals with certain particular features that the lover finds important. If that would be the case, love would be similar for example to the possession of useful objects. For example, if my bike gets stolen, I might want a very similar bike to replace it. Or to take Frankfurt's example, love is very different from things like a desire to help the sick and the poor. Someone who has this desire might find any sick or poor individual to be a satisfactory

[154] Frankfurt *The Reasons of Love*, 43.
[155] Ibid., 44. I will come back to this claim in the next section while discussing the problem of *fungibility* or the view that if love is based on some qualities of the beloved he is replaceable by another individual who possesses the same qualities.

object of his care. Humanitarian love, because of its impartial nature, can be directed at any individual and in this sense its objects are "replaceable". But while in these examples the objects of our concern are replaceable, in case of love the beloved cannot be substituted by any other individual.[156]

And the last, third feature of love is that "it is not under our direct and immediate voluntary control".[157] Love cannot be a matter of conscious decision; we cannot simply choose to love someone. Frankfurt compares this to things that we cannot help caring about as for example staying alive. It is simply a *volitional necessity*[158] that follows from our biological nature that we care about staying alive. When our life is in danger we cannot help making the appropriate steps to defend it against the danger because we cannot help wanting to stay alive. In Frankfurt's view love is similar: "what we love and what we fail to love is not up to us".[159] Rather, love involves volitional necessities. This concept is a very important for Frankfurt; indeed, it is the very heart of his concept of love. A volitional necessity is something that constrains the will from within itself, and thus is the basis of Frankfurtian autonomy. Therefore, it is very different from compulsion which, as a result of its coercive nature, constrains the will from the outside and consequently defeats it. Thus, just like caring in general, love makes agents autonomous in the sense that when they act out of love, they are not determined heteronomously by something outside themselves, but only by something within themselves.[160]

[156] Thus Frankfurtian love is personal and *partial*, and thus is very different from humanitarian, impersonal and *impartial* love. The same is true of Frankfurtian caring, and this feature might be a reason why caring and love are not the basis of morality but rather, are in a sense contrasted with it.
[157] Ibid.
[158] Watson also discusses the concept of volitional necessity. See Watson, Gary "Volitional Necessities", in *Agency and Answerability*, 88-122.
[159] Frankfurt, *The Reasons of Love*, 46.
[160] It is interesting to note here that though Frankfurt is opposed to Kantianism, his claim about autonomy is surprisingly similar to it. Willigenburg notes that "Frankfurt holds, in agreement with a Kantian line of thought, that the will can only be necessitated by what originates from inside the will itself". Willigenburg, Theo Van, "A Non-Reductive Analysis of the Normativity of Agent-Relative Reasons", *Ethical Theory and Moral Practice* 8 No.1 (April 2003), 52.

But if it is absolutely not up to us whether we come to love a certain individual or not, how exactly does loving come about? Frankfurt has the following very disturbing answer: "[W]hat we love is shaped by the universal exigencies of human life, together with those other needs and interests that derive more particularly from the features of individual character and experience".[161] In the end, what we love is determined by "biological and other conditions, concerning which we have nothing much to say".[162] But actually, Frankfurt has a bit more to say about that. Frankfurt suggests that the reason why we come to love an individual is that we have certain *needs*.[163] His claim seems to imply that on the one hand, we have *biological* needs, on the other we also have *psychological* needs that are rooted in our character and in our experience. At first approximation this implies that we have the love objects that we have because they are certain objects with certain characteristics, such that, it is *them* that we need to satisfy our biological and psychological needs. If this is true, we attach some kind of value to our love objects at least partly on the basis of what characteristics they have objectively.

However, Frankfurt insists that as in case of caring in general, it is not the inherent worth of the object that moves us when we come to love something but we come to value only by loving it. I will discuss this problem in detail in the next section.

4.4. Love, Worth and Value

In this section I will discuss the problem of worth, value and love. The main question is this: does value precede caring and love or is it derivative of it? In other words, do individuals have a value that precedes and is independent from, caring and love or is value created by

[161] Ibid., 47.
[162] Ibid., 48.
[163] In an earlier paper Frankfurt already discussed the problem of *volitional needs* in connection with volitional necessities, and he argued that "meeting needs merits priority over satisfying desires". See Frankfurt, Harry G., "Necessity and Desire", *Philosophy and Phenomenological Research* 45, No. 1 (September 1984), 7.

caring and love itself? *Appraisal accounts* of love argue that love is created by the value that an individual has prior coming to love her and *bestowal accounts* of love claim that the value of the beloved is created by loving her.[164]

However, I see no reason why we should follow one of these lines. Why should all value connected to love be rooted either in the lover or the beloved? Rather, I will argue that we can talk about two types of value in connection with love: 1) *worth*, or a specific type of value that precedes or contributes to the development of love and 2) *derivative value* that is derivative of love or in other words the value the beloved acquires as a result of love itself. I will argue that both types of value can be divided further: 1a) everybody has an *intrinsic worth* to be loved and 1b) there is a *relational* worth that contributes to love. In the second group, we can find 2a) value that is constituted by the *relationship* itself and 2b) values that love can *bring about*. Thus, in what follows, when I talk about the worth of an individual I will mean its value that contributes to the development of love and when I talk about the derivative value I will mean the value it has as a consequence of love.

Let us look at the concept of *worth* first. Frankfurt straightforwardly refuses the view according to which love is a response to the perceived worth of the beloved.[165] He admits that though it might come about this way, it typically does not. We can come to love something even if we perceive that it is worthless. What's more, we can come to love an individual even if we recognize that he is "utterly bad".[166] Frankfurt's basic claim is that the perception of the value of an object is not a *formative* or *grounding* condition of love, but rather, the relationship is just the other way round: "what we love necessarily *acquires* value for us *because* we love it".[167] Thus, Frankfurt's view can be characterized as a bestowal account of love: by loving her, the lover bestows some value on the beloved.

[164] For a recent detailed discussion of these different views on love, see Helm, 21-34.
[165] Frankfurt, *The Reasons of Love*, 38.
[166] Ibid.
[167] Ibid., 39, italics in the original.

Frankfurt defends his thesis according to which the value of the beloved is derived from loving her with the help of an example that also serves as an ideal for love: parental love. He refers to his very own love for his children: "[I] can declare with unequivocal confidence that that I do not love my children because I am aware of some value that inheres in them independent of my love for them".[168] It is obvious for a parent that he would love his child in the same way regardless of what valuable qualities he has or has or does not have. The parent would love his son or daughter even if he or she was the worst person one could imagine confirmed by the phrase "only a mother could love him".[169] Though Frankfurt admits that we can say that somebody or something is unworthy of our love, for the most ideal type of love, parental love, it cannot count as a consideration against love.

Now, Frankfurt thinks that this relationship between love and value holds not only in case of parental love but quite generally. I believe that the reason why he stresses that value does not precede but is the result of love is an important feature of love I have not talked about so far, namely that love is *unconditional*. We do not love an individual *because* she has certain properties and it would not count as genuine love if we made our love conditional upon these characteristics. This is why parental love is a good example here: the parent does not love her child because she is clever, talented or well-behaved and it would obviously be wrong to deny love from the child because he lacks any of these. It would be even worse if the parent told her child that she will not love her unless she complies with some requirements. This kind of conditional "love" is obviously not genuine love. Thus, Frankfurt is right to say that a parent should love his children even if they "should turn out to be ferociously wicked".[170] Genuine parental love is unconditional.[171]

[168] Ibid.
[169] Helm discusses the case of Aicha el-Wafi, who, after having heard that her son, Zacarias Moussaoui had participated in the 9/11 attacks, declared that he was no longer her son. But reportedly, she changed her mind later on, saying in an interview that "he's my son, and I love him no matter what". See Helm, 195-196.
[170] Frankfurt, *The Reasons of Love*, 39.

Nevertheless, one might argue that though parental love does not depend on objective qualities, in other types of love, this is not always so. One could claim that though we usually love persons because for some reason they are worthy of our love, parental love is an exception because there is a biological basis for it that makes it impossible for the parent not to love the child even if he finds him unworthy of love. To see if this is true, let us look at another example. Suppose a friend of yours is "ferociously wicked": he drinks, takes drugs and commits crimes. And suppose that somebody gives you the advice to break with this person because you just should not love him. But suppose now, that your son drinks, takes drugs and commits crimes. In this case my advice that you should not love him seems to be wrong. Whereas in the first case stopping to love your friend would be arguably acceptable or even preferable, it would not be in the second.

Now, what does all this prove? Does it prove that after all, parental love is just an instinctive urge and that in other cases love should be guided by considerations about whether the object of love is worthy of love? Or rather, does it prove that parental love is a higher instance of love that is closer to ideal love? I think it shows that whereas it is acceptable or permissible to terminate a friendship as a result of some considerations,[172] parental love should not be responsive to these. And one might argue that the closer other types of love get to this ideal, the more genuine they become. To go back to the example above, what would be the case if you loved your friend "no matter what", in the same way a parent loves his child? How should we evaluate your love? Suppose you refuse my advice and say that you will not brake with your friend because you want to help him to change his unhappy way of living.

[171] Of course, unconditional love is not limited to parental love. For instance, Rosati notes that "unconditional love is not unique to the parent-child relationship" and we might love a great many people "no matter what" *and* without loving that as if they were our own children. Rosati, Connie S., "Preference-Formation and Personal Good," (Royal Institute of Philosophy Supplements 81, Supp. 59, 2006), 57.

[172] Though of course, once you are involved in a friendship, it might be argued that it is your duty to help your friend even if you have to face some objectionable qualities in her. Though these might count as a reason against loving her, the friendship itself constitutes further reasons that might override them. I will discuss the reasons created by a relationship later on.

You are just unable to leave him and let his situation become even worse. You are not sensitive to the consideration that he is "unworthy" of your love. Many people would say that you are doing something good; you should not just leave your friend to his bad faith. And this would suggest that Frankfurt is right. Love is more ideal or should be evaluated more positively if it is not based on the worth of the person. Therefore, it could be argued that the closer love gets to unconditional parental love, the more praiseworthy it becomes.

Another important objection to the view according to which worth is constituted by objective qualities is the *fungibility problem*. As Helm puts it, "to be fungible is to be replaceable by another relevantly similar object without any loss of value".[173] Thus, if we love someone because of her qualities, it implies that she is replaceable by another similar individual. Also, since the quality or appraisal view seems to involve that love is justified and sustained by some objective qualities, there are a number of disturbing consequences. For instance, if you notice that your beloved has changed in some respect, this might lead you to stop loving her, which, at least to some extent, is apparently contrary to our experience. People change, but this does not have to lead to such consequences. Another problem is that if you love someone because of her qualities, why aren't you supposed to love everyone else with these? I will not discuss these problems here in detail;[174] I only want to point out that a virtue of Frankfurt's view is that it can account for the fact that the beloved cannot be substituted by another individual.

But though this shows that Frankfurt is right in the sense that love should not be conditional on the qualities of the beloved and that the beloved should not be replaceable by another individual, it still does not imply that worth of an objective kind does not play an important role in the development of love. To go back to the example again, you could say that I am mistaken to say that your friend is unworthy of your love. After all, you know him

[173] Helm, 20.
[174] See Helm for a detailed discussion of the problem.

better. And you know that though his lifestyle suggests that he is a "bad" person, you know that this is superficial, and that he is potentially a much better person; he has many valuable traits latently and you want to help him to make these more effective. And certainlty, there is some intuitive truth in the appraisal account of love. Most of the time, when we come to love someone, this is at least partly due to her qualities, for example her sense of humor, intelligence, bravery or any similar qualities.

Susan Wolf also thinks that there is worth that is not merely derivative of love and she criticizes Frankfurt's concept of caring and love and his subjectivism about values in a recent paper.[175] She refuses Frankfurt's claim according to which the worth of the object of care or the beloved is of no central importance. She claims that an individual has to have objective worth, and that this is not simply a minimal condition but there is a "proportionality requirement" according to which "[P]eople should care about only what is somewhat worth caring about; and how much people should care about things, both in themselves and relative to other things they care about, depends somewhat on *how* much worth caring about the objects in question are".[176] More precisely, she thinks that there are three sorts of considerations that one has to take into account when thinking about the question whether he should care about or love an object: "whether (and how much) the object in question is itself worth caring about, whether (and how much) the person has an affinity for the object in question, and whether (and how much) the relation between the person and the object has the potential to create or bring forth experiences, acts, or objects of further value".[177] Therefore,

[175] Wolf, "The True, the Good and the Lovable", in *Contours of Agency*, 227-244, ed. Buss and Overton. The paper was written before the publication of *The Reasons of Love*, so Wolf focuses on an earlier paper by Frankfurt: "Autonomy, Necessity and Love", in Frankfurt, *Necessity, Volition and Love*, 129-154
[176] Wolf, "The True, the Good and the Lovable", 232.
[177] Ibid., 235. In his comments about Frankfurt, MacIntyre argues for a similar view. He states that "[T]o be a mature agent is to have educated one's feelings appropriately. It is to care the right amount for the right people and things in the right way at the right time and place". See MacIntyre, Alasdair, "Comments on Frankfurt", *Synthese* 53 No. 2 (November 1982), 292. This claim is obviously contradictory to Frankfurt's views since it implies that we have to apply a standard (one that is independent from caring itself) to the way we care.

she argues that it does make sense to give advices to someone for example, to find a more worthy friend or partner.

What is it that makes an individual worth of care and love? Wolf argues that the individual has to be objectively good, or lovable. But can we really determine what the worth of each individual is to be loved? It seems that she suggests that we somehow weigh reasons and deliberate about loving or not loving an individual. Wolf's view seems to imply that we can somehow evaluate individuals against some kind of standards to see whether they are lovable or not. In this sense it involves a quality view of love that makes love conditional upon qualities of the individual. But as I will try to point out, though in a sense it is objective, *lovability* does not depend on the evaluation of an individual against a certain standard.

So it seems that both Frankfurt's and Wolf's view, that is, both bestowal and appraisal accounts of love are right in some sense. That is, Frankfurt is right to think that love is not conditional upon evaluating an individual against a certain standard and that value is created through love. But as I have been trying to show above, it cannot be the case that all value connected to love is merely subjective; love has to have some kind of objective basis. At the beginning of this chapter I suggested solving this problem by dividing value connected to love into 1) *worth*, that precedes love and 2) *derivative value* that the result of it. In what follows I will discuss this division in detail. As I have mentioned, we can talk about worth in two senses: 1a) *intrinsic wort*h and 1b) *relational worth*. To look at the first, my claim is that everyone has an intrinsic worth to be loved. To see what exactly this means let me quote at length Darwall who makes a distinction between merit and worth:

> The contrast I have in mind between what I am calling "merit" and "worth" is that between a kind of value (*merit*) that persons and actions have in being worthy of admiration or emulation, on the one hand, and a kind of significance, importance or "mattering" (*worth*) that something can have by virtue of being appropriately deemed intrinsically significant or important, for example, as an appropriate object of care or (recognition) respect.[178]

[178] Darwall, Stephen: *Welfare and Rational Care* (Princeton: Princeton University Press, 2002), 78.

> The value someone has just by virtue of being a person […] is a kind of *worth*, dignity or "mattering," a value status that makes appropriate certain forms of valuing conduct toward the person and certain feelings that are *as of* someone who is to be treated and regarded in those ways.[179]

Each person has a dignity and thus deserves respect. Velleman's view is similar: he claims that when we love a person, "we are responding to the value that he possesses by virtue of being a person".[180] He thinks that being a person is linked to the capacity to be autonomous and everybody has an intrinsic value as a result of this capacity. Of course, in case of love the way we respond to this capacity is very different from the way we do when we respect someone. But the point is that all persons, just by virtue of being persons are worthy of respect and love. Velleman stresses that this value is not comparative. Persons have a *dignity* as a specific form of value as different from price. The price of an object is a value that we can compare with the price of other objects; but dignity is different since it is not a comparative value.[181] As a result, in this sense everybody deserves love regardless of any standards, moral or non-moral.

Any person has dignity or worth; in other words, any person "matters" and deserves to be valued in different ways, that is, deserves respect, care and love. Thus, in this sense worth is not an evaluation of the agent against a certain objective standard. An agent's worth to be cared for is independent from any facts about him relating to a certain set of standards, and thus if an agent fails to comply with any moral or non-moral standards, he does not lose his intrinsic value to be cared for or loved. This is one of the reasons why a quality theory of love cannot work. It is not a certain set of qualities that makes an agent worthy of love; rather, the intrinsic worth that he has.

[179] Ibid., 78-79, italics in the original.
[180] Velleman: "Love as Moral Emotion", in *Ethics* 109, No. 2 (January 1999), 365.
[181] Ibid., 364-366.

But of course, intrinsic worth in itself is not sufficient; rather, it is only a "minimal condition" that anyone meets. We obviously do not care about and love everyone. The reason for this is that persons have a particular relational worth by virtue of which we come to care about and love them. Wolf claims that we need to have a certain *affinity* toward an individual, which, as she puts it, means "the suitability of an object for our affection".[182] She does not explain in detail what exactly she means by suitability, though she compares it to the suitability of a pair of shoes. Perhaps this comparison means that just as the shape of a shoe and that of the foot has to be similar in order for the shoe to be comfortable, there must be some kind of match between the lover and beloved.[183] For some reasons some persons and objects are more suitable for us to care about and love than others. This is similar to Velleman's view who thinks that that there has to be a contingent fit between the lover and the beloved, and as a result, some people are more loveable for us than others.[184]

I think that the concept of affinity or contingent fit is in vein with Frankfurt's claim according to which love is shaped by our individual character and experience. These can explain why we have affinity for some objects and not for others. Our character and our personal history explain why it is the case that some individuals are more lovable for us than others. However, lovability is not dependent upon the evaluation of an individual against a certain set of standards, moral or non-moral.

But though affinity is not based on such kind of assessment, it is still an objective concept. Whether or not I have an affinity for an object depends on some facts about me and the object, or in other words, some facts about how I am *related* to the object. Thus, inasmuch as affinity is part of worth, it is an objective concept. It refers to some relational properties that contribute to the development of love. At the same time, it does not imply

[182] Wolf, "The True, the Good and the Lovable", 233.
[183] This could remind one to Aristophanes' discussion of love in Plato's *Symposium*. See Plato, *The Symposium* (London: Penguin Books, 1951). Indeed, Solomon calls relational reasons for (romantic) love *aristophanic reasons*. See Solomon, *In Defense of Sentimentality*, 208-213.
[184] Velleman, "Love as a Moral Emotion", 372.

making love conditional upon certain properties. It only explains how love develops. We are simply incapable of loving everyone and the concept of affinity explains why this is so. It does not imply that when we come to love an individual, we do so because of those qualities that explain our affinity towards her and it does not imply that we would cease to love her if she lost those qualities. Love is a process, in which reasons are created, thus, such change in the beloved does not have to lead to such results. Affinity is just something that places certain limitations on us and it might even make sense to overcome them (though we do not necessarily have an obligation to do so).

Thus, relational worth is due to a contingent, objective match between the lover and the beloved. But though relational worth involves value of a certain kind, it does not involve evaluating the individual against any kinds of standards, moral or non-moral. Affinity might even work *against* such evaluations. For example, one might recognize that though someone is evaluated poorly against some certain standard, one is still drawn to him regardless, or even because of this negative evaluation. For example, though one might recognize that a certain person is a lousy chess player, he is awkward or forgetful, this does affect negatively love. A person better evaluated against these standards is not any more worthy of love. Such negative qualities could even have a positive effect on the development of love inasmuch they contribute to affinity.

We have seen that the concept of value in connection with caring and love is complex. On the one hand, we can talk about *worth*, or the value that precedes love, on the other, *derivative value* that is the result of it. I divided these two types of value further and I talked about intrinsic and relational worth in the first group and value constituted by the relationship itself and value brought about by that relationship in the second. That is, we can find two types of value in the second group as well. First, the beloved acquires a special value for the lover; he cherishes her in a way that is the result of love or the relationship itself. Second, the

relationship contributes to the creation of values. Since the first is more important, I will first look at the second of these briefly and discuss the second in detail.

A relationship can bring about certain values and is itself valuable. To quote Wolf, it is important "whether (and how much) the relation between the person and the object has the potential to create or bring forth experiences, acts, or objects of further value".[185] It is fairly uncontroversial that caring and love can bring about desirable consequences and thus it is legitimate to consider them as counting in favor or against a certain individual. It is also quite obvious that being loved can make someone a better person. And maybe most importantly, love contributes to meaning in life and is part of what makes living life worthwhile. Wolf notes that "[I]f there is nothing we love or are able to love, a meaningful life is not open to us".[186] Frankfurt would agree and say that in order to have meaning in life we need final ends[187] and love provides us with them. But how is worth connected to meaning? According to Wolf, "in addition to wanting to live in the real world, we want to be connected to it – that is, we want our lives to have some positive relation to things or people or ideas that are valuable independently of us".[188] This is "the core of the desire to live a meaningful life".[189] We want to live a meaningful life because we want to connect to it, and meaning arises from this connection with the world.

This brings me to the more important value in the second group, value constituted by the relationship itself. The relationship itself is valuable and it constitutes a reason to continue it. As Solomon puts it, sometimes it is argued that "loving is itself a reason for love".[190] Once you love someone, this provides you with a good reason to continue loving her. A similar

[185] Wolf, "The True, the Good and the Loveable", 235. For another critic of Frankfurt and a more psychoanalytic discussion of the claim that love should aim at something good see Lear, Jonathan, "Love's Authority", in *Contours of Agency*, 275-292.
[186] Ibid., 237.
[187] See Frankfurt, "On the Usefulness of Final Ends", in Frankfurt, *Necessity, Volition and Love*, 82-95, reprinted from *Iyyun, the Jerusalem Philosophical Quarterly* 41 (1992), 3-19.
[188] Wolf, "The True, the Good and the Loveable", 236.
[189] Ibid., 236-37.
[190] Solomon, *In Defense of Sentimentality*, 199.

account has been recently defended by Niko Kolodny.[191] He understands love as valuing a relationship. He criticizes Frankfurt's theory of love in the following way:

> In sum, Frankfurt's argument that there can be no reasons for loving a person or thing depends on the implicit assumption that any such reasons would have to be intrinsic properties of that person or thing. Once it is seen that such reasons might be extrinsic properties, such as historical relations of the lover to the lover of a certain type, the argument is no longer so clearly decisive.[192]

In his view value connected to love is constituted by the relationship and not by intrinsic properties of the object. First, he claims that from the first-person perspective of the lover the emotions and motivations constitutive of love seem to be appropriate or make reflexive sense. Second, from the third-person perspective we can judge love or the absence of it inappropriate; for example, we judge negatively a wife's love for her abusive and uncaring husband or a parent's lack of love for his child. And third, love consists of psychological states, and as such, these might be responses to reasons. But are these reasons any more than explanatory? They might explain why it is the case that loving comes about in a certain case, but are these reasons normative, too? According to Kolodny, they are. Though he agrees with Frankfurt that love is non-voluntary insofar as "[O]ne does not decide to love on the basis of considering reasons, and one should not be blamed for loving or failing to love",[193] the reasons listed are normative because they have to do with the *appropriateness* of love. The three types of considerations help us to assess how appropriately one loves or loves not. They do not simply explain how loving comes about, but they also tell how it should come about or be avoided.

But what makes love appropriate and in what way is it a response to reasons? Kolodny's relationship theory is the following: "[M]y reason for loving Jane, I suggest, is my relationship to her: that she is my daughter, or my mother, or my sister, or my friend, or the

[191] Kolodny, Niko, "Love as Valuing a Relationship", in *The Philosophical Review* 112, No. 2, (April 2003).
[192] Kolodny, Niko Review of *The Reasons of Love* by Harry G. Frankfurt (Princeton: Princeton University Press), in *The Journal of Philosophy* 103, No. 1 (January 2006), 50.
[193] Kolodny, "Love as Valuing a Relationship", 138.

woman with whom I have made my life".[194] It is the relationship itself that renders love appropriate. Being related to someone in the ways listed above automatically makes love appropriate. Loving one's child is appropriate from both the first-, and third-person perspective because the psychological states it involves are responses to reasons constituted by the relationship itself. Love, on the other hand, would be inappropriate in cases in which the lover believes that some relationship renders his love appropriate but his belief is false; to take Kolodny's example it would be inappropriate to start to love a stranger's child as one's own whom one sees now and then passing by.

This theory is opposed to both appraisal and bestowal account of love. Since Kolodny's reasons for love are relational and not intrinsic properties of the love object, he refuses the "quality theory" according to which the object's qualities are decisive. As we have seen, an appraisal account has to face several problems. If one loved people or things for their qualities, one would stop loving them as soon as they lost their qualities and since this would be obviously wrong, the quality theorist should maintain that love should not be responsive to its own reasons. Furthermore, if the quality theorist were right I would have reasons to love anyone with the same qualities and as Kolodny notes, "an attitude that would accept just as well any *Doppelganger* or swamp-Jane that happened along would scarcely count as love".[195] On the basis of these arguments of Kolodny, it is not the intrinsic qualities of an object that constitute the reasons for love.

Kolodny's theory has an advantage to both appraisal and bestowal accounts of. If the relationship itself is the reason for love, it cannot be objected to the relationship theory that if we love a person with some certain relational features, we should love any other with the same because it is simply true that one should love all of his children or friends. Consequently this theory, unlike quality theory, does not have to face the fungibility problem.

[194] Ibid., 146.
[195] Ibid., 141.

Also, unlike Frankfurt, it can explain why it is normatively appropriate to love in some cases and not in some others.

But it is not the case that once you love someone, you should continue loving her simply because of the presence of love. As I have argued earlier, there should be a contingent fit between the lover and the beloved and its presence or absence provides reasons independently of whether the relationship exits or not. As Solomon puts it, "love should be accepted and respected *so long as there are not good reasons against it*".[196] Kolodny himself claimed that the wife's love for her abusive husband is inappropriate which suggests that the presence of her love is not enough to justify continuing loving him. Thus, one important question remains: what is the relationship between the different types of reasons I discussed in this chapter? Which one is the most important? If we have a reason to love in one sense, but a reason against it in the other, which one should be more important? In the example, the wife has a reason to continue loving her husband constituted simply by the presence of the relationship, and in the minimal sense of intrinsic worth. On the other hand, she has good reason not to love him in terms of the lack of positive values the relationship brings about, and in the sense of the lack of contingent fit as well. In the last section I will examine the relationship between these different types of reasons. More specifically, I will look at the most important question of this chapter. And this was not the concept of love itself but how it influences our agency and practical reason.

4.5. Love, Reason and Action

To see what the relationship is between the different types of reasons involved in love let us go back to Kolodny's example. As we have seen, though there are reasons for the wife to perpetuate the relationship with her husband, there are also reasons that count against it. The

[196] Solomon *In Defense of Sentimentality*, 200.

question is the following: which of these reasons is more important? Which one of them should override the other? The wife has reasons to continue loving her husband in the minimal sense of intrinsic worth and the relationship itself might be a reason. But here are also considerations that count against it in terms of lack of relational worth and the lack of values brought about by the relationship. As I said above relational worth is about a contingent match or fit between the lover and the beloved and that an individual has a relational worth in case she is related to an agent in a particular way. One could argue that there is some kind of contingent fit in this case but I do not think that the argument goes through. One could say that the husband likes abusing and his wife "likes" being abused, thus, the contingent fit is present. But this surely cannot be right. After all, there is a contingent fit between the sadist and the masochist as well. The sadist likes torturing, and the masochist likes being tortured. But this does not give genuine reasons to act. And it is not true after all that the wife likes being abused; we should rather say that for some subconscious reasons, she needs to be abused. But then, this example belongs to the category of perverse reasons for love.[197] Solomon thinks that some reasons are perverse because "they fail to take love as the love of a person and in a relationship".[198] In the example we cannot really talk about a fit between two persons, only between two perverse desires that complement each other. Thus the relevant contingent fit is missing. What is required is that the individuals involved be related to each other *as persons*, and not just that some of their desires complement each other. In the example, the contingent fit would be present only if the wife's whole personality could be reduced to the desire to be abused, and the husband's to abuse. But surely, we cannot reduce persons to some of their desires.

Thus, it seems obvious that the contingent fit in the sense of two persons as a whole is missing from this relationship. Plus, we can add that there are other types of reasons to

[197] For a detailed discussion see: Solomon, *In Defense of Sentimentality*, especially 213-218.
[198] Ibid., 217.

terminate the relationship, namely reasons based on the values that the relationship brings about. An abusive relationship hardly brings about any values and thus the husband's severely abusive conduct is in favor for terminating the relationship. But why exactly should these considerations override the husband's intrinsic worth and his value constituted by the relationship itself? I myself suggested in the previous example that even though it is permissible to terminate the friendship in case of the drug addict and criminal friend, there are considerations that count in favor of helping him, namely the unconditional nature of love. Should not the wife love her husband in an unconditional way and try to amend their relationship?

One way of answering these questions is considering a further element in love, namely *self-love*, and the relationship between love and self-love. Frankfurt himself think that love and self-love are tied together. We could argue that in the example love and self-love have come apart: the wife's love for her husband reflect some kind of self-hatred and this shows that the relationship is perverse. Though I do not think that it is the best way to solve the problem, I will look at it since Frankfurt discusses it in detail I will argue that though this link between love and self-love is important, this is not the most important reason for or against love. And I will also try to show that due to its subjectivism, Frankfurt's view cannot differentiate between genuine self-love and self-hatred.

Frankfurt discusses the concept of self-love in detail in his book, *The Reasons of Love*. He begins by noting that though self-love is often dismissed as a defect, properly understood, it does not refer to selfishness.[199] The reason for this is that self-love and love for the other is tied together. Loving one's self means that one wants the good for oneself, but what is good for someone is defined by what he loves. As a result, Frankfurt claims that "someone who loves himself displays and demonstrates that love just by loving what he

[199] Frankfurt, *The Reasons of Love*, 70.

loves".[200] Thus, what is good for me and what contributes to my well-being depends on the well-being of the beloved. This explains how it is possible that for the sake of love, one jeopardizes what otherwise one would regard as his well-being. For instance, one can sacrifice oneself for one's beloved without this implying the lack of self-love. Take the following example: someone starts shooting in a supermarket and Mark, out of love for her, jumps in front of Janice to protect her. Though with his action he threatens his life, it does not imply that he does not love himself and does not consider his own well-being. Frankfurt would say that the reason for this is that his well-being is tied to the well-being of his beloved and her death in the shooting accident would obviously influence negatively his well-being. Also, he might add that not saving her out of cowardice would imply that he betrayed himself and did not live up to what the most important was for him.[201]

The same does not work in the case of the abused wife. One might try to argue that she, recognizing that her husband's well-being depends on torturing her, sacrifices her physical well-being out of love for him and let him abuse her. In some less serious cases this might work, as in the case in which the wife lets his husband enjoy the annoying habit of playing the trumpet loudly and falsely now and then. But after a point, self-sacrifice becomes self-torture and the relationship perverse. It would be a pathological form of love, whereas self-sacrifice in the first example is not.

Now, why does not self-sacrifice work in the second case? One might argue that the problem here is that one's well-being cannot be defined solely on the basis of what he loves. For if his love reveals self-destruction, this shows that self-love and love for the other has came apart: "love" in such cases is rooted in self-hatred or some other perverse reasons rather than self-love. Though it is true that the ability to love and to love one's self is closely tied

[200] Ibid, 8.
[201] For a different view on love and sacrifice see Helm, *Love, Frienship and the Self*. There he argues that that love, as a commitment to the import of an individual rationally requires self-sacrifice. See especially 168-170.

together, noted by other authors as well,[202] it is not the true that love, and thus one's welfare, is merely a subjective phenomenon as Frankfurt suggests. Care and love, and self-care and self-love *are* tied to well-being, but well-being is not a subjective phenomenon that depends only on what one cares about and loves. The example above shows that Frankfurt's concept of love fails: we need some objective or rational constrains to be able to define these concepts properly and their relationship to each other. It is useful here to look at Darwall's theory of welfare and rational care. He claims that there is a strong relation between welfare and care:

> What it is for something to be good for someone *just is* for it to be something one should desire for him for his sake, that is, insofar as one cares for him. The relevant sense of 'should' again, is its most general normative sense. We might equivalently say that what it is for something to be good for someone is for it to be something that is rational (makes sense, is warranted or justified) to desire for him insofar as one cares about him. This is a *rational care theory of welfare*. It says that being (part of) someone's welfare is being something that it would be rational to want for him for his own sake.[203]

Thus, insofar as one cares about someone or oneself, one should desire for him not just anything that he happens to desire or want, even in the sense of caring and love. There are rational constraints constituted by the close relationship between welfare and care. It is not the case that care involves promoting anything that the individual cared for desires; for caring is *regulated by* what is good for the person in question, in other words, by his welfare.[204] And as Helm stresses[205], one's welfare is *just partly* constituted by what one wants, cares about or loves. In some cases what an agent cares about does not contribute to his well-being, on the contrary, it affects it negatively.

This constraint on caring places a limitation on love in cases in which love cannot be warranted or justified. This way, we can exclude options that are outright insane or self-destructive. The constraint tells us what cannot be part of caring and love as for example possessiveness, self-destruction and similar disordered forms of love in which the intimate

[202] See for example Erich Fromm's classic account: Fromm, *The Art of Loving*.
[203] Darwall, Stephen, *Welfare and Rational Care* (Princeton and Oxford: Princeton University Press, 2002), 8-9.
[204] Ibid., 9.
[205] Helm, 86.

relationship between love and well-being are compromised. But it does not prescribe what exactly one's welfare should consist in, only what it should not. This would make caring and love paternalistic, which, as Helm puts it, is warranted only incases in which the beloved "displays a diminished capacity for autonomy or a diminished understanding".[206] Consider for example the case of someone who thinks that drinking a bottle of vodka every day is part of his well-being. Caring about a person like him certainly does not amount to buying vodka for him. In this case the drinker obviously has a diminished capacity to act autonomously and to understand what it is that really contributes to his well-being. But paternalistic love is unjustifiable in cases in which the beloved's capacities are not impaired. Thus, though Bach might be more worthy than a rock band in some aesthetic sense, this does not imply that it is better to listen to the former or more importantly, that one should force somebody one loves to go to the Bach concert instead of a rock festival. To a large extent, then, care and love remains subjective, and in a sense Frankfurt is right that it is not better to love something that is worthier by some objective measure. Nevertheless, because of his subjectivism, he cannot differentiate between cases as the abused wife's situation and genuine self-sacrifice.

But the primary reason why self-sacrifice does not work in the second case is not that love and self-love have come apart there. Even though we might agree that love and self-love is tied together, we should take care not to treat self-love as centrally important to why one loves. Self-love cannot be a primary reason for love because it would make love egocentric. If love and self-love are tied together because, as Frankfurt claims, the lover identifies with the beloved, then the close connection between the two is due to incorporating the beloved's well-being into the lover's well being. This makes this conception some kind of *union account*, which means that the identity of the lover and the beloved somehow is merged together. As a result, we can only understand the concern involved in love in

[206] Ibid., 227.

egocentric terms. Helm calls this kind of identification the *egocentric conception of intimate concerns*.[207] In *chapter* 6 of my dissertation I will argue in detail that Frankfurt's concept of love involves an underlying egocentrism. As Helm puts it, this is due to the fact that in this account, "love of another in ultimately grounded in a concern for oneself".[208] I will try to show that this concern for oneself involved here is due to Frankfurt's definition of caring and thus love as a second-order process. If they are primarily involve second-order desires, their primary focus is *not* what the agent cares about or loves. Since they are of the second-order, they are directed at certain desires, and *not* on the object of care or beloved.

Nevertheless, I will develop a revised hierarchical theory of motivation. I will argue that caring and love are complex, multi-order phenomena: they involve a hierarchical structure of desires. However, they primarily belong to the first-order. Caring and love are constituted by first-order desires directed at something and *not* on some other desires. But, due to their typically enduring presence, they are capable of guiding our action. I will also argue that though sometimes this guiding role can manifest itself in second-order desires, volitions, and the like, these are of secondary importance. This shift towards the significance of first-order motivation is indebted to Frankfurt's later work on caring and love. Though he defined caring as a second-order desire, he also stressed that in a basic sense, is finding something important and this implicitly involves first-order desires.

In the last part of the dissertation I will develop in detail my own claims that are based on a revised Franktean hierarchical theory. More precisely, I will defend three separate, but interrelated claims, in three chapters. In *Chapter 5*, I will develop in detail the concept of caring as a "multi-order desire", that is, as a certain kind of first-order desire with an action guiding role and a tendency to generate higher-order desires. In *Chapter 6* I will discuss the significance of returning to the first-order and I will develop the concept of freedom as self-

[207] Helm, 10.
[208] Ibid., 149-150.

transcendence as opposed to self-expression. In *Chapter 7* I will discuss the problem of skepticism about self-control. I will try to defend my view by arguing that though first-order desires are spontaneous and not under our direct, voluntary control, we can nevertheless exercise some amount of control over them by forming certain kind of beliefs about our motivation.

Part III: A revised Hierarchical Theory

Chapter 5: Caring as a Multi-Order Phenomenon

In the first two parts of my dissertation I discussed and criticized in detail two distinct phases of Frankfurt's thinking and the shift in his thinking from emphasizing second-order acts of will to deeper, not voluntarily chosen motivation. At first his hierarchical theory implied that we are agents and persons because we are capable of forming higher-order desires and volitions and perform second-order acts of will, such as endorsement and identification. This view led to a Sartrean view of freedom according to which the agent makes his decisions "arbitrarily", cutting off the process of deliberation by a decisive commitment at a certain point. However, as I have already showed, by introducing the concept of caring, Frankfurt's view on agency, freedom and identity dramatically changed. Autonomous decisions of agents no longer depend on second-order acts of will. On the contrary: caring, and its most important variant, love are both unchosen. They are not the result of acts of will that we perform. We do not control them directly by our decisions and yet it is them that provide us with autonomy. They define who we are and thus when we act in accordance with what we care about and love we act autonomously.

Stefaan E. Cuypers has recently developed a similar interpretation of the changes in Frankfurt's thinking.[209] He differentiates between Frankfurt-1 and Frankfurt-2, the first emphasizing voluntary and the second non-voluntary elements in action. He argues that the two different views on agency lead to two different views on autonomy and he calls "the conception of autonomy largely associated with the Enlightenment tradition 'voluntaristic', and that primarily affiliated with the Romanticism tradition 'non-voluntaristic'".[210] The first one corresponds to Frankfurt's classical view and emphasizes voluntary acts of will as the basis of autonomy and the second corresponds to Frankfurt's view that caring as something that we do not choose by a decision is the basis of our autonomy. Cuypers argues for a hybrid view of autonomy that combines voluntary and non-voluntary elements and also claims that "the voluntaristic conception *asymmetrically depends* upon the non-voluntaristic conception – that is to say, voluntarism depends upon non-voluntarism, but not vice-versa".[211] As I will discuss it more in detail, this means that caring, which is involuntary, is more important and fundamental that our capacity to form second-order desires voluntarily. Our will is constituted by what we care about and thus our second-order desires should depend on what we care about. This way, Cuypers combines the voluntaristic thinking involved in Frankfurt's earlier views and the non-voluntarism of his later theory of caring.

My view is hybrid as well, but in a different way. Instead of using the voluntary/non-voluntary division, I operate with Frankfurt's original distinction between first-, and second order desires; and I will develop a view that defines caring as a *multi-order* desire, in other words, a complex desire that involves a hierarchical structure of desires. Though the two distinctions partly overlap, they are not the same.[212] My view is also different from Cuypers

[209] Cuypers, "Beyond Voluntarism".
[210] Ibid., 226.
[211] Ibid.
[212] One can have second-order desires as a matter of fact, without actively forming them and one can have a degree of (indirect) control over his first-order desires, as I am going to argue in the last chapter. Thus, the voluntary/non-voluntary distinction does not completely overlap with the first order/second order distinction.

inasmuch as his asymmetrical dependency thesis seems to involve skepticism about the possibility of having an influence on our motives when it is needed. In the last chapter I will develop view of self-control and self-management.

5.1. Deep Caring and Reflective Caring

Cuypers claims that since the publication of the paper "The Importance of What We Care About", Frankfurt has "implicitly been working on the […] the non-voluntaristic conception" of autonomy.[213] Cuypers is right to say that there is something implicit in Frankfurt's shift from his earlier views to a quite different one in his later writings inasmuch as the change is not articulated clearly. I will try to grasp this change in terms of first-, and second-order desires and say that while earlier Frankfurt stressed second-order desires in his thinking about agency and autonomy, later on, by emphasizing the importance of what we care about, he partly (and implicitly) shifted the focus to the significance of first-order desires. I will point out that the reason why the shift in Frankfurt's thinking was implicit is the ambiguity of his use of the term 'caring'. Thus, in order to understand the implications of Frankfurt's later thinking to the problem of agency, we have to redefine this concept. In this chapter I will attempt to do this and develop my own concept of caring as a multi-order phenomenon.

We saw in the earlier chapter that the significance of caring is that it provides us with a diachronic volitional unity. It performs this task by including not only a higher-order desire, but also, a desire to preserve it diachronically. So it seems, that in one sense, caring is a special higher-order desire; or in other words it is on the *reflective level*. I will call this phenomenon *reflective caring*. But I would like to point out that caring, besides being a higher-order attitude, is also of the first order. What's more, I will argue that the most

[213] Ibid, 227.

important element of caring is basically a first-order phenomenon, which I will call *deep caring*. Using this distinction I will keep Frankfurt's term *wholehearted caring* and redefine it as a complex desire that *unifies* the different orders.

To see that caring is primarily a first-order phenomenon, we simply have to ask the following question: what does it exactly mean to care about something? As we saw, Frankfurt defines caring as a diachronic commitment to some of our desires. In this sense, caring is a higher-order phenomenon, since it is basically a reflective attitude about some of our own desires. It answers questions as "What is it that is really important for me?" and "What do I really want in life?" So caring is a very important, if not the most important, part of our ability to reflect on our motivation and actively influence our actions.

But in my view, besides this meaning of caring, it also has a more basic, common-sense usage, and what's more, the reflective power of caring is rooted in this. If you ask someone what it means to care about something, he might reply with Frankfurt's basic notion of caring and say: "caring about something means that this something is important to me." But he might not (only) mean by this that he endorses some of his desires on a higher-order in a certain way. Rather, it is more commonsensical to think that when someone cares about, say, his garden he means that he has an attitude, or a desire that is *primarily* directed at the thing in question, which is in this case his garden. Caring about my garden means that I plant there trees, water my flowers, etc; that is, caring is primarily a first-order desire: it is not directed at other desires but at the object of care.

This ambiguity can be found in the simpler, commonsensical definition of caring according to which when we care about something it means that we regard it as important to us. For one can say: "caring about something means that this something is important *to me*" or, putting the emphasis differently, "caring about something means that *this* something is important to me". Putting the emphasis in one or the other part of this sentence reveals which

part of caring one regards as more important. Caring in the most commonsensical meaning of the word is a concern directed at something or someone in the world and thus it simply follows that caring is a first-order phenomenon. The ability to care about something rests on our capacity to have enduring, diachronic desires with an action-guiding role. Thus, in order to be able to care about something we do not necessarily need to form second-order desires and volitions. We can imagine "naïve" agents who have the relevant diachronic desires and wholeheartedly care about something and act in accordance with what they care about without being reflective.[214] However, caring usually does involve second-order processes. Though a naïve agent is capable of caring and yet does not form any second-order desires, we are not typically like that. We reflect about our motivation and while in conflict or deliberating we sometimes try to make use of second-order desires.

To see how are both first-, and second-order desires are involved in caring, let us look at the following example. If we ask someone what it means that he cares about his dog, he will not answer that he wants to sustain some kind of desires, but rather, he will say that caring about his dog means that he feeds his dog, takes him for a walk, plays with him, etc, or in other words, he would naturally think that caring is an attitude that is primarily directed at his *dog* and not at some of his desires. Of course, if he starts to reflect about what he should do in a certain occasion, his deliberation might involve the formation of higher-order desires. Suppose it's cold outside, and he wonders if he should take his dog for a walk or stay at home and watch TV. Now, his decision will reveal what it is that he really cares about. If it happens to be the case that he really cares about his dog and finds watching TV only a pleasant activity that he happens to do now and then, he will probably decide to take his dog for a walk. If, on the other hand, he decides that he will stay at home and watch TV in his

[214] Recently Agniezska Jaworska has been arguing for the same claim in a series of papers. See her "Caring and Internality" and two of her other papers as well: "Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value", in *Philosophy & Public Affairs* 28, No. 2 (Spring 1999), 105-138 and "Caring and Full moral Standing", in *Ethics* 117 (April 2007), 460-497.

comfortable chair in the warm room instead of taking his dog for a walk, this might suggest that he does not care about his dog after all. But of course, one single occasion will not prove that this is really the case. Since we ask here what it is that a person really cares about, and since caring is a diachronic attitude, one's failure to do on a particular occasion something does not entail that one does not care about it after all. Rather, if one *repeatedly* fails to perform the actions relevant to caring about a particular thing, we might suspect that one does not care about it after all. That is, if someone repeatedly fails to take his dog for a walk and stays at home watching TV instead, the poor doggy might be justified in his feeling that his owner does not care about him. Of course, it might be the case that the owner of the dog is simply hindered in taking his dog for a walk because he has broken his leg, or is seriously agoraphobic. But even a dog will be able to understand whether his owner cares about him or not. One's actions simply reveal what one cares about. If the owner is incapable of taking his dog for a walk for some reason, he will still be able to express it in a different way, say, by playing with his dog at home.

True, if at some point, one starts to wonder what it is that one really cares about, it will naturally involve an ascent to the reflective level. But as we have seen, Frankfurt argued in *The Reasons of Love* that in order to answer the normative question about how we should live, we first have to answer the factual question about what we already care about. When we are trying to find out what we care about we find the answer by *discovering* what we really want.[215] Self-discovery is an important process which involves trying to find out what kind of first-order desires we really have.

As we have seen earlier, the basis of Frankfurtian practical reasoning is the following: once one finds out what one cares about (or loves) it gives one a *reason* to act. Of course, this claim is very controversial if we understand "reason" in the sense that it gives us normative

[215] See also Cuypers, 237.

justification to do something. But at least it seems to be clear that one has more reason to do what one cares about than to act on an urge or on a transient desire. Caring, as a desire with a diachronic nature is qualitatively different from a merely momentary, passing desire. Though higher-order desires are superior as well in the sense that they express the reflective authority of the agent, this is not the only way how desires can be superior to other desires. They can be qualitatively different on the first-order as well, and thus ground the authority of higher-order desires. Some of our desires might just come and go and have little significance, and they cannot offer much guidance for our actions. Some others might push us in the wrong direction. Caring, as a set of diachronic first-order desires, has the potential to guide our actions and has a natural tendency to generate second-order desires.

Now, what does the diachronic nature of caring consist in if all this is true? We saw that Frankfurt claimed that caring has a diachronic nature because it is primarily a higher-order desire which is directed at preserving some of our first-order desires. Their diachronic nature, then, would be due to the fact that we want them to be preserved. But this is not the only way how desires can be diachronic, that is, by being kept alive by a higher-order desire. Frankfurt claims that the continuity of a desire has to be the result of our own activity; it should not simply persist without our contribution. This point is crucially important for Frankfurt and is in vein with his emphasis on being active with regard to our motivation. As we saw, he claimed that some of our desires might preserve themselves due to "their inherent momentum" but this is not what he has in mind when he talks about caring, since being active is crucially important for him, and the agent is passive with regard to these desires.

But one can look at this matter in a different way: if some of our desires need to be preserved through our conscious efforts, this fact, at least to some extent, questions their genuineness. If my caring about my dog consists solely in my second-order desire to preserve my relevant first-order desires, do I really care about my dog? Isn't it the case that when I

care about something, this involves desires with an enduring presence; and not just as the result of my own activity but because it is part of their nature that they tend to remain stable as opposed to a passing whim? If the reason why my love for someone is kept alive is that I put constant effort in preserving it, one would naturally think that my love is not real. When we care about something this should not mean that our desires constituting caring are preserved *only* thanks to our efforts to preserve them. In order for caring to be sincere and genuine, the desires that constitute it have to be persistent by nature, or in other words they have to persist "by their own inherent momentum". For after all, what is wrong with having desires that tend to remain an important part of our motivational lives? This is precisely part of what makes them more important: they are not just passing whims and wishes but things that we want, have wanted and supposedly will keep wanting as well.

Frankfurt's claim that caring cannot simply be the matter of some of our desires tending to persist without our contribution might have something to it but we should be careful in avoiding the opposite excess, namely the claim that our desires should be kept alive primarily by our conscious efforts. In order for a desire to be a good basis for caring, it has to have the "inherent momentum", that is it has to have the tendency to remain stable and influence our agency for an extended period of time. Caring is first-order desire like that. As I have argued, it is a more or less coherent set of diachronic, first-order desires. Since it is directed at its object, and not another desires, it primarily belongs to the first order. It is also diachronic, since if someone cares about something, say, his dog, his attitude is not limited to having momentary desires, but it involves desires that one has over a period of time. Thus, caring for example about your dog essentially consists in wanting to do several things, and wanting to do them not only at this or that moment, but tending to want to do them in general. Obviously, what you care about might change through time, but it hardly makes sense to say

that one can care about something different in each moment; we would rather say that such a person does not care about anything at all.

Of course, it is not enough for a desire to be persistent to be a legitimate candidate for caring. Some desires like the addict's craving for the drug is persistent and obviously, this is not to say that the addict cares about getting the drug. Caring consists of desires that besides being persistent also express what the agent finds important or valuable. These features explain why what one cares about overrides what one happens to desire at a certain moment. The owner of the dog might have a momentary or even a persistent desire to watch TV, but when he makes up his mind about what to do, he will find taking his dog for a walk more important, at least in case he really cares about him. His desire to watch TV is not something that expresses his long-term concerns. In contrast, caring has a potential to guide one's actions since it expresses one's long term goals. And because of its guiding role, it can enter into higher-order deliberation by manifesting itself in the form of the various higher-order phenomena, that is, higher-order desire and volition, endorsement or identification.

My claim that caring is a multi-level desire implies that first-, and higher-order processes are not separate phenomena isolated from each other; rather, they form a bigger structure. Caring is a complex psychological phenomenon that includes phenomena of different orders; that is, caring is essentially a *multi-order* phenomenon. On the one hand, Frankfurt is right to say that caring is a manifestation of reflective self-evaluation and as such, it is a higher-order desire or attitude that is directed at first-order desires. I called this *reflective caring*, since it is found on the higher, reflective level. On the other hand, in a basic sense of the word, caring is essentially a first-order phenomenon; it is primarily directed at its *object* and not at any desires. To distinguish this sense of caring from reflective caring, I called it *deep caring*. It is natural that when one cares about something, this *necessarily* involves caring as a special first order desire, and it may also give rise to different higher-

order phenomena, including higher-order desires and volitions, endorsement, identification, and rejection.

5.2 Caring, Valuing and Integration

The view on caring, that is, a type of valuing that is constituted by multi-order desires could be regarded as an answer to Watson's criticism which I discussed in *Chapter 2*. As we have seen, his argument against Frankfurt was that second-order desires in themselves do not have the special status that Frankfurt thinks they do, and our value judgments are primary and our second-order desires are based on them. My view can both answer to this criticism and yet preserve the Humean assumption about agency and the hierarchical theory of motivation as well. Our values are indeed very important and our actions should rest on them. However, I do not believe that our values have to be based on explicit value judgments. As Watson himself admits, earlier he overlooked the difference between valuing something and judging it valuable.[216] Judging something valuable does not entail that one personally values the given thing: as he puts it, "one can in an important sense fail to value what one judges valuable".[217] I can for example judge that a healthy life is valuable but nevertheless continue smoking, because, regardless of my judgment I personally value smoking as such. Thus, just as value judgment does not entail valuing, the latter does not entail the former either. Nevertheless, Watson still wants to reject the Humean view according to which valuing should be reduced to desiring. The reason for this is that though sometimes one might value something that one does not judge the best, the problem is that when one values something this way "one needn't see it as expressing or even conforming to a general standpoint one would be prepared to

[216] Watson, "Free Action and Free Will", 168.
[217] Ibid. Bratman also stresses that judgments of value underdetermine action. See Bratman, *Structures of Agency*. Helm makes a similar point, see Helm, 69.

defend".[218] Watson thinks that such cases can be called "perverse cases". In these cases there is a serious lack of integration between one's value judgments and what one happens to value in a situation. As we have seen, in his view our motives (or motivational system) and values (or valuational system) have to overlap. In case of perverse valuing we cannot talk about value that the agent regards as his own and expressing our general standpoint. In the most ideal case we both value something and judge it valuable.

I think that using the redefined concept of caring, it is possible to work out a Humean concept of valuing which can solve this problem. While I am aware that Watson and many others will still find defining our values in terms of desires unconvincing, I believe that this view can answer to the criticism above and is not a crude reductionism of values. Valuing something as caring about it in the redefined sense of caring I discussed in the previous section does not amount to valuing things in the perverted way mentioned in the previous paragraph. When we value things in the perverted way, we do not care about them in the sense of having stable, action guiding desires complemented with second-order desires. These cases are more appropriately described by Watson's words: "it may not be thought best, but is fun, or thrilling; one loves doing it, and it's too bad it's not also the best thing to do, but one goes for it without compunction".[219] For example, I might light a cigarette against my best judgment because I value smoking in certain situations. I would say that I do not really care about smoking; true, in some situations I do value it and enjoy indulging in it. However, I do not care about smoking; on the contrary, I care about my health. In these cases I do not smoke because I have an irresistible desire that simply overcomes me as an external force; I personally value smoking a cigarette say, after dinner, in good company and with a glass of wine. Nevertheless, it does not express what I care about, that is, it does not express

[218] Ibid.
[219] Ibid.

my stable, action guiding desires. Such values only refer to occasional, passing desires that I like to satisfy in certain situations.

Caring is more that just valuing something even if it has a temporal dimension, as for example a tendency to value smoking a cigarette occasionally. It is not valuing in the "pervert" sense but it expresses one's long term concerns and values. More importantly, caring involves the integration of the different aspects of a person since it involves both stable first-order concerns and second-order desires. Perverse values do not have an action-guiding role, that is, they do not have the tendency to create motivation on the reflective level. Reflective evaluation is also insufficient in itself; it is ineffective in action and does not express what the agent really wants. I argued earlier that there is reason to think that our first-order desires are central part to what we are, so it is inappropriate to say that one's real self is defined by one's second-order desires. This is the reason why reflective self-evaluation is not enough in itself; we need to have the relevant first-order motivation as well.

Caring involves the integration of the different levels, since it includes both first and second-order desires. Caring is a kind of personal valuing of things. It is not a matter of making judgments or forming beliefs about something: it is more personal and expresses not our beliefs or judgments but what we find important and truly want in our life. Caring, being a complex multi-order desire expresses what we really want. These features of caring explain its importance in deliberation: its special action-guiding role.

But is action guiding the goal of caring? Should we care about things in order to guide our actions? Caring itself is very important for Frankfurt; he stresses the "importance of what we care about". He thinks that caring and love are important because they provide us with final ends and guide our actions due to this fact. So it is important for us to care about things because they give us goals and meaning in life. I will argue that in case of genuine love and especially genuine love this cannot be correct. True, caring and love provide us with goals

and meaning in life, but, as I will argue, they are similar to what Elster calls "states that are essentially byproducts of action".[220] In case of caring this means that we do not care about things *in order to* make our lives meaningful. In other words, we do not use caring as a means to an end; caring, and especially love are directed at something or someone and not on achieving something.

To see that caring is primarily a first-order phenomenon, let me look at an important aspect of caring that Frankfurt notoriously neglects but which is often acknowledged by other authors. As we have seen, for Frankfurt, caring and love are primarily a volitional; they do not necessarily involve emotions. Even if they do, it seems that these emotions are not constitutive of caring. However, many philosophers would argue otherwise. It is often stressed for example that caring involves some kind "investment" in a certain object.[221] When one cares about something, one is emotionally invested in it, and is vulnerable to feel pain and joy depending on the fortunes and misfortunes of the cared-for object. Shoemaker thinks that the word caring simply refers to such dispositions: "talk of caring is simply a way of referring to the range of emotional reactions one is expected to have with respect to the fortunes of the cared-for object".[222] He also thinks that "genuine caring about something involves a package deal: one must, along with the possibility of joy (and other positive emotions), accept the possibility of distress (and other negative emotions) when things are not going well with the cared-for object in order for one truly to be said to care for it".[223]

I agree with Shoemaker that caring involves a set of emotional dispositions and that these form a centrally important part of caring. However, this should not mean that caring is reducible to such dispositions. For the most important question is the following: why are we

[220] Elster, Jon, *Sour Grapes: Studies in the Subversion of Rationality* (Cambridge: Cambridge University Press, 1991).
[221] Watson, "Volitional Necessities", Cuypers, "Autonomy beyond Voluntarism" and Shoemaker, "Caring, Identification and Agency".
[222] Shoemaker, "Caring, Identification and Agency", 94.
[223] Ibid., 92.

disposed to have certain emotional responses in connection with the things we care about? Identification is a possible candidate to solve this problem: we are disposed to have these emotional reactions because we identify with the cared-for object. But still, the very essence of caring is still missing from the picture. As I argued earlier, the concept of identification with one's motives is incomplete as well: we need an explanation *why* we identify with some of our motives and not others. Identifying with what one cares about or loves does not provide us with a real explanation. It might be true that we identify with what we care about and it might account for the fact that we are disposed to a range of emotional reactions when we care about something. However, these points seem to be irrelevant when we want to explain what it means to care about something. It would be strange to say that the point of caring is to have certain emotional dispositions. No doubt, having such dispositions makes our life richer. But caring itself is not identical with a set of dispositions. What Solomon says about love applies here as well: "[L]ove involves many dispositions, including a disposition to feel protective or jealous as well as a disposition to experience moments of passionate affection, but it is not itself a disposition".[224] Love involves all kinds of emotional dispositions but it is very strange to define love as a set of dispositions. Rather, Solomon argues, love is "a protracted process".[225] That is, it has a temporal dimension: it is a long-term process that involves (among other things) emotions as well. The same thing is true about caring: it is not to be defined as a set of dispositions. Frankfurt's view has the very same misplaced focus. When we care about something we might want some of our desires to prevail, but again, it would be very strange to say that the point of caring about things is to preserve some of our desires. As I mentioned, it would be unsatisfying to say that the point of caring is to preserve some of our desires even if we believe that we do that in order to have some final ends and make our lives meaningful.

[224] Solomon, "On the Passivity of Passions", In Solomon, *Not Passion's Slave*, (Oxford: Oxford University Press, 2003), 203.
[225] Ibid.

Thus, though certain second-order desires, identification and emotional dispositions might be an important part of caring, these cannot tell us what the real point of caring is. We want to know what it means to care about something and in order to have a full answer to that question it is not enough to talk about the mental phenomena through which caring works. For it is not these that are the real meaning of caring, in the same way as the movements of the heart's muscles are not the real point of having a heart. Rather, we need a purposive explanation: the point of having a heart is that it pumps blood, in order to help us stay alive. Similarly, the point of caring is not the emotional dispositions that we have when we care, or our identifications, even if these make our life meaningful.

The problem with concepts of caring I discussed so far is that they try to give an explanation of this concept by analyzing *only* the psychology of the agent, focusing on what kind of desires, mental states, or dispositions there are in our psychology when we care about something. Why do we have any of these? Humans are not isolated: they are actively engaged with the world, and a very important part of this engagement is that they care about things. And when one cares about something the crucial issue is that gets connected to the world in a certain way. I quoted earlier Wolf who thinks that we need a real connection to the world. Caring is an important way of how we do this. Sustaining some of our desires, identifying with the cared-for object and emotional dispositions are a part of this process but we cannot reduce caring to these. For the point of caring about something is not to preserve some desires, to feel any emotions or to identify with an object, or even, to live a meaningful life but to connect to the world.

It is true that when we care about something we are vulnerable to certain emotional reactions but the story does not end here. Our emotions only reveal that some things or persons matter to us, thus, caring and emotions are connected together. As Cavell puts it, "emotions are orientations to the world that show up how things matter to us, revealing the

world in its relation to us, and ourselves in relation to the world".[226] She even thinks that caring is so essential to emotions that it is "implicit in all emotions".[227] If we did not care about anything, why would we feel any emotions? There is a close relationship between what we care about and what kind of emotions we have, and for this reason Frankfurt's neglecting the discussion of emotions make his view incomplete. The value that things have for us in not a matter of judgments of reason or the result of what we want, on whichever level.

In order to understand caring it is not enough to look at the psychology of a subject. When philosophers try to understand caring, they analyze mental states, such as cognitions, desires or volitions in detail. Then they claim that to care about something is to have certain desires, invest in something, have emotional dispositions, etc. But such explanations will always remain incomplete. As I mentioned above, the question why we care about things always remains open. But whatever the answer is to this question, it is certainly not that we want to have diachronic coherence, emotional dispositions or whichever features in our mental lives. Rather, caring is an engagement with the world (and ourselves) and establishing, as Wolf puts it, a positive relationship with those persons or ideals that we care about. What it means to care about something is to establish, or initiate a connection to the world. Caring about things is the most ultimate mode of our engagement with the world.

[226] Cavell, Marcia, *Becoming a Subject* (Oxford: Oxford University Press, 2007), 126.
[227] Ibid., 136.

Chapter 6: Caring, Love and Self-Transcendence

In the previous section I sketched a revised theory of caring. I claimed that caring is multi-order by nature, that is, it involves a hierarchical structure of desires. However, I also claimed that the most important part of caring is a set of stable, diachronic first-order desires with an action-guiding role. Also, I argued that second-order desires which are part of caring are generated by the first-order level of caring. It is part of the action-guiding role of the first-order desires that constitute caring that they tend to generate second-order desires in deliberation. Thus, when one is faced with a choice between doing something one cares about and doing something one does not care about but feels a strong desire for, a second-order desires will be generated in favor of the thing we care about. For instance, when I have to choose between taking my dog for a walk and watch a TV program, a second-order desire might be generated as a result of the fact that I care about my dog but I find watching TV an unimportant activity. This might especially the case when I, say, out of laziness, feel a strong inclination to stay at home and watch TV instead of taking my dog for a walk. In such cases, there is a need for a second-order desire to promote the option which I find more important.

The claim according to which caring is primarily a first-order desire has an important consequence. It reinterprets the concept of the self embedded in Frankfurt's views and the concept of freedom. In turn, these two are connected, so I will consider them together in this chapter. I showed in the first part that Frankfurt's earlier views are consistent with a Sartrean type of existentialism according to which we do not have an essence and that we are able top make arbitrary free choices. However, as Cuypers noted it, Frankfurt turned from voluntarism to non-voluntarism. As a result, our choices became less important. Instead, he stressed the importance of unchosen motivation: caring and love. These involve volitional necessities, and as we have seen, these are centrally important to who we are. We are defined by what we really care about or find important in life, or to put it in Frankfurt's terms, what we cannot help caring about. I will argue that this view implies a *real-self view* of personal identity according to which we have a *volitional essence* that defines who we are. Frankfurt thinks that one's identity depends on the personal characteristics of one's will, which are "reflexive, or higher-order, volitional features".[228] Also, this view implies a self-expression view of freedom, according to which we are free if our actions are determined by or express our real self.

I will argue against the self-expression view of freedom in this chapter and suggest instead that there is another notion of freedom which is more important theory built on caring and love: freedom as self-transcendence. I have already argued against the real-self view but I will consider some more arguments against it by trying to show that we do not have a volitional essence in the sense Frankfurt thinks we do. This claim is centrally important for the problem of self-control and self-management which I will discuss in detail in the last chapter.

[228] Frankfurt, Harry G., "On the Necessity of Ideals", in *Volition, Necessity and Love*, 113.

But before I begin, I would like to consider a general problem in connection with Frankfurt's treatment of the concept of caring and love. I have already mentioned in the chapter discussing his theory of love that his concept of love is too wide. It includes many concerns that otherwise most people would not consider to be an example of love. Rather, love should be restricted to personal relationships and when we talk about concern directed at other objects than persons as an ideal, a country, etc. it is better to talk about caring. Frankfurt talks about love as the most important mode of caring but he never specifies what other modes of caring there are. It is better to restrict love to personal relationships and talk about caring in other cases. Thus, since Frankfurt's concept of love is too wide, it could be argued that a lot of claims in *The Reasons of Love* can be applied to caring in general. For example, Frankfurt argues that love is liberating but if this is true, wholehearted caring in general is liberating as well for the same reasons.

6.1. Two Views on the Self and Freedom

As I mentioned above, there are two different concepts of the self and freedom. On the one hand there are 1) those who think that freedom amounts to be determined by, and *express one's essence or real self*; on the other hand, 2) there are those who think that *transcending ourselves* in the sense of being determined by the objectively good is what makes us free. We can call the first view the *self-expression view,* and the second one the *self-transcendence view.* In my view Watson and Frankfurt belong to the first category, Susan Wolf, Berofsky[229] and others who stress objectivity to the second.

Let us look at the self-expression view first. As I mentioned, according to this view being determined by, and expressing one's real self amounts to freedom. There is something

[229] Berofsky, Bernard: *Liberation from Self: A Theory of Personal Autonomy* (Cambridge: Cambridge University Press, 1999).

attractive or intuitive about this view. For what else could be freedom if not such a state of affairs in which I am determined by nothing but myself? This view is built on the assumption according to which if someone is determined by something *external* to him in his actions, he loses his freedom. Instead, if an agent is determined only by what is *internal* to him, he is free. There might be several versions of this theory, depending on what one thinks about human nature. If one thinks that humans are basically rational and that the essence of a person is his evaluative faculty, one might say with Gary Watson that freedom rests on an ability to act in accordance with one's judgments that are based on one's most deeply held values. One must be able to make evaluative judgments and make these effective in action. On the other hand, there are philosophers as Frankfurt who build on a contrary theory of human nature. Frankfurt claims that humans are basically moved by desires. True, human psychology is quite complicated and less than transparent, but Frankfurt's theory can deal with this difficulty. Desires form a hierarchical structure synchronically and these in turn form diachronic unities, which are manifested by what we care about and love in life. Thus, if one's actions are determined by and express what one loves most and cares about most in life, one is free. Both theories claim that if one's actions are determined by one's real self, one is free. In both views, our real self is defined by our most deeply held values. Thus, they agree about this main point, they only disagree about the nature of this real self, whether it has a rational on non-rational nature; whether our values are expressed by our judgments or desires.

I have already argued against the real-self view while discussing Frankfurt's concept of identification in *chapter 2*. There I argued that it is not the case that identification makes a desire and internal and rejection makes it external. In this chapter I will consider different arguments against the real-self view. Several philosophers objected to the this view, especially Susan Wolf, who claimed that one's real self might be "insane" as a result of

negative developmental circumstances, or coercion, brainwashing, addiction and the like, and that insane agents are not free.[230] Rather, she argues for the Reason View[231] according to which in order to be free, we have to be able to act in accordance with the objectively true and good, which implies self-transcendence. If we are unable to do this we are not free, even if our choices express what we truly are. Recently Gideon Yaffe also argued against the self-expression view and claims that Locke has a self-transcendence theory as well, according to which the self-transcendent agent has a superior freedom to the freedom of the merely self-expressive agent.[232] He argues that "the full fledged free agent has both freedom of action – she expresses herself in her conduct – and freedom of will – she transcends and escapes herself, and thereby becomes like God".[233]

In this chapter I will focus on Frankfurt's views to show how exactly self-expression and self-transcendence are connected. I will especially focus on his theory of *caring* and *love*. Caring and love have a power to determine one's actions, and consequently they can provide one with freedom as self-expression. Frankfurt's views imply that love makes us able to express ourselves and be free in that sense. If one loves wholeheartedly, it means that one has a real self and is able to express it in action. However, I will argue that the essence of love is self-transcendence and not self-expression, and as a result, freedom as self-transcendence is a more important effect of love than freedom as self-expression. Frankfurt himself claimed that self-transcendence *is* involved in his theory: love as a volitional configuration necessarily involves an obligation to transcend one's self.[234] However, it seems that he focuses more on self-expression than on self-transcendence, and his subjectivism seems to preclude real self-

[230] See: Susan Wolf: "Sanity and the Metaphysics of Responsibility" in Schoeman, Ferdinand, *Responsibility, Character and the Emotions* (Cambridge: Cambridge University Press, 1987), 46-62.
[231] Wolf, *Freedom within Reason.*
[232] Yaffe, Gideon, *Liberty worth the Name. Locke on Free Agency* (Princeton: Princeton University Press, 2000).
[233] Ibid., 74.
[234] Frankfurt, personal communication.

transcendence. To use Helm's term, the real-self view implies an "egoistic concept of intimate concerns" and an "individualist conception of autonomy".[235]

6.2. Freedom as Self-Expression

In *The Reasons of Love* Frankfurt claimed that love liberates us. When we are able to love wholeheartedly and without any self-doubt we go through an experience of liberation, enhancement and expansion of ourselves. Love imposes volitional necessities on us, and these "are themselves liberating"[236] because they are capable of "freeing us from ourselves".[237] By this crucial phrase he means that love frees us from the painful states of *ambivalence* and *indifference*. When we are ambivalent or indifferent we do not have a volitional essence, and as a consequence we are impotent agents. Love provides us with volitional necessities and consequently it makes us potent agents. It provides us with a real self that we can express in our actions. Therefore, it seems that Frankfurt's claim about the liberating effect of love implies that love makes us free because it makes self-expression possible.

Now, there is determination involved in love in two senses. First, "it is a necessary feature of love that it is not under our immediate and voluntary control".[238] Love is directed at an irreplaceable object, but in such a way that we cannot voluntarily control which object it will be. We simply end up loving some objects and not loving others, without much choice in this matter. Second, once we are lovers, we have no choice but to promote the well-being of the beloved. This is a volitional necessity that we cannot escape. We necessarily want to benefit the beloved, without being able to want to harm her. What we truly care about and

[235] Helm, 9.
[236] Frankfurt, *The Reasons of Love*, 65.
[237] Ibid, 66.
[238] Frankfurt, 44.

love seems to be part of, if not the essence of our practical identity, and thus Frankfurt's view involves a motivational or volitional essence.

Thus it seems that we have a motivational or volitional essence that defines who we are, or in other words a real self. Before going on, I would like to consider some problems in connection with this problem. First, Velleman criticized Frankfurt by arguing that we do not have a real self, at least not in the sense of a "singular entity waiting to be found".[239] In his reply to Velleman, Frankfurt declares that he does not talk about a self in the sense of a singular entity. Rather, volitional necessity draw only the boundaries of the self, but Frankfurt stresses that his view does not have any ontological implications and that he is "not inclined to construe the self as an "entity" at all".[240] This might suggest that Frankfurt's theory in not a real-self view after all, since we do not have a real self or essence in the sense of an "entity". However, as we have seen, he does assert that our continued existence depends on the continuity (or diachronicity) of our volitions. Thus, though he thinks that there is no such thing as "a single entity to be found", he does think that our volitions define who we are.

Another point about volitional necessities that Frankfurt sometimes talks about them as if they were merely negative, that is, they make it impossible for us to act in a certain way: "[V]olitional necessity constrains the person himself, by limiting the choices that he can make".[241] This suggests that we do not have a volitional essence after all; there are only some possibilities that we cannot bring ourselves to want. But it is easy to show that Frankfurt's view does involve a volitional essence. To use his example of parental love, consider a mother who cannot bring herself to intentionally harm her child. But it is not the case that her love is simply defined by what she cannot do. This negative volitional necessity is only derivative of the positive one that is the essence of her love. The mother cannot bring herself to intentionally harm her child simply because she loves her, and she wants what is good for

[239] Velleman, "Identification and Identity". 99.
[240] Frankfurt, "Replies", 124.
[241] Frankfurt, "On the Necessity of Ideals", in Frankfurt, *Necessity, Volition and Love*, 138.

her; "she is guided by the interests of the beloved",[242] and as a consequence she cannot want the bad for her. Thus, after all, volitional necessities do not simply draw the boundaries of the self by making it impossible for the agent to want some possible courses of action. They are dependent on, and the consequence of, the positive determination involved in caring and love, in this case the interests of the child.

But if love has such nature, how can it make us free? Some might have he intuition that it takes away our freedom by making us subject to the two kinds of determinations above. But Frankfurt thinks that the alternative to the necessity by which love moves us is either *indifference* or a state of *ambivalence* and it is these states that take away out freedom. Both of these, to a different degree, are characterized by uncertainty, self-doubt and hesitancy. Love is able to liberate us from these negative states precisely because of the volitional necessities it creates:

> […] the necessity with which love binds our wills puts an end to the indecisiveness concerning what to care about. In being captivated by our beloved, we are liberated from the impediments of choice and action that consist either in having no final ends or in being drawn inconclusively both in one direction and in another. Indifference and unsettled ambivalence, which may radically impair our capacity to choose and to act, are thereby overcome. The fact that we cannot help loving, and that therefore we cannot help being guided by the interests of what we love, helps to ensure that we neither flounder aimlessly nor hold ourselves back from definitive adherence to a meaningful practical course.[243]

In this text Frankfurt compares three types of experiences: being indifferent, being ambivalent and to love or care wholeheartedly. All three terms have a double reference, that is, indifference means indifference toward both one's self and the world, ambivalence means ambivalence both toward one's self and some objects in the world, and as we have seen, caring about something involves caring about one's self, loving individuals wholeheartedly

[242] Frankfurt, *The Reasons of Love*, 66.
[243] Ibid., 65-66. Frankfurt also discussed the liberating effect of volitional necessity in an earlier paper, "The Importance of What We Care About, in *The Importance of What We Care About*, 80-94, reprinted from *Synthese* 53, No. 2 (1982), 257-72.

implies loving one's self. In order to see how love and freedom are connected in Frankfurt's thinking we have to compare these states.

If somebody is indifferent, he has no goals in life and "flounders aimlessly". What can be the cause of this state? Many know the experience of emptiness characterized by not having goals in life and not being able to see any meaning in life. Such indifference might be a severe form of depression in which one's capacity to make choices and act is severely compromised: actually, in such a state one might be completely impotent as an agent. To see this, it is important to stress that here indifference is understood in the strictest possible sense: not caring about anything. After all, as I discussed it chapter 2 in connection with Strawson's argument against a narrative view of the self, there might be people who claim that they do not care about anything but their attitude to life reveal that they somehow care about not caring about anything.

What Frankfurt means by not caring about anything is quite a painful state in which one is completely lacking of any motivation and hence unable to make any decisions. And this is the key point: the *lack of capacity to make decisions* is what characterizes the state of indifference. This lack of ability to make decisions manifests itself as a painful experience of boredom, emptiness and meaninglessness. This is not quite the state Strawson has in mind. It is important to stress that Frankfurt's view, though based on caring and love, does not imply a narrow view on what is the best way to live our life. As we have seen, this view is neutral as regards morality, and it is neutral with regard to what vision of the good life is correct, at least in the narrow sense. It only involves that there has to be something that one wholeheartedly cares about, but as I argued in *Chapter 2*, Strawson's vision of the episodic life can fit well into this picture. As I tried to show episodic agents do care about something: it is important for them to always live in the present.

The other alternative to Frankfurtian love is ambivalence which is characterized by being drawn into incompatible directions in life. In this painful state one is unable to settle about basic issues in his life, together with a concern about solving them. The ambivalent person is not at all indifferent, he finds it important to solve his conflicts, but he lacks the ability to do so. The reason why such a person is an impotent agent is the same as in case of the indifferent: he *lacks the capacity to make decisions*; in this case he is unable to solve a serious internal conflict. And again, this impotence manifests itself as a painful experience of being torn between alternatives. The ambivalent person lacks *wholeheartedness*; his heart is divided between incompatible alternatives.

Now, Frankfurt claims that the reason why love makes us free is that it makes us able to either avoid or overcome the states of indifference and ambivalence. It makes us potent agents who can make decisions:

> It may seem, then, that the way in which the necessities of reason and of love liberate us is by freeing us from ourselves. That *is*, in a sense what they do. The idea is nothing new. The possibility that a person may be liberated through submitting to constraints that are beyond his immediate voluntary control is among the most ancient and persistent themes of our religious and moral traditions. "In his will," Dante wrote, "is our peace". The restfulness that Russell reports having found in the discovery of what reason required of him evidently corresponds, at least up to a point, to the escape from inner disturbance that others profess having discovered through accepting as their own the inexorable will of God.)[244]

According to Frankfurt, love and reason are very similar because they both impose necessities upon us, but they do so without impairing us. When we submit to the necessities of reason and love – when we follow reason and our hearts – we experience ourselves as free and potent agents. Once I know what the truth is I cannot want to believe otherwise, and once I love wholeheartedly I cannot want to do otherwise than love commands me.

To sum up, though Frankfurt's theory on love and its liberating effect implies self-transcendence, there seems to be too much focus on self-expression. According to him, love's

[244] Ibid., 66-67 (italics in the original).

liberating effect is due to its power to liberate us from some painful internal states, and to create a harmonious real self in us. I will argue that this liberating effect relies on the fact that love necessarily involves self-transcendence. But if that is true, it provides us with a different and more important variety of freedom: self-transcendence makes it possible for us to escape the prison of our solipsism, which is more than merely transcending some mental states.

6.3. Freedom as Self-Transcendence

I quite agree that wholehearted caring and love helps us to overcome the painful internal states of ambivalence and indifference and as a result, it makes it possible for us to express ourselves in action. It liberates us from ourselves in the sense that it liberates us from indifference and ambivalence. And it is also true then, that love provides us with freedom understood as self-expression. But I would like to emphasize that the most important variety of freedom that love provides us with is freedom as self-transcendence, and the liberation that Frankfurt talks about relies on it. Love frees us from ourselves in a more important sense: it frees us from the prison of our solipsism. My main argument is connected to the nature of love itself. The primary goal of love is to transcend ourselves; overcoming some internal states and self-expression are only its effects. Or, to take Elster's term, they are *by-products* of love.[245] These are such states that come about as a result of love, but we cannot intentionally choose to love in order to achieve some certain internal state. This would mean that we make love instrumental to this goal. However, love should not be an instrument of self-regulation. If this would be the case, it would not be genuine love. Genuine love implies that the primary focus of our actions is no longer our own selves but some other. Its goal lies outside of the lover, namely in the good of the beloved. Since, as Frankfurt himself claims,

[245] Elster: *Sour Grapes*. I will discuss this concept in detail in the last chapter.

love should be directed at self-transcendence and not self-expression, freedom as self-transcendence is a more important aspect of love than freedom as self-expression.

Frankfurt himself claims that the essence of love is a volitional configuration directed at the good of the beloved. Consequently, we should not focus on how love liberates us from some painful mental states. If I love wholeheartedly, it no longer matters for me what my mental states are. On the contrary: genuine, selfless love means that one does not focus on oneself any more, but on something external, and its good. If I want to love just because I want to overcome a messy, painful state of internal ambivalence, it is not love at all. If I want to love because my life is empty and meaningless and I want to give meaning to it or to fill in a hole in myself, it is not love at all. Attempts to love in such ways fail because of their misplaced inward focus. The beloved would recognize that after all, it is not her, who is important for her lover, but her role in his psychology. His actions might be directed at her well-being but not for its own sake, but because it is instrumental to creating harmony and meaning in life.

I do not claim that it is not desirable to reach a harmonious motivational structure through love. But I claim that creating such structure should not be our goal in loving. Rather, our goal should be the well-being of the beloved for its own sake, and this necessarily involves that we abandon all self-regarding goals. Love should work this way if it is to be genuine. And if we successfully love this way we will able to reach that internal harmony anyway. What is more, this harmony cannot be reached any other way. To use again Elster's notion, some states are essentially by-products of action. One cannot fall asleep by intentionally trying to do so, because such attempts are self-defeating.[246] The more I try to fall asleep the more I reduce my chances of falling asleep. This is because intentionally trying to fall asleep keeps my consciousness working and this way it prevents me from reaching the

[246] Elster, 45-46.

desired state, since sleeping implies leaving the state of consciousness. Something very similar is at work in case of love. If I find that I am like the Frankfurtian non-lover either in the state of emptiness or boredom or in the state of a messy ambivalence, I might become desperate about this fact. I might realize that the reason why I am in this painful state is that I do not care about or love anything in life. I might want to change this and adopt the same strategy as some people might try to employ in case of insomnia. That is, I might try to start to care about something or to love someone. However, the more I try, the more I will fail as in case of falling asleep. The bigger my concern for changing my internal states, the less chance I have to achieve genuine caring or love. Genuine love is selfless and consequently it implies that one is not aware of oneself any more, or at least that it is not in the center of his attention. A central concern for changing oneself is the enemy of love in the same way as concentration is the enemy of sleep. Not only it is an obstacle to it, it is just the opposite of the state that we are supposed to reach. I cannot lose my consciousness and fall asleep by keeping my consciousness working. And I cannot selflessly direct my efforts on the well-being of someone else by selfishly focusing on my internal states.

This comparison of love with sleep reveals that they have an important feature in common. In both cases, there is a *transition* from one state to another. In case of sleep, I begin with a conscious state, and I lose my awareness gradually. I cannot observe this transition after a while because that would bring me back to consciousness. Insomniacs know that once you give up trying to fall asleep, you succeed. And this is because in such state of resignation one's consciousness does not interfere with falling asleep any more, and the transition can happen. The transition is preceded by a conscious state, and it begins only when consciousness starts to fade. In the same way, there is a transition in case of caring and love as well. I begin with a state of either indifference or ambivalence, and I end up as someone who is capable of selfless love. Just as in case of falling asleep, I cannot move from

one state to the other by simply intentionally trying to do so. Trying to do that will perpetuate my self-regarding goals and thus prevent love from coming about. Thus, just as the transition in case of falling asleep is preceded by consciousness and begins with the fading of it, the transition from solipsism to love might be preceded by an incapacity to love characterized by a focus on our own selves and begin with the fading of this focus. What I have to reach is a state of self-forgetfulness that makes it possible for me to put the object of love into the center of my attention instead of my own internal states.

Of course, some people fall asleep easily, and just as well, some people can achieve genuine love easily. The transition in such cases occurs more smoothly and quite naturally. Nevertheless, the difficult cases reveal that it is necessary to eliminate something to go through a transition from one state to the other. And this thing is conscious effort in case of falling asleep, and it is the focus on one's self in case of love. This excess of focus on one's self has exactly the same role in preventing love as conscious effort has in preventing sleep. We can also have another word for this state. As in case of a lack of ability to fall asleep we can talk about *insomnia*, the state that prevents love can be called *solipsism*. It involves solipsism in the moral sense that one's primary concern is oneself, and in the more actual sense that one gets isolated from the world around him. The most important aspect of love is that it makes us able to step out of solipsism understood in this sense. We are no longer prisoners of ourselves, and not being a prisoner is obviously a certain type of freedom.

According to Murdoch[247], modern Anglo-Saxon philosophy tends to treat the individual as a solitary will. She thinks that philosophers since Hume and Kant isolate the individual from his environment and treat him as a lonely self-contained individual. Arguing against this tradition she claims that "[W]e are not isolated free choosers, monarchs of all we

[247] See Murdoch, Iris, "Against Dryness", in Murdoch: *Existentialists and Mystics* (London: Penguin books, 1997), 287-295.

survey, but benighted creatures sunk in reality".[248] Defining love, she claims that it is "the extremely difficult realization that something other than oneself is real".[249] The virtue of Frankfurt's thinking is that it also reveals that one's will is not isolated but it is bound to the environment, and in love it is bound to the reality of some other. However, as Helm notes his concept of love is still grounded in an underlying egocentrism. His views seem to apply that "love of another is ultimately grounded in a concern for oneself".[250] To use Gary Foster's words, he "characterizes love from the perspective of the lover (or sometimes the beloved), whereas love between mature individuals can only be properly understood and characterized in dynamic relational terms".[251] In this chapter I argued that because of the relational nature of love, we must stress that love implies self-transcendence, and the positive effects of love that Frankfurt talks about are only secondary and are based on this self-transcendence.

Now, what is the upshot of this discussion for Frankfurt's concept of second-order desires and reflective self-evaluation? Forming desires of the higher order and evaluating ourselves reflectively by definition involves focusing our attention on our own first-order motivation. As a result, Frankfurtian reflectivity can be an obstacle to reaching genuine caring and love. These involve that we turn our attention just into the opposite direction: on an individual in the world. An excessive exercise of our reflective powers interferes with the development of a genuine concern with something outside ourselves just in the same way as a conscious effort to fall asleep interferes with falling asleep. And since what we care about is central to our agency, an excessive self-focus interferes with agency as such. In order to be agent in the full sense we have to be able to "put to sleep" our concern with ourselves and we have to be able to focus our attention on the right thing. And this is not focusing on our on desires but rather, on what action to perform, which is in turn depends on what we care about.

[248] Ibid., 293.
[249] Murdoch, Iris, "The Sublime and the Good", in Murdoch, *Existentialists and Mystics*, 215.
[250] Helm, 149-150.
[251] Foster, Gary, "Bestowal without Appraisal: Problems in Frankfurt's Characterization of Love and Personal Identity", *Ethical Theory and Moral Practice* 12 (2009), 167.

The latter involves focusing our attention on what we care about. I believe that we can reach the freedom that love can provide us with only this way.

Real love involves real self-transcendence, and it liberates us from our solipsism. But what if a solipsist claims to have a different kind of freedom in the sense that he is free from the external world? Whether such freedom is more valuable than freedom from our selfish desires, the reader should decide on his own. Or perhaps those can decide the best who have experienced both.[252]

[252] I owe this point to Ksenija Puskaric.

Chapter 7: Caring and Self-Control

7. 1. The Problem of Self-Control or Can You Make Yourself Care About It?

In the last chapter I will like to focus on what I think is the biggest problem with Frankfurt's views: there is a theory of human nature embedded in his recent views on caring that seems to be a bit too dark and pessimistic. Frankfurt does have an ideal of human life, namely the life of a wholehearted agent whose actions are determined by what he cares most about in life. However, there are hardly any Frankfurtian wholehearted agents, and it might be natural to ask: is it in our power to attain this ideal? Human life is often troubled by ambivalence and indifference, and some people might simply unable to overcome these states. And even if some people manage to reach this ideal, they might end up wholeheartedly caring about things that seem to be obviously wrong: the destruction of other people and themselves. In both cases, it makes sense to say that these people should change themselves. But are they capable of it? If caring is really centrally important, the question of self-change and self-management has to be put this way: can we make ourselves care about something? Or is this

question meaningless? Is it the case that people simply care about what they happen to care about and thus they are incapable of changing themselves: why would they want to change if they do not care about it? Thus, it seems that caring – something which we cannot simply directly control – is the ultimate point from where everything begins in our motivational life.

But I would like to argue that this theory on human nature and human psychology embedded in Frankfurt's views is too pessimistic. We can make sense of cases in which people recognize that they do not care about the right thing and yet they are able to make a change. That is, a certain type of self-management is possible. Suppose for example, that an addict realizes that he should change himself. However, he reached a point in his addiction at which he no longer genuinely cares about changing himself. If Frankfurt is right, his lack of caring implies that he cannot change the situation. The reason why we cannot be able to change ourselves if Frankfurt is right is that in his view, we will not do so unless we care about changing ourselves. Now, if changing ourselves depends on caring, in cases in which caring is missing, it further depends on our ability to make ourselves care. And this is the main question: can we make ourselves care about something? Frankfurt's answer to this question would be this: you can make yourself care about something, if you care about making yourself caring about it. But this regress cannot go on infinitely. That is, there is a point where the regress has to stop, and we are left only with an insight about what we happen to care or not care about on the deepest level, and the recognition that it is simply a fact that is given, something that itself is not voluntarily chosen, and which we cannot simply change at will.

As I mentioned, Cuypers argues that the voluntaristic conception of autonomy (the theory of autonomy stressing reflectivity and the active formation of second-order desires) asymmetrically depends on the non-voluntaristic conception (the theory of autonomy which claims that we are autonomous when we act in accordance with what we care about), in other

words, “voluntarism depends upon non-voluntarism, but not vice versa”.[253] He defends a very similar interpretation of Frankfurt’s views that I have been developing so far. He puts it in the following way: “the nature of a person's will is fundamentally constituted by his caring about, loving or being committed to certain things”.[254] Accordingly, one’s ability to form second-order desires depends on what onee cares about. Thus, to use Cuyper’s terminology, the voluntaristic conception of autonomy asymmetrically depends on the non-voluntaristic conception. He defines this thesis more in detail in the following way: “[G]enuine and effective identification (with desires) through forming decisive second-order volitions asymmetrically depends upon identification (with desires) through caring about something as restricted by volitional necessity”.[255] In other words, what kind of second-order desires and volitions we form depends on what we care about. It is not the case that we can identify with any of our desires or that we can form a second-order volition supporting any of our desires; rather, our reflective, second-order agency has to be informed and restricted by what we care about. Thus, Cuypers argues that “the non-voluntaristic conception is more fundamental than the voluntaristic one” and consequently “caring about something as restricted by volitional necessity constitutes the ultimate foundation upon which the organization of a person's will and the execution of his acts of will rest”.[256]

I agree with Cuypers that caring is more fundamental than forming second-order desires but my views are also significantly different inasmuch as while he uses primarily the voluntary/non-voluntary distinction, I work with the first-order/second-order distinction instead. In disagreement with Frankfurt, I argued that caring is more fundamentally a first-order desire than a second-order one. Thus, in my revised theory of caring and agency our first-order desires are more fundamental that what second-order desires we form. A more

[253] Cuypers, 226.
[254] Ibid, 240.
[255] Cuypers, 248.
[256] Ibid., 249.

important difference is the following: though I agree that caring is non-voluntary inasmuch we do not have a direct, voluntary control over what we care about, I strongly disagree with Cuypers' claim according to which "personal autonomy and self-identity require *pre-fixed boundaries and restrictions*".[257] If this means that in order to be autonomous, we have to care about something I agree, but if it means that what we care about is fixed completely independently of our activities I disagree. I will argue that for effective self-management, sometimes it is necessary to believe that we do not have a volitional identity; it is necessary to think that it is not the case that we have a pre-fixed identity. Though I agree with Cuypers that caring is non-voluntary, we do have a say in what we care about. If we believe that though our actions depend on what we care about what we care about is not completely fixed and determined beforehand, we open up a world of possibilities. Humans are capable of openness and as a result, we can, to a degree, control what we care about.

Now, how exactly can the revised hierarchical theory I sketched handle the problem of self-management? In the previous two chapters I argued that caring is essentially a multi-order phenomenon that has both a first-order and a higher-order element. But I suggested that the first-order element is more important. I have tried to show, what the real purpose of caring (and love) is. One answer to this question is Frankfurtian, according to which the goal of caring and love is that by caring, we organize our volitional essence through time; we create a diachronic volitional unity, or to use Frankfurt's term, we *constitute* ourselves. In this sense of the word, caring is a reflective, higher-order process, which has a specific function. But as I have argued, caring essentially involves first-order desires, since it is directed at a certain object and not on other desires. In this sense of the word, caring belongs to the first-order and I called this deep caring. I have tried to show in the previous chapter that if one cares about something or loves someone *in order to* perform the function specified above,

[257] Ibid., italics in the original.

one's care or love ceases to be genuine. This is more obvious in case of love: if a person loves a person only because he would be miserable (or indifferent or ambivalent, to use Frankfurt's terms) without her, his love is certainly not genuine. And if I care about something, say, my garden or my dog only because my life would be empty and meaningless without it, this does not seem to be genuine caring.

Thus, so far I have been trying to restore the significance of the first-order, as opposed to Frankfurt's view that stresses the importance of the reflective, higher-order level. But this view leads to a disturbing problem: if our first-order desires are so crucial and if the most important type of first-order desires – first-order, or deep caring – is not under our voluntary control, it seems that self-control has became very difficult, if not impossible. But in this chapter I will try to show that a certain type of self-control and self-management is consistent with the revised hierarchical model of agency.

I argued earlier that first-order or deep caring is more important and this is what gives rise to second-order phenomena, that is, reflective caring. Because first-order desires are more important, second-order desires have only a limited role. Take the case of someone who wholeheartedly cares about something or wholeheartedly loves someone. For wholeheartedness to be present, the relevant first-order desires themselves have to be in place. That is, there have to be stable, enduring desires directed at the object of caring or the beloved. These desires have to be part of one's identity and have to be able to guide one's actions. When this level of caring or love is present, second-order desires have the following role: they add a further affirmation to these first order desires and thus contribute to their enduring presence and efficiency.

When one is ambivalent, second-order desires might be helpful as well. In an ambivalent case, one has two desires opposed to each other. And when he tries to solve his conflict by forming a second-order desire, this second-order desire will be directed at one of

the desires involved in the conflict. Thus, second-order desires can have a special role in handling an internal conflict: the agent can try to form a conflict by committing himself to one of his desires and overcome his inner state of opposition.

The cases of the wholehearted and ambivalent agent demonstrate that second-order desires can have a functional role in exercising control over one's motivation. But what is the case when we have second-order desires but the relevant first-order desires are not present? The case of the unwilling addict is a classic example of this case: he has a second-order desire directed against his desire to take the drug but a genuine first-order desire for abstinence seems to be missing. To take another example, what if someone says, "I wish I cared more about myself" or "I wish I wanted to be kind more often", etc. but without really caring about himself or wanting to do kind things? In these cases we can clearly talk about the presence of second-order desires without genuine first-order motivation. Second-order desires are by definition desires to have or not to have desires and this is exactly the case here: the agent wishes that he had certain first-order desires that constitute caring about himself. In such cases the agent has a second-order desire without having the relevant first-order desire in the robust sense.

It seems that in cases as the one described above second-order desires are the symptom of a serious problem with our motivation. Though in some cases they express an affirmation of what we want, as in the case of the wholehearted person, in cases in which self-management is needed, they are rather just a symptom of a *volitional disorder*. When one suffers from a volitional disorder, one wants to do something one wishes one did not want to do and one does not want to do something else one wishes one wanted to do; typically both happen, complementing each other. To go back again to a typical example of a volitional disorder, the unwilling addict wants to take the drug badly (which he wishes he didn't want to do so badly) and he does not want to try to quit enough (which he wishes he wanted to do

more). In cases like this the presence of second-order desires is the painful symptom of an acute problem: a volitional disorder.

Thus sometimes, rather than being a cure and a solution to a problem, second-order desires can be a symptom of a serious volitional disorder. We cannot use them strategically; on the contrary, they can interfere with the cure itself. Their presence just contributes to, or even constitutes the problem itself. In the next section I will try to explain why second order desires are inefficient in solving a volitional disorder by comparing it to what Jon Elster calls states that are essentially by-products of action.

7. 2. First-Order Caring as a State that is Essentially a By-Product

Sometimes one's wishes can become self-defeating. Sometimes the more we want something to be the case, the more we interfere with its occurrence. The same can happen with second-order desires, especially in case of a volitional disorder. Sometimes the more we wish to want something genuinely or not want it that badly, the less chance we have that our wish will come true. Having or not having first-order desires can be something that is not dependent upon our second-order desires but rather, are similar to the phenomena that Elster called by-products of action. As I mentioned in the previous chapter, he uses this term to describe some desirable states that we cannot reach intentionally; these states can only come about as the by-products of other actions.[258] It is not possible to achieve these states intentionally, by deliberately trying to bring them about. I argued that in case of love this means that we cannot reach the desirable states that Frankfurt talks about in connection with love – overcoming ambivalence and indifference – if we love *in order to* reach these states. Genuine love implies that we do not focus on our own interests but on that of the beloved; and it is precisely this feature of love that will bring about the desirable consequences of love. The

[258] Elster, 43.

same is true about caring in general: genuine wholehearted caring about something implies focusing on the object of care. In other words, both caring and love are of the first order: they are directed at their object and not on any desires. I argued that they can generate second-order desires but these are of secondary importance.

Now, it seems that second-order desires can interfere with genuine caring and love. Thus, Frankfurtean self-evaluation as such – reflective agency that is built on second-order desires, second-order volitions and other reflective phenomena as commitment, identification and rejection – has a misplaced focus. It makes us focus on our own motivation – and in the end our own self – compromising our agency. In the last section I will suggest that instead of forming second-order desires, it is by forming beliefs through which effective self-control and self-management works.

But before I sketch this view, let me look at Elster concept of By-products a bit more in detail. In his book *Sour Grapes*[259] he devotes a chapter to this concept. He begins the chapter by claiming that "[S]ome mental and social states appear to have the property that they can only come about as the by-product of actions undertaken for other ends".[260] And he goes on by stressing that "[T]hey can never, that is, be brought about intelligently or intentionally, because the very attempt to do so precludes the state one is trying to bring about".[261] He also mentions two typical mistakes people make in connection with such states. First, he talks about the *moral fallacy of by-products*, which refers to the case when people try to bring about desirable states that are essentially by-products, and the *intellectual fallacy of by-products* which refers to the case in which people explain the presence of a state as "the result of action designed to bring it about – even though it is rather a sign that no such action was undertaken"[262] – since it is a state that is essentially a by-product.

[259] Elster, Chapter 2, 43-109.
[260] Ibid., 43.
[261] Ibid.
[262] Ibid.

Before I go on, let me clarify one point: Elster's term, "states that are the by-products of actions" suggests that there is a certain action that we perform, of which the given desired state is a by-product. For example, one exercises in order to be healthy, not because of getting tired. This would suggest that the state of being tired is a by-product of our action (exercising). However, this is not what Elster has in mind. This is rather the case of foreseen consequences of an action. One can easily achieve the state of being tired, by exercising for example. That is, foreseen consequences of an action can be easily brought about by performing the given action. But the case is radically different with the states that are essentially by-products. They are not foreseen consequences of a single action which we can simply perform and thus bring about the state. Rather, it might be better to say that some states are the by-product of our agency as such; not of a single action of ours but of the general features of our agency. To put it simply, they are not the by-products of any certain actions performed but rather, they are the result of *the way we perform our actions.*

To show what exactly this means, let me look at two paradigm examples of states that are essentially by-products. Elster discusses two examples of such states, one of them – as I already mentioned – is the case of insomnia; the other is taken from Stendhal's life history. As we know it from his diary, Stendhal was obsessed with "the idea of *becoming natural*".[263] This is a paradigmatic case of "willing what cannot be willed":[264] one cannot intelligently, intentionally become natural, since the state of being natural is characterized by spontaneity and the lack of trying. Strangely enough, there is a need to be *indifferent* about being natural, since it is precisely a concern for it that prevents its occurrence. Stendhal himself tried to solve the problem by adopting the following maxim: "I shall be certain to succeed if only I learn to show my indifference".[265] But, as Elster notes, this concept is "contradictory in

[263] Ibid., 44 (italics in the original).
[264] Elster takes this phrase from Leslie Farber, see Elster, 44.
[265] Quoted by Elster, 45.

terms"[266], since being indifferent implies a lack of intentionality, so one cannot act intentionally and be indifferent at the same time. Perhaps one might try to fake indifference to trick the process but Elster thinks that this cannot work either: one has to be *genuinely indifferent*, not just pretend that one is.

Elster's other paradigm example of a state that is essentially a by-product is sleep: the more someone tries to fall asleep, the less chance he has that he will do so, since his attempt itself to bring it about interferes with its occurrence.[267] A strategy would be to appear to be indifferent, but again, this cannot work for the same reason as noted above in connection with being natural. Faking cannot work either; faking itself is an activity that will prevent the desired state from occurring.

Now, Elster thinks that we cannot bring about states that are essentially by-products directly or indirectly. Thus, it seems that we cannot bring them about at all; we are at mercy whether they occur or not. Elster discusses in detail technologies of self-management devised to bring about such states. I will not discuss in detail his arguments; I would only like just to emphasize that they point to a strong skepticism towards the efficacy of such self-management methods. What seems to be the upshot of this discussion that there is a condition to reach a certain desired state: in case of being natural it is spontaneity and in case of sleep it is the gradual loss of consciousness. Both of them can be characterized as a state of *indifference*: spontaneity as the lack of a concern with what one should do and the loss of consciousness as not focusing on falling asleep. Now, the most important point here is the following: this indifference has to be *genuine*. One cannot fake it or command oneself to have it; one genuinely has to be in such an indifferent state.

This discussion has an important and surprising consequence for a Frankfurtian theory action. As we have seen, Frankfurt argues that *reflective self-evaluation* is necessary for

[266] Ibid., 45.
[267] Ibid., 45-46

being a person and having freedom of the will, which suggests that it is necessary to exercise some kind of second-order self-control over our motivation. In other words, for Frankfurt, the biggest enemy of full-blown agency seems to be indifference, or at least a special type of indifference he dubs wantonness. However, it is not at all obvious that indifference is always a big enemy of agency. As Elster's examples show, sometimes it is precisely a type of indifference that we need in order to be effective agents, and sometimes it is conscious, second-order control that interferes with agency.

Nevertheless, Elster's claim is too strong. Though there are states that we cannot simply achieve intentionally, it is not true that there aren't *any* methods to reach them at all. Donald C. Hubin argues that Elster's conclusion is too strong; it is not true that there are no methods at all to bring about certain desirable states. He also argues that these states can be reached through "sufficiently sophisticated methods of psychosurgery".[268] In the last section I will suggest a method, but it is not intended to be a sophisticated method of psychosurgery. Rather, I will describe a rather natural attitude which is based on openness and self-trust and claim that we can reach this kind of attitude by forming certain kind of beliefs (rather than second-order desires). Thus, the possibility of self-management and self-control is not jeopardized completely.

7.3 A Method for Self-Management

Second-order desires can make us aware of the fact that our will is integrated and that our volitions are in place as in the case of a wholehearted agent. But in less lucky cases they make us aware of the opposite: that we lack integrity and that our volitions are misguided. In this role second-order desires are ultimately important, but they are ineffective as a self-management strategy. Rather, they are the symptom of a volitional disorder which we cannot

[268] See Hubin, Donald C., "Of Bindings and By-Products: Elster on Rationality", *Philosophy and Public Affairs* 15, No. 1 (Winter 1986), 93.

cure with the help of them. Second-order desires in unhappy cases are not therapeutic. What we need is something different, as I argued in the last chapter. We need to give up the belief that we have a volitional essence, that is, we need to give up the belief that some desires are somehow the part of our nature, that they are constitutive of our personal identity. Instead, we need a certain kind of openness, a belief that it is not the case that our desires are fixed and that we just have to accept them.

When someone wishes that he wanted something – to care about himself more, or want to do kind things more often – it is a wish in the sense that he does not have the relevant first-order desires, at least not in the robust, action-guiding form. In such cases the right strategy does not depend on our second-order desires, which, as I claimed earlier are the symptom of a volitional disorder. Rather, one should simply *believe* that one is *capable of* having those desires at which one's second-order desires are directed. That is, the person who wishes that he cared more about himself has to believe that he is capable of caring more about himself and the person who wishes that he was a better person has to believe that he is capable to becoming a better person.

We saw in the earlier section that our second-order desires might have a negative effect on our agency. Notoriously, as in Elster's examples, wanting something to be the case can interfere with its occurrence. However, our beliefs – whether they are in place or not – can do two things: they can either 1) make it possible for some first-order desires to occur spontaneously or 2) they can prevent their spontaneous occurrence. I will argue that the biggest obstacle to using this indirect strategy above to "install" or "remove" first-order desires is the belief that we have a motivational or volitional essence. Such belief makes positive self-change impossible by postulating a set of desires which are constitutive of who we are. Rather, what we need is a certain kind of openness, that is, one should understand oneself as someone who is capable of having the "desired desires". I argued in my

dissertation that what we care about is not under our voluntary control but this does not have to imply that self-control and self-management are impossible. There is an effective tool at our disposal: to form our beliefs in a particular way. A theory of self-control and self-management can be developed based on forming beliefs instead of forming second-order desires.

In cases of volitional disorder, as for example the case of the unwilling addict, one forms second-order desires without having first-order motivation (at least in the robust sense). For example, someone might wish that he wanted to live in a more healthy way and yet he might recognize that he does not care about himself sufficiently enough to make this desire effective in action. As I defined it in the first section of this chapter, when someone suffers from a volitional disorder, one wants to do the thing he wishes he did not want to do and one does not want to do the thing he wishes he wanted to do. For example, the unwilling addict wants to take the drug badly (which he wishes he didn't want to do so badly) and he does not want to fight his addiction enough (which he wishes he wanted to do more).

Can we handle situations as this one by forming second-order desires? As I indicated above, in some cases this strategy is self-defeating. The more I want to have robust, genuine first-order caring, the less chance I have to develop it. Thus it is not possible to install or generate first-order motivation this way. What are we left with in this situation? Does not the theory on the relevance of caring to agency imply a too dark, pessimistic view on agency? Does not it imply that after all, people happen to care or not to care about certain things, and that in turn this is what determines what they will end up doing? What is more, this theory also seems to imply that one's values, and also how one evaluates one's own self also depend on what one happens to care about. But this view does not necessarily involve such a dark, pessimistic view on agency. One can employ a certain type of cognitive strategy, which is based on forming beliefs of a certain kind and directing one's attention in a particular way.

For effective self-management, the only thing we need is that the agent believes that he is *capable of* caring about something in the deep sense and is able to direct his attention in a way that makes it possible for the relevant first-order desire to occur spontaneously. If you believe that you are capable of wanting the right thing, you make the spontaneous occurrence of a first-order desire possible. This way, you are capable of having a certain degree of indirect control over you desires, while preserving their spontaneity.

Now, the biggest obstacle to using this strategy effectively is the belief that we have a motivational or volitional essence. That is, when for example the addict believes that after all, his desire for the addictive substance has become an important part of his identity to the point that this is what he cares about most, the result will be that he increases the presence of the desire to take the drug in his mind, he gives it power, and lets it overcome him and move him all the way to action, that is, to taking the drug. Besides, the addict might also believe that he simply does not care about himself enough to change himself; and this belief prevents the spontaneous occurrence of relevant first-order motivation. Now, to believe that we have a more or less stable motivational essence is an obstacle of self-management. In this example, the addict simply believes that his desire for the drug is central to his identity, and that a concern for his own self is not. But what if he realizes that *there is no such thing* that we happen to care about most? Once we realize that our desires are not necessarily part of a fixed volitional essence, we also realize that they do not have the power that we attribute to them. We will also recognize that our desires can be different and we can have a great deal of self-control over our own motivation by forming certain beliefs.

Of course, in positive cases it makes no sense to question the power of caring or love, for example if somebody thinks that his love for his children is central to his identity and that his actions are determined by it, and that he cannot but do what is good for them, there is nothing bad about his belief, whether or not it is correct. But the addict's case is different. He

might think that after all, he *is* an addict; he might feel so overwhelmed by his desire to take the drug that he feels himself being simply reduced to this desire. Letting his desire to overcome him, he will end up taking the drug. But what if, in a moment of insight, this addict simply realizes that he is not identical to his strong desire to drink in the sense that it does not constitute his volitional essence? What if he realizes that he does not have such an essence? As a result of such insight, he might realize that he is not at mercy of what he happens to care about or not. If we simply attribute a power to our desires, we simply let them use it over us. If one surrenders to a desire, it need not be because it has such an irresistible power, but by believing that it does we make it behave as it really did. The addict should think that his desire to take the drug does not have the power he attributes to it and that he is capable of developing a contrary concern not to take the drug and care about himself more and thereby he can, make the occurrence of a spontaneous desire not to take the drug more probable.

To conclude, though first-order desires and thus deep caring are similar to Elster's states that are essentially by-products and as a result we might think that we cannot achieve it merely by a decision, a caring-based view does not necessarily involve skepticism about self-control and self-management. True, we cannot simply decide to start to care about something. But as I suggested, we can try to use a different strategy through which we can change ourselves. Through forming beliefs, we can create circumstances that permit the spontaneous occurrence of some desires and prevent the spontaneous occurrence of some others. It is not tight second-order self-control. It is a subtle way of having control over our desires that makes it possible to preserve their spontaneity.

Conclusion

This dissertation centered around two keywords: reflectivity and caring, and their relevance for human life and human agency. Of course, with such simplification, it is impossible to do justice to the richness of Harry G. Frankfurt's work. However, I tried to follow this simple framework to interpret his concepts and the complex relation between them and I tried to grasp his ideas in one single statement: human beings are reflective and they care about things.

These two abilities are centrally important for human life and agency as such. Also, closely related and connected together. I agree about this basic statement, but differ from Frankfurt as regards their definition and the relation between them. For Frankfurt, the ability to care about things requires reflectivity, that is, it requires the ability to form second-order desires. During the course of this dissertation I have been trying to argue that for a number of reasons, the significance of second-order reflectivity is not as centrally important as Frankfurt supposed. Though we are reflective creatures, our reflectivity is rooted in what we care about, or in other words, in our personal values. Caring in this sense is not constituted by second-order desires but belongs to the first order. I have also been arguing that this particular ability

– to care about things in the sense of first-order evaluation – is the basis of our second-order desires. The fact that we care about things in the first-order sense is the reason why we have second-order desires and develop reflective caring. Our second-order desires can be the manifestations of something deeper: what we care about most in life. Of course, one might have second-order desires without really caring about the particular thing; and this raises important issues as I tried to show in the last chapter. I am aware that I might not have been able to fully discuss the complex and problematic relationship between reflectivity and first-order caring, but I hope I have shown that the latter is deeper and more significant.

Towards the end of the dissertation I argued that second-order desires can be particularly unhelpful in certain cases, as in the classic case of the unwilling addict. In such cases they are not something positive, but rather, the symptom of a deep problem with the volitions of the agent. Though they are revealing in these cases, they cannot be used as part of a self-management strategy. Rather, I suggested a more subtle way to do that, which is more similar to a type of art that second-order self-control.

In one of the sections of his chapter on states that are essentially by-products, Elster focuses on the problem of choice and intention in art.[269] He thinks that "making a work of art is an intentional action, a series of choices guided by a purpose", which is "to condense and convey some specific aspect of human experience within the discipline created by a technical framework".[270] An attempt to impress the audience distracts the artist from the real purpose of art. If in the end he manages to impress his audience, it will be the by-product of his attempt to create a genuine work of art. That is, the constant aim of the artist has to be, as Elster puts it, "getting it right".[271]

[269] Elster, 77-86.
[270] Ibid, 77.
[271] Ibid.

When the artist works toward his goal, he has to reduce a huge set of words, musical notes or brushes of paint. Initially, there is "an infinity of possible configurations"[272] that has to be first reduced to a feasible set to work with. Let me quote Elster to illustrate his view on artistic creation:

> First, the artist must cut down the feasible set to a more manageable size, by imposing additional constraints. Secondly, he has to exercise his creative gift proper, by choosing within the reduced feasible set some specific configuration of elementary units.[273]

However, he stresses that we cannot talk about choosing directly among the alternatives; this would not be rational, "for there are just too may possibilities to survey".[274] How does the artist choose then? Obviously, this is what artistic gift is good for. There is a need for artistic inspiration; that is, there is a need for a certain kind of spontaneity. Not only would it be irrational to survey all the possibilities because of their large number, it would also kill spontaneity and thus inspiration.

On the basis of my discussion of second-order desires and caring, it seems that second-order desires themselves do not facilitate spontaneity. First-order desires can spontaneously develop and disappear; but the more I want this to happen through forming a second-order desire, the less chance I have that it will. As I argued in the last chapter, second-order desires cannot be used effectively as part of a self-management strategy; on the contrary, they can be a symptom of a volitional disorder.

Instead of second-order control, *art* might be a good model for agency and for the method of self-management I tried to sketch in the last chapter. We cannot say that the artist is not in control of his actions, especially not in the case when he fulfills his goal and manages to express an aspect of life and his experience. However, being in control does not involve a second-order process; on the contrary, a second-order concern with his first-order

[272] Ibid, 78
[273] Ibid, 79.
[274] Ibid.

processes might be an obstacle to accomplish the goal at hand. Being in control does not have to involve second-order control, only that the agent's beliefs are in place. In the artist's case this means focusing on his goal and being open to the particular experience or aspect of life which we is trying to express. This way he makes it possible to become inspired. It could be argued that one also needs to become inspired in order to be a successful agent. In case of action in general, this means we need a certain type of openness that makes it possible for the relevant spontaneous motivation to develop. This raises the problem that some might be more gifted and talented for life than others. But art can always be cultivated and one can always improve one's skills.

When the artist is able to let his inspiration work for him, he will be "getting it right". In the title of his Tanner Lectures, *The Importance of Taking Ourselves Seriously and Getting It Right*, Frankfurt uses the very same expression. But when will we be able to get it right? Perhaps the two phrases in the title are mutually incompatible. When we take ourselves (too) seriously, we will not be able to get it right. That is, when we are too much involved in our agency and motivational processes through second-order control, spontaneity will be missing from our first-order processes and we will also lack openness that makes it possible for the "desired desires" to occur. Second-order processes still have an important role. Sometimes we have to know what kind of a person we wish to be, that is, sometimes we form second-order desires; we have desired desires. But it is a delicate process of letting the desired first-order desires develop spontaneously and work for us. This requires not exercising too tight self-control and a certain type of openness as well: a belief and self-trust that after all, we can genuinely be the person we would wish to be. Thus, sometimes we will get it right when we do not take ourselves too seriously.

Bibliography

Arpaly, Nomy. "On Acting Rationally Against One's Best Judgment". *Ethics* 110 (April 2000), 488-513.

——— *Unprincipled Virtue: An Inquiry into Human Agency*. Oxford: Oxford University Press, 2002.

Bratman, Michael E. "Hierarchy, Circularity and Double Reduction". In *Contours of Agency: Essays on Themes from Harry Frankfurt*, ed. Buss, Sarah and Overton, Lee, 65-85. Cambridge, Massachusetts: The MIT Press, 2002.

———. "Identification, Decision and treating as a Reason". *Philosophical Topics*, 24 (1996), 1-18.

———. *Structures of Agency*. Oxford: Oxford University Press, 2007.

Cavell, Marcia. *Becoming a Subject*. Oxford: Oxford University Press, 2007.

Clement, Grace. *Care, Autonomy and Justice*: *Feminism and the Ethics of Care.* Boulder, Colo: Westview Press, 1996).

Cuypers, Stefan E. "Autonomy beyond Voluntarism: In Defense of Hierarchy". *Canadian Journal of Philosophy* 30, No. 2 (June 2000), 225-256.

Damasio, Antonio R., *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Avon Books, 1994.

Darwall, Stephen. *Welfare and Rational Care*. Princeton: Princeton University Press, 2002.

Davidson, Donald. *Essays on Actions and Events*. Oxford: Claredon Press, 1980.

Dennett, Daniel C. "I Could Not Have Done Otherwise – So What?" *The Journal of Philosophy* 81, No. 10 (October 1984), 553-565.

Elster, Jon. *Sour Grapes*: *Studies in the Subversion of Rationality* Cambridge: Cambridge University Press, 1983.

Erikson, E. *Young Man Luther: A Study in Psychoanalysis and History*. New York: Norton, 1958.

Ferrero, Luca. "What Good is a Diachronic Will?" *Philosophical Studies* 144, Vol. 3 (2009), 403-430.

Foster, Gary. "Bestowal without Appraisal: Problems in Frankfurt's Characterization of Love and Personal Identity". *Ethical Theory and Moral Practice* 12 (2009), 153-68.

Frankfurt, Harry G. "Alternate possibilities and moral responsibility". In *The Importance*

of What We Care About, 1-10. Cambridge: Cambridge University Press, 1988, reprinted from *Journal of Philosophy* 96 No. 23. (December 1969), 829-839.

———. "Autonomy, Necessity and Love". In *Necessity, Volition and Love*. Cambridge: Cambridge University Press, 1999, 129-154.

———. "Concerning the freedom and the limits of the will". In *Necessity, Volition and Love*, 71-81, reprinted from *Philosophical Topics* 17 (1989).

———. "Disengaging Reason". In *Reason and Value: Themes from The Moral Philosophy of Joseph Raz*, ed. Wallace, Pettit, Scheffler and Smith, 119-129. Oxford: Claredon Press, 2004.

———. "Three Concepts of Free Action". In Frankfurt, *The Importance of What We Care About*, 47-57, reprinted from *Proceedings of the Aristotelian Society* 49 (1975), supplementary volume, 95-125.

———. "The Faintest Passion". In *Volition, Necessity and Love*, 95-107, reprinted from *Proceedings and Addresses of the American Philosophical Association* 66/3 (1992).

———. "Freedom of the Will and the Concept of a Person". In *The Importance of What We Care About*. 11-26. Cambridge: Cambridge University Press, 1988, reprinted from *Journal of Philosophy* 98, no. 23 (January 1971), 5-20.

———. "Identification and Externality". In *The Importance of What We Care About*, 58-69, reprinted from *The Identity of Persons*, ed. Rorty, Amelie, 239-253. Berkeley: University of California Press, 1977.

———. "Identification and Wholeheartedness". In Frankfurt, *The Importance of What We Care About*, 159-176, reprinted from *Ethics* 98 (1987) 170.

———. "The Importance of What We Care About". In *The Importance of What We Care About*, 80-94, reprinted from *Synthese* 53, No. 2 (1982), 257-72.

———. "Necessity and Desire". *Philosophy and Phenomenological Research* 45, No. 1 (September 1984), 1-13.

———. "On Caring". In *Necessity, Volition and Love*, 155-180.

———. "On the necessity of ideals". In *Necessity, Volition and Love*, 108-116, reprinted from *The Moral Self*, ed. Noam, G. C. and Wren, T., 16-27. Cambridge, Massachusetts: The MIT Press, 1993.

———. "On the usefulness of final ends". In *Necessity, Volition and Love* 82-94. Reprinted from *Iyyun, the Jerusalem Philosophical Quarterly* 41 (1992), 3-19.

———. "The Problem of Action". In *The Importance of What We Care About*, 69-79, reprinted from *American Philosophical Quarterly* 15 (1978).

———. *The Reasons of Love*. Princeton: Princeton University Press, 2004.

———. Replies to Michael E. Bratman, J. David Velleman, Gary Watson, T. M. Scanlon, Richard Moran, Susan Wolf, Barbara Heman and Jonathan Lear. In *Contours of Agency: Essays on Themes from Harry Frankfurt*, ed. Buss, Sarah and Overton, Lee, 86-90, 124-28, 160-64, 184-88, 218-25, 245-52, 275-78, 293-28.

———. *Taking Ourselves Seriously and Getting It Right*. Stanford: Stanford University Press, 2006

———. "What We are Morally Responsible For". In *The Importance of What We Care About*, 95-103, reprinted from *How Many Questions? Essays in Honor of Sidney Morgenbesser*, ed. L.S. Cauman, Isaac Levi, Charles D. Parsons, Robert Schwarz. Indianapolis: Hackett, 1983.

Fromm, Erich. *The Art of Loving***.** New York, Harper: 1956.

Green, O. H. "Is love an Emotion?" In *Love Analyzed*, ed. Lamb, Roger E., 209-225. Boulder, Colo: Westview Press, 1997.

Haji, Ishtiyaque and Cuypers, Stefan E. "Moral Responsibility, Love and Authenticity". *Journal of Social Philosophy* 36, No. 1 (Spring 2005), 106-126.

Helm, Bennett W. *Love, Frienship and the Self: Intimacy, Identification and the Social Nature of Persons.* Oxford: Oxford University Press, 2010.

Hume, David. *A Treatise of Human Nature*. Oxford: Oxford University Press, 2000.

Inwagen, Peter van. *An Essay on Free Will*. Oxford: Claredon Press, 1983.

Hubin, Donald C. "Of Bindings and By-products: Elster on Rationality". *Philosophy and Public Affairs* 15, No. 1 (Winter, 1986), 82-95.

James, William. *The Varieties of Religious Experience: A Study in Human Nature*. New York: Simon & Schuster, 2004.

Janaway, Christopher. "Schopenhauer's Pessimism". In *The Cambridge Companion to Schopenhauer*, ed. Janaway, Christiopher, 318-344. Cambridge University Press, 1999.

Jaworska, Agieszka. "Caring and Internality". In *Philosophy and Phenomenological Research* 74, No. 3 (May 2007), 529-568.

———. "Caring and Full moral Standing". *Ethics* 117 (April 2007), 460-497.

———."Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value". *Philosophy & Public Affairs* 28, No. 2 (Spring 1999), 105-138.

Kalis, Annemarie. *Failures of Agency*. PhD Dissertation. Utrecht University 2009.

Kane, Robert, *The Significance of Free Will*. New York: Oxford University Press, 1996.

Kolodny, Niko. “Love as Valuing a Relationship”. *The Philosophical Review* 112, No. 2, (April 2003), 135-189.

———. Review of *The Reasons of Love* by Harry G. Frankfurt (Princeton: Princeton University
Press). *The Journal of Philosophy* 103, No. 1 (January 2006), 43-50.

Lear, Jonathan, “Love’s Authority”. In *Contours of Agency*, ed. Buss, Sarah and Overton, Lee 275-292.

Lehrer, Keith. *Self-Trust: A Study of Reason, Knowledge and Autonomy*. Oxford: Claredon Press, 1997.

Lippert-Rasmussen, Kasper. “Identification and Responsibility”. *Ethical Theory and Moral Practice* 6, No. 4 (December 2003), 349-376.

Locke, John. *An Essay Concerning Human Understanding*. Oxford: Claredon Press, 1894.

MacIntyre, Alasdair. “Comments on Frankfurt”. *Synthese* 53 No. 2 (November 1982), 291-294.

Mele, Alfred R. *Autonomous Agents*. Oxford: Oxford University Press, 1995.

———. *Motivation and Agency*. Oxford: Oxford University Press, 2003.

Moran, Richard: “Frankfurt on Identification: Ambiguities of Activity in Mental Life”. In *The Contours of Agency*, ed. Buss, Sarah and Overton, Lee, 189-217.

Murdoch, Iris. “Against Dryness”. In *Existentialists and Mystics*. London: Penguin Books, 1997, 287-295.

———. Murdoch, Iris, “The Sublime and the Good”. In *Existentialists and Mystics*, 261-287.

Noddings, Nel. “Caring”. In *Justice and Care. Essential Readings in Feminist Ethics*, ed. Held, Virginia, 7-31. Boulder, Colo: Westview Press, 1995.

Nussbaum, Martha C. “Nietzsche, Schopenhauer and Dionysus”. In *The Cambridge Companion to Schopenhauer*, 344-375.

———. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press, 2001.

Pink, Thomas. *The Psychology of Freedom*. Cambridge: Cambridge University Press, 1996.

Rosati, Connie S. “Preference-Formation and Personal Good”. *Royal Institute of Philosophy Supplements* 81, Supp. 59 (2006), 33-64.

Plato. *The Symposium*. London: Penguin Books, 1951.

Sartre, Jean Paul: *Existentialism and Humanism*. London: Methuen 1955.

Scanlon, T. M. “Reasons and Passions”. In *The Contours of Agency*, ed. Buss, Sarah and Overton, Lee, 165-183.

Schechtman, Maya. “Self-expression and self-control”. In *The Self?*, ed. Galen Strawson, 45-63. Malden, MA: Blackwell, 2005**.**

Shoemaker, David W. “Caring, Identification and Agency”. *Ethics* 114, No. 1 (October 2002), 88-118.

Schopenhauer, Arthur. *The World as Will and Representation*. Translated by E. F. J. Payne. New York: Dover Publications, 1969.

Simmel, Georg. *Schopenhauer and Nietzsche*. Amherst: University of Massachusetts Press, 1986.

Slote, Michael. “Moral Sentimentalism”. *Ethical Theory and Moral Practice* 7, No. 1 (March 2004), 3-14.

Soble, Alan. “Union, Autonomy, and Concern” In *Love Analyzed*, ed. Lamb, Roger E., 65-93.

Solomon, Robert C. *In Defense of Sentimentality*. Oxford: Oxford University Press, 2004.

———. “On the Passivity of Passions”, in *Not Passion’s Slave: Emotions and Choice* (Oxford: Oxford University Press, 2003

———. “Emotions and Choice”, in *Not Passion’s Slave*, 3-24, reprinted from *Review of Metahysics* 28, No. 1 (September 1973)

de Sousa, Ronald. *The Rationality of Emotion*. Cambridge, Mass: The MIT Press, 1987.

Strawson, Galen. “Against Narrativity”. In *The Self?*, ed. Strawson, 63-87. Malden, MA: Blackwell, 2005.

Stump, Eleonore. “Sanctification, Hardening of the Heart, and Frankfurt’s Concept of Free Will”. *The Journal of Philosophy* 85, No. 8 (August 1988), 396-420.

Vellemann, David J. “Identification and Identity” in *The Contours of Agency. Essays on Themes from Harry Frankfurt*, ed. Buss, Sarah and Overton, Lee, 91-124.

———. “Introduction”. In *The Possibility of Practical Reason*. Oxford: Oxford University Press, 2000, 1-31.

———. “Love as a Moral Emotion”. *Ethics* 109, No. 2 (January 1999), 338-374

———. “What Happens when Someone Acts?” In *The Possibility of Practical Reason*. Oxford: Oxford University Press, 2000, 123-143.

Wagner, Zoltan. “Intentions and Agency”. *Filozofia*, 64 (2009), 739-747.

———. “Practical Reason and the Work of the Will”. *Croatian Journal of Philosophy* 19 (2007), 93-102.

Watson, Gary. “Free Action and Free Will”. In *Agency and Answerability*. New York: Oxford University Press, 2004, 161-197, reprinted from *Mind* 96 (April 1987), 145-72.

———. “Free Agency”. In *Agency and Answerability*, 13-32. Originally *Journal of Philosophy* 72 (1975), 205-20.

———. “Soft Libertarianism, Hard Compatibilism”. In *Agency and Answerability*, 197-215, reprinted from *Journal of* Ethics 3 No. 4 (1999) 351-65.

———. “The Work of the Will”. In Watson, *Agency and Answerability*, 123-157. Reprinted from *Weakness of the Will and Practical Irrationality*, ed. Stroud, Sarah and Tappolet, Christine, 172-200. Oxford: Oxford University Press, 2003.

———. “Volitional Necessities”. In: *Agency and Answerability*, 88-122. Reprinted from *The Contours of Agency*, ed. Buss, Sarah and Overton, Lee, 129-159.

Wolf, Susan. *Freedom within Reason*. Oxford: Oxford University Press, 1990.

———. “Sanity and the Metaphysics of Responsibility”. In *Responsibility, Character and the Emotions*, ed. Schoeman, Ferdinand, 46-62. Cambridge: Cambridge University Press, 1987

———. “The True, the Good and the Lovable”. In *The Contours of* Agency, 227-244.

Willigenburg, Theo Van. “A Non-Reductive Analysis of the Normativity of Agent-Relative Reasons”. *Ethical Theory and Moral Practice* 8 No.1 (April 2003), 45-62.

Yaffe, Gideon. “Free Will and Agency at its Best”, *Nous* 34, No. 4, Supplement 14, 2000, 203-229.

———. *Liberty Worth the Name: Locke on Free Agency*. Princeton: Princeton University Press, 2000.