THE ROLE OF THE HYDRAULIC MODEL OF DESIRE IN ACTION:
EXPLORING MINDFULNESS AS A MEANS TO ENHANCE
AUTONOMY IN ADDICTS

A Thesis
Presented to the
Faculty of
San Diego State University

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts
in
Philosophy

by
Tonya L. Warren
Fall 2011
The Undersigned Faculty Committee Approves the

Thesis of Tonya L. Warren:

The Role of the Hydraulic Model of Desire in Action: Exploring Mindfulness as a

Means to Enhance Autonomy in Addicts

Mark Wheeler, Chair
Department of Philosophy

Steve Barbone
Department of Philosophy

Wilbur Hansen
Department of Religious Studies

2 June 2011
Approval Date
DEDICATION

For my Mother, who always told me I could be anything I wanted to be.
ABSTRACT OF THE THESIS

The Role of the Hydraulic Model of Desire in Action: Exploring Mindfulness as a Means to Enhance Autonomy in Addicts
by
Tonya L. Warren
Master of Arts in Philosophy
San Diego State University, 2011

How we understand human action and define autonomous agency has ethical, moral, and legal implications for our theory of addictive motivation. Questions of responsibility often turn on whether the agent is deemed to have been in control of his or her actions at the time that they were performed. Understanding the relations among motives, autonomous agency, and action can also inform discussions regarding possible limitations and/or enhancement of willpower and self-control.

This thesis examines the possible impact addiction may have on agency and autonomy. It draws on theories from philosophy, psychology, addiction theory, and neuroscience. It explores the hydraulic model of desire, which posits desires as the causal vectors of force that move us to act. I reject this model based on its exclusion of the agent’s will in action. Instead, I develop a framework for a motivational hydraulic model that includes various motives besides desires that act as generic causal vectors of force. I distinguish between the “force” of a motive and its “strength.” Motivational strength or power refers to the relative effectiveness of a motive, which is a function of its basic force and the basic forces of other competing motives in a given motivational structure. In this model, we do not do what we most desire, as in the hydraulic model, but rather, what we are most motivated to do. More importantly, while some motives in the motivational system I propose are given, some are self-generated and allow for the agent to exert control over them. This model allows the agent’s will to play an active role and leaves open the possibility for autonomous action.

I distinguish between addiction and physical dependence and show that in some cases, addiction, which may or may not coincide with physical dependence, can impair autonomy. Failures to comply with one’s reflectively endorsed desires can happen autonomously or can result from impaired autonomy. I investigate the roles of self-control and willpower in action. While these can be used to support one’s intentions and resolutions about how one wishes to act, for the most part they do not improve autonomy. I argue that mindfulness meditation can be used to improve autonomy. I draw on a dual-processing model to help explain why we experience conflicting motives. This model also supports my argument that one can improve autonomy.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT ............................................................................................................................... v</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS ........................................................................................................ ix</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 INTRODUCTION ................................................................. 1</td>
<td></td>
</tr>
<tr>
<td>The Problem ............................................................................................................. 1</td>
<td></td>
</tr>
<tr>
<td>The Structure .......................................................................................................... 2</td>
<td></td>
</tr>
<tr>
<td>The Scope ................................................................................................................. 3</td>
<td></td>
</tr>
<tr>
<td>2 THE HYDRAULIC MODEL (HM) AND CRITICISMS ....................... 6</td>
<td></td>
</tr>
<tr>
<td>What Is It? ................................................................................................. 6</td>
<td></td>
</tr>
<tr>
<td>How Is It Supposed to Work? ............................................................. 6</td>
<td></td>
</tr>
<tr>
<td>Why Does It Fail? .................................................................................. 7</td>
<td></td>
</tr>
<tr>
<td>3 THE MHM AND HOW IT AVOIDS THE CRITICISMS OF HM .......... 12</td>
<td></td>
</tr>
<tr>
<td>How Is It Defined? .................................................................................. 12</td>
<td></td>
</tr>
<tr>
<td>Positive Motivational Base of Desire ............................................. 14</td>
<td></td>
</tr>
<tr>
<td>Motives versus Desires ............................................................................. 16</td>
<td></td>
</tr>
<tr>
<td>Desires as Motives ................................................................................... 17</td>
<td></td>
</tr>
<tr>
<td>Reasons as Motives .................................................................................... 21</td>
<td></td>
</tr>
<tr>
<td>Intentions Are Motives ............................................................................. 23</td>
<td></td>
</tr>
<tr>
<td>How Is It Supposed to Work? ............................................................. 25</td>
<td></td>
</tr>
<tr>
<td>Why Doesn’t It Fail in the Same Way HM Fails? .............................. 25</td>
<td></td>
</tr>
<tr>
<td>4 MHM AND AUTONOMY ........................................................................ 26</td>
<td></td>
</tr>
<tr>
<td>Chapter</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>5</td>
<td>WILLPOWER AND THE SELF</td>
</tr>
<tr>
<td></td>
<td>The Nature of the Will</td>
</tr>
<tr>
<td></td>
<td>The Nature of Willpower</td>
</tr>
<tr>
<td></td>
<td>The Two Decision-Making Systems</td>
</tr>
<tr>
<td></td>
<td>The Fragmented Self</td>
</tr>
<tr>
<td></td>
<td>The Authority of the Will</td>
</tr>
<tr>
<td></td>
<td>Willpower and Self-Control: A Means to Self-Unification?</td>
</tr>
<tr>
<td>6</td>
<td>MHM AND MEDITATION</td>
</tr>
<tr>
<td></td>
<td>Meditation</td>
</tr>
<tr>
<td></td>
<td>Goals of Meditation</td>
</tr>
<tr>
<td></td>
<td>Defining Mindfulness</td>
</tr>
<tr>
<td></td>
<td>Deautomatization</td>
</tr>
<tr>
<td></td>
<td>Improved Autonomy through Meditation</td>
</tr>
<tr>
<td>7</td>
<td>MHM AND ADDICTION</td>
</tr>
<tr>
<td></td>
<td>Defining Addiction</td>
</tr>
<tr>
<td></td>
<td>The Disease View</td>
</tr>
<tr>
<td></td>
<td>The Willpower View</td>
</tr>
<tr>
<td></td>
<td>Defect of Will?</td>
</tr>
<tr>
<td></td>
<td>Defect of Willpower?</td>
</tr>
</tbody>
</table>
Defect of Desire? .................................................................64
Denying Defects of the Will .........................................................65
Strength and Weakness of Will ......................................................65
Motives Change over Time .............................................................68
Addictive Desires ............................................................................70
The Unwilling Addict .......................................................................71
Impulses and the Impairment of Reflective Agency .........................74
8 CONCLUSION .................................................................................77
REFERENCES ..................................................................................80
ACKNOWLEDGEMENTS

This research project would not have been possible without the support of many people. The author wishes to express her gratitude to her supervisor, Prof. Dr. Wheeler who was abundantly helpful and offered invaluable assistance, support and guidance. Deepest gratitude is also due to the members of the SDSU philosophy department without whose knowledge and assistance this project would not have been successful.

The author wishes to express her love and gratitude to her beloved friends and family; for their endless support and understanding through the duration of her studies.
CHAPTER 1

INTRODUCTION

THE PROBLEM

How we understand human action and define autonomous agency has ethical, moral, and legal implications for theories of addictive motivation.¹ Questions of responsibility often turn on whether the agent is deemed to have been in control of his or her actions at the time that they were performed. Understanding the relations among motives, autonomous agency, and action can also inform discussions regarding possible limitations and/or enhancement of willpower and self-control.

In this thesis, I will defend a version of the hydraulic model and argue that, based on the conception of autonomy I present, meditation can improve autonomy. The hydraulic model aims to explain human action and autonomous agency in terms of desires that act as causal vectors of force that determine our actions. Harry Frankfurt and Donald Davidson are two notable philosophers who support this model. Others, such as R. Jay Wallace and Michael L. Corrado, charge that because this theory does not take the agent’s will into account...

account, it cannot provide an adequate explanation for human action. Later, I examine Wallace’s objections in further detail and provide an objection to his criticisms.

**THE STRUCTURE**

My thesis is organized as follows: in chapter 2, I begin by outlining the basic hydraulic model of desire (HM) and show how it fails. In chapter 3, I develop my version of the hydraulic model, which I will henceforth refer to as the Motivational Hydraulic Model (MHM). In it, I provide an account of motives that explains the roles of desires, self-control, and willpower in moving one to act. Having outlined my view, I consider a possible objection presented by R. Jay Wallace. I argue that his view mischaracterizes the role of the agent in the hydraulic model and is accurately represented in the MHM. In the third chapter, I describe a dual-information processing system that best describes the basis for motives in the MHM. I argue that the will has authority based on how it derives from this system. In the fourth chapter, I explore how autonomy cashes out in the MHM I propose. It is here that I introduce the notions of self-generated motives and argue that we have control over some of our motives. I also discuss features of self-control and how they relate to human action. I argue that self-control-skills (SCS) are a means to extend one’s will over time. Chapter 5 explores the nature of willpower and the will. I present a dual-processing model of decision making that helps explain simultaneous conflicting desires. Chapter 6 consists of discussions of mindfulness meditation as a method to improve autonomy. In chapter 7, I provide an overview of the major features of addiction and attempt to show how addiction might impair autonomy.
THE SCOPE

This thesis draws on theories from philosophy, psychology, addiction theory, and neuroscience. From the field of philosophy, I will include discussions from the theory of action. I will primarily concern myself with relevant theories of agency and autonomy. I will touch on questions of free will but have tried to present a view that does not conflict with the main schools of thought in that field. In terms of psychology, I draw mainly on cognitive-behavioral learning theories in providing explanations for the development of maladaptive behavior. I present the main theories of addiction to provide an account for the basic features of addiction. I also include the current definitions from the Diagnostic Statistical Manual of Mental Disorders to highlight the differences between the phenomena of addiction and substance dependence. I will not directly address questions of the etiology of addiction either in terms of responsibility or in terms of different individuals‘ possible susceptibility to addiction or substance dependence. This is because I am mainly concerned with how addiction might impair autonomy. To this end, I will focus on the features of addiction that apply equally to addictions to substances and also those that are termed behavioral addictions. For the same reason, I will not consider the effects of withdrawal from either physical or behavioral addictions. I will offer a brief description of the neurological aspects of addiction in the section on failures but will focus only on those that I think can be directly impacted by mindfulness. References will be provided for those who wish a more in depth exposition of the areas that I do not thoroughly treat in this thesis.

In terms of autonomy, there are two main ways to conceptualize it in the literature. First, one might conceive of autonomy in terms of self-government: to govern oneself means having and acting upon the preferences one wants to have. Or one might conceive of autonomy in terms of self-determination: to determine oneself means that one is the source of
the preferences upon which one acts. Self-government is a less demanding way to understand what Levy and others refer to as basic autonomy: “the minimal status of being responsible, independent and able to speak for oneself.” This is contrasted with ideal autonomy, “the state of maximal authenticity to which many aspire, and few achieve.” To say that having and acting on the preferences one wants to have is required for autonomy entails that anytime one acts on preferences one doesn’t want to have, one does not act autonomously is a mistake. I will argue that there are times when this situation occurs that the agent did act autonomously, though he or she may not have acted freely. However, it seems that the alternative might be interpreted in two ways. One way is to say that we must be the source of all of our preferences. This seems unrealistic. We must concede that at least some of our myriad preferences are given. Because we are not the source of all our preferences, I must reject this way of conceptualizing autonomy as well. The second way we might interpret self-determination is to say that we are the source of some of our preferences. While some would deny this altogether, I will argue that there are two ways in which the agent can be the source of his or her own preferences. This idea will be further explored in a later section of the paper. For now, I will assume the notion of self-determination entails that the agent is the source of some of his or her preferences. Those preferences that have their source in the agent I will call self-generated preferences.” I will offer this modified definition of autonomy: autonomy consists in being the source of some of one’s preferences and acting upon them in such a way that one attempts to bring it about that the self-generated

---


preferences are realized. It is not relevant for the purposes of this paper whether the addict had predispositions to addiction or whether he or she autonomously acquired the addiction. What matters is whether, once addicted, his or her autonomy is impaired.
CHAPTER 2
THE HYDRAULIC MODEL (HM) AND CRITICISMS

In this chapter, I begin with a description of the basic hydraulic model of desire. I first define the theory and what it aims to accomplish. Following this, I will show why it fails to achieve that aim and examine an argument by R. Jay Wallace. Wallace also argues that the hydraulic model fails, but I will show that while I agree that the HM fails, his argument is flawed.

WHAT IS IT?

The basic hydraulic model of desire belongs to a class of theories of action called the choice-theoretic view of human action. On this view, human action is the result of choices and these choices depend on the actor's preferences. The actor seeks to maximize his or her preferences and will always act accordingly. Preferences in the hydraulic model are desires. Acting against one's desires is not possible on this view. The hydraulic model says that in cases of competing desires, action will ultimately be determined by the strongest desire.

HOW IS IT SUPPOSED TO WORK?

The basic hydraulic model of desire sees action as the result of choosing among competing desires. On this model, whichever desire is the strongest is the one we act upon. Chief among proponents of this theory are Donald Davidson and Harry Frankfurt. Davidson

writes, —P1If an agent wants to do x more than he wants to y and he believes himself free to do either x or y, then he will intentionally do x if he does either x or y intentionally.‖ One of the benefits of this theory is that it avoids the need to rely on motives other than belief and desire to explain human action. This simplified theory reduces all motives, including judgment, to desires. For example, even when the agent judges what he or she has most reason to do, supporters of the hydraulic model say that if the agent acts on this judgment, it is because his or her strongest desire was to act in accordance with that judgment. The theory appeals to our intuitive understanding of human action. We do seem to have competing desires, and our actions often appear to ourselves and others to be guided by what we most wanted to do at the time. Clearly, there are some aspects of the hydraulic theory that are grounded in experience; therefore, it is worthwhile to investigate further the model and see what it may have to offer to our understanding of human action. I will begin with an argument against the view proposed by R. Jay Wallace.

**WHY DOES IT FAIL?**

In this section, I will explore an objection to the hydraulic model of desire. In his paper, —Addiction as a Defect of the Will: Some Philosophical Reflections,” R. Jay Wallace argues that addiction can be understood in terms of a volitional defect. He shows how this is possible through a careful investigation of the ways in which desires provide motivations to actions and how those motivations relate to one’s choices and intentions. He first critiques

---


7 Corrado, —Addiction and the Theory of Action,” 16.

8 R. Jay Wallace, —Addiction as a Defect of the Will: Some Philosophical Reflections,” Law and Philosophy 18, no. 6 (1999), 622.
the traditional hydraulic conception of desire and then goes on to provide an account of his
volitionalist model,” which he thinks solves some of the major problems he cites. The main
problem with the HM is that the agent’s will doesn’t enter the equation. “The actor’s
judgment, his evaluation of the alternatives, plays no active role.”⁹ I agree with these
criticisms, however, I aim to expose a flaw in Wallace’s argument.

Wallace has two main lines of argument against the hydraulic model: first, he charges
that it deprives us of the capacity for self-determination. Second, he says it does not
distinguish between desires we are merely subject to and the motivational states that we have
control over. This view, according to Wallace, entails that agents always do what they most
want to do. On this model, he sees deliberative agency as consisting, essentially, in two
distinct moments. There is first, the agent’s practical or evaluative judgment about what there
is most reason to do, or what it would be best to do on the whole; and there is, second, the
agent’s motivational state, which is a function of the causal strength of the desires to which
the agent is subject.”¹⁰ Wallace says what we need instead is a model that accounts for
weakness of the will but does not rely on the idea that it is the causal strength of the desire
that determines whether it is conformed with or not.

In his paper, Wallace considers a version of the hydraulic model that allows for self-
control techniques. He says these enable one to achieve control by devising strategies to
influence causally the motivational strength of the desires.”¹¹ He refers to this as the cold

---

¹⁰ Wallace, Addiction,” 635.
¹¹ Ibid.
shower paradigm” in which, for example, one might counter an overwhelming sexual desire by taking a cold shower.\textsuperscript{12} Wallace writes,

\begin{quote}
To suppose that self-control must always conform to the cold shower paradigm turns us into passive bystanders at the scene of our own actions. We don’t really determine which action we perform directly, rather we attempt to manipulate the psychological influences [What he calls psychological influences, here I take to be motives.] to which we are subject; in the hope that they will eventually bring it about that we do what we judge best.\textsuperscript{13}
\end{quote}

He says that this reduces agency, whatever is left of it, if anything, to a form of indirect self-manipulation in which we do not directly determine our actions.\textsuperscript{14}

In order to overcome the aforementioned shortcomings of the hydraulic model, Wallace proposes we instead adopt a volitionalist motivational psychology.\textsuperscript{15} As opposed to the hydraulic model, which consists in two moments, Wallace adds a third moment that goes beyond deliberative judgments and given desires. He calls this third moment “volition” and cites intentions, choices, and decisions as typical examples of volitional states. He distinguishes between two kinds of motivations: (1) Passive motivations: states or conditions we find ourselves in and are passive in regards to, and (2) Other motivations: choices and decisions that are directly up to us. These “other motivations” are distinct from the passive motivations because, he says, they are directly under the control of the agent.\textsuperscript{16} For Wallace, Agency […] is manifested in our exercising the capacity to choose for ourselves what we are to do.”\textsuperscript{17} Therefore, by making choices and decisions a volitional state that is up to the

\textsuperscript{12} Ibid., 636.
\textsuperscript{13} Ibid.
\textsuperscript{14} Ibid.
\textsuperscript{15} Ibid., 632.
\textsuperscript{16} Ibid., 636.
\textsuperscript{17} Ibid., 637.
agent, he thinks that we bring agency back into the equation. According to Wallace, this addendum to the hydraulic model solves the two main problems he cites earlier: the hydraulic model deprives the agent of the capacity for self-determination and fails to distinguish between motives that are passively received and those we have control over. In the next section, I will present an objection to Wallace’s criticisms of the hydraulic model.

In response to Wallace’s criticisms of the hydraulic model, I argue that if he allows for the implementation of self-control in the hydraulic model, it does not deprive one of the capacities for self-determination or fail to distinguish between motives that are given and those that are under our direct control as he supposes. I contend that self-control entails choices and decisions, which on his model constitute motivations that are under our control. If this is correct, then his third ‘volitional’ moment is already contained in the hydraulic model and his account does not add anything to the hydraulic model. Moreover, this also means that he has miscategorized the notion of a motivational state, which he describes as ‘function of the causal strength of the desires to which the agent is subject.” The motivational state is in part a function of the causal strength of desires but also includes other motivations.

Wallace charges that when an agent relies on self-control strategies such as the cold shower paradigm, the agent does not determine his or her actions but rather attempts indirectly to manipulate the mental landscape. The key word here is ‘attempt.” Contrary to Wallace’s charge that this diminishes or even destroys the possibility of agency, I argue that the capacity for planning, devising, and enacting these self-control strategies that ‘attempt’ to manipulate the influences we are ‘subject to’ is precisely what autonomy is based on. Self-control as Wallace describes it here presupposes reflective endorsement. Reflective
endorsement, I argue, expresses autonomy because it is a self-generated desire and self-control strategies that arise from self-generated desires likewise express one’s autonomy. Moreover, self-control strategies require a decision on the part of the actor to put them into effect. While reflective endorsement is merely another desire in the motivational system, the decision to enact self-control strategies is a kind of motive that will enter into the motivational system. Since these self-generated motives impact the motivational strength of other motives including those passively received, it is not clear how Wallace thinks we can reconcile the ideas of (1) being passive bystanders with (2) deliberately exerting self-control strategies. The fact that we can determine what we judge best and decide to attempt to manipulate these motives is in itself an exercise of our agency. Whether we succeed or not, the decision to attempt to struggle against a competing motivation and develop a self-control regime to that end presupposes self-generated motives that exert causal influence over one’s motivational state. To say that devising self-control skills whose purposes are to counter motives one does not endorse is a form of “indirect self-manipulation” is wrong. If anything, the reflective endorsement and the intentions and self-control regimes formed from that are what express our will and allow us directly to manipulate that which is alien to us. In the next section, I develop an account of the hydraulic model that incorporates the notion of motives as causal forces, which I will call the Motivational Hydraulic Model.
CHAPTER 3

THE MHM AND HOW IT AVOIDS THE CRITICISMS OF HM

HOW IS IT DEFINED?

The hydraulic model of desire to explain action has been criticized for not allowing for deliberative agency.\(^\text{18}\) This criticism is based on the hydraulic model’s interpretation of desires as causal vectors that determine our actions. This would entail that any action taken on the part of the agent was a direct result of forces that were irresistible to the agent based on the agent's actual psychological circumstances at that time. I agree with Wallace's assertion that this model fails to account for motives that are under the direct control of the agent.

The model I propose here expands upon the basic framework of the hydraulic model, introducing reflective endorsement, self-control, and willpower as parts of the motivational framework. I will argue that my appended model allows for deliberative agency. We are merely subject to the force of some desires. Whether or not we are aware of them, there are motives that subject us to their force.\(^\text{19}\) However, just because we are subject to certain motives in this way, we are not therefore impotent in relation to them. Since the strength of a given motive is a function of its force in relation to the other forces in the motivational system, as long as the motives over which we do have control (each of which has its own

\(^{18}\) Ibid., 633.

\(^{19}\) Phillip Pettit and Michael Smith, "Backgrounding Desire," The Philosophical Review 99, no. 4 (1990), 565.
force) are the strongest motives (relative to the others), we can effectively exert some degree of control over our actions.

I intend to show that we do not do what we most want to do, as the original HM proposes but, rather, what we are most motivated to do. On my version of the hydraulic model, “motives” are the generic type for vectors of causal force related to actions. Not all vectors of causal force are motives. Motives are psychological vectors of causal force and, more specifically, psychological vectors of causal force the possible effects of which are actions. Desires are one kind of motive in the motivational framework the relative forces of which depend on other motivational factors within that framework.

I begin with a description of how the strength of motives relates to the MHM. I distinguish between the “force” of a motive and its “strength.” Motivational strength or power refers to the relative effectiveness of a motive, which is a function of its basic force and the basic forces of other competing motives in a given motivational structure. Insofar as all motives have a basic force and operate in a similar way within the hydraulic system, they are of a like kind. They are all forces that push or pull the actor to act. The motives exert their basic forces on the agent, and the strength of a given motive depends on the forces of the other motives with which it is competing. Competing motives all play a role in how one will act, but which factors are effective depend on their relative strengths.

A motivational state of a motivational system is defined as the set of all the motives within a given agent at a given time. The synchronic relative strength of the motives is a function of the basic forces of the motives present in the mind of the agent at that time. A motivational system is defined as the set of all the motives within a given agent over a given interval of time, where the diachronic relative strength of the motives is a function of the
basic forces of the motives over that interval. The relevant set of motives includes all conscious and unconscious motives and both passive and active motives. The set includes dispositions, plans, some desires, intentions, and so on. I argue that agents can always retain regulative and guidance power over motivational systems and states because every motive is part of a wider network of motivations and dispositions that constitute the agent’s notion of self and, importantly, the relative strength of certain motives over which agents have control can always in fact be the strongest. Motivational strength is relevant to the discussion because it helps explain the causal influence of the actor on the motivational state.

**Positive Motivational Base of Desire**

If an agent can generate and control any of the motives constituting the motivational system, or if the agent can come to control such motives in spite of not having generated them, then the agent may have the capacity to control the total motivational system or, at least, significant dimensions of it. In other words, although there are surely factors of the motivational system that are passively received, such as some desires, once they enter the motivational system, it may be that they can become parts of the agent’s motivational system over which the agent exerts control. As such, they can contribute to one’s autonomous decisions by becoming a part of what Alfred R. Mele calls the “positive motivational base” of a desire. 20

Mele describes the positive motivational base of desire as “the collection of all occurrent motivations of the agent that make a positive contribution to the motivational

---

strength of a desire.” This is more or less the same as my concept of a motivational state, defined above, except that Mele defines his concept ultimately in terms of desire, whereas I define a motivational state in terms of motives only some of which are desires. Mele’s conception of the positive motivational base of a desire can be expanded to include all motivational factors in an analogous way. As with the factors of a motivational state, the composite of motivational factors that make up Mele’s positive motivational base of desire may include motivational factors of which the agent is not conscious but which still have motivational power over the agent. In other words, an agent may or may not be aware of all of the motives influencing the agent’s actions.

I will argue that some motives are “self-generated.” Unlike given motives, self-generated motives are under the direct control of the agent. This will be discussed further in the following chapter. If we accept that some motives are under the direct control of the agent, then agents can causally influence their own motivational states and systems. There are two kinds of motives that agents have control over in this way. First, there are self-generated motives. Second, there are given motives in the form of desires, which once reflectively endorsed, become the agent’s own. Such motives enter into the positive motivational base of a desire. Agents, I will argue, can directly affect the force of these motives, and by directly affecting their force, the agent is able to modify the relative motivational strength of the motive in question.

Two distinct ideas claims are important here: one is the idea that agents can directly control some of their motives. The other is that agents can directly affect the force and

---

relative strength of their motives. These ideas help explain how agents can change their motivational state and, hence, their motivational system.

**Motives versus Desires**

A key difference between my view and the basic version of the hydraulic model, the one criticized by Wallace, is that on my view not all motives are desires, although some desires are motives. While many philosophers equate the notions of desires and motivations, it matters whether we describe them as motivations or desires.\(^{22}\) A motive is — the disposition within the person to strive to approach a certain class of positive incentives (goals) or to avoid a certain class of negative incentives (threats).\(^{23}\) A desire can be seen as a sort of want or hope that a certain circumstance (or set of circumstances) comes to be, however, desires must be paired with belief in a relevant way in order for them to have motivational force. Equating desires and motives seems to entail that all desires are motives (however weak or strong) to act and that all motives are desires. It does seem strange to say that one would have a desire that did not entail at least a weak motivation, yet that is the case with some desires. For example, I may desire to attend X University, but if that desire is paired with the belief (right or wrong) that the university in question will not be accepting new Ph.D. students in my discipline, I will not act on the desire. If the desire does not push or pull me to act upon it, then it is not a motive. That does not mean, however, that all of our motives are desires. Even if we take it that avoiding negative threats consists of a desire to avoid pain, motives, as I will show, come in many forms, only one of which is desires. On my view, while some desires,

---


namely those paired with the correct beliefs, are motives, there are other kinds of motives that can also move one to act. To continue with the previous example, I may not have a desire to attend X University, however, I may have a motive to do so in the form of a sense of duty to my parents to attend this particular school, or a guarantee of a desirable job upon graduation from this school. In this case, even though I do not desire to attend the school, I have a motive that moves me to act in such a way that I bring it about that I gain admission to X University. Motives include some desires, but there are other types of motives that are equally important in the motivational equation that determines how we will act.

**Desires as Motives**

In the previous section, I made a distinction between desires and motivations. This distinction should not be taken to mean that desires have a special place in the realm of motives in terms of their potential for causing action. On the contrary, I argue here that the desires that act as motives do so in the same way as other kinds of motives.

Desires are often differentiated on the bases of their root cause, strength, how we experience them, or other various factors. Discussions usually aim to understand desires in order to shed light on why we succeed or fail in mastering them. The hydraulic model is no different in that regard. What is different about this model is that instead of creating distinctions between types of desires to explain the way they operate, it subsumes all desires under the category of possible motives. All desires have the potential to act as motives. What determines whether the desire in question acts as a motive is whether there is an concurrent belief that there exists the possibility that the object of desire is attainable or not. This way of understanding desires helps to explain why one might act on a desire that seems irrational or fail to act upon one that seems rational. For example, an addict who claims he or she wants to
abstain may have a concurrent belief (again, right or wrong) that abstinence is impossible will not be motivated to try to abstain. While it may seem rational to an outsider for the addict to take measures to abstain, if the addict truly believes it is impossible to do so, it would in fact be irrational for him or her to do so. Likewise, an agent might believe that by chanting the word “chocolate” throughout the day, he or she will win the lottery. While this belief might be completely delusional, what would seem like irrational behavior to an outsider would actually be completely rational in light of the beliefs of the agent. This way of conceptualizing motives and desires also helps to explain why some agents fail to act upon desires that they describe as persistent.

What is relevant about desires in terms of agency and autonomy is how we respond to them, and that depends upon their strength as motives. If one has a desire that is paired with a belief that the object of desire is attainable, then that desire acts as a motive. Lacking such a belief, or if the desire is paired with a belief that the object of desire is unattainable, the desire will lack causal force and cannot be included in the agent’s set of motives. For example, I may desire to go to another galaxy. If I believe that traveling to another galaxy is impossible, the desire will not move me to act. That does not necessarily mean that the desire is extinguished. It is possible that the desire will persist, and if my belief regarding this desire changes, it may become a motive. If I come to believe that traveling to another galaxy is in fact possible, the desire becomes a motive for me to act in such a way that I bring it about that I do so. These beliefs need not be rational, for even if they are not, it is rational for me to act upon them given my (possibly deluded) belief. I will call actions paired with beliefs that the object of desire is (more or less) attainable motivational desires (MD). From now on, when I refer to desires, it should be assumed that I am describing motivational desires. I
should also note that the motivational force of the desire relates to the strength of the belief. Beliefs vary in strength, and I may believe that there is absolutely no chance to realize the desire, 100% chance that I will, or anywhere in between. The stronger my belief is, the stronger the force of the desire of the motive, and the converse is also true. However, belief alone is not a motive. It must be linked to an appropriate desire in order to have causal force in the motivational system. Next, I will explain the motivational strength of desires in the MHM.

Like all motives, desires vary in motivational power, and some seem to be so powerful that they might even be seen as irresistible. Addictive desires are sometimes portrayed this way, or if they are not viewed as literally irresistible, at the very least theorists tend to place them in a separate class of desires because of their peculiar strength. For example, one feature of addictive desires that is often pointed to is the fact that they appear unresponsive to reasons. —Reason-responsiveness, [the capacity to regulate behavior based on one’s judgments about what one’s best reasons are for doing so] in turn, seems necessary for the feeling of being ‗in control‘ because desires, beliefs or attitudes which appear unresponsive to reasons are the ones over which one typically feels one lacks control.”

Even when one deliberates well, these desires do not seem to yield to one’s judgments about them.

---


25 See George Loewenstein, —Out of Control: Visceral Influences on Behavior,” Organizational Behavior and Human Decision Processes 65, no. 3 (1996), 272-292; Levy, —Autonomy and Addiction,” 427-448. A notable opponent to the idea that addictive desires are different than other types of desires is found in Bennett Foddy and Julian Savulescu, —A Liberal Account of Addiction,” Philosophy, Psychiatry and Psychology 17, no. 1 (2010), 1-22.

26 Edmund Henden, —What Is Self-Control?” Philosophical Psychology 21, no. 1 (2008), 73.
On the MHM, reasons are one kind of motive. Reason-responsiveness turns on whether reasons are what lead to actions. Because on the MHM, reasons are only one kind of motive among many that can move one to act; they cannot be seen as the single factor that causally moves one to act. It may be the case that the agent has good reason to act in such a way to bring about a certain outcome and he or she might be responsive to that reason, however, on the MHM, the agent may or may not act on that reason. That is because on the MHM, what will move the agent to act is the strongest motive. So while the agent may act contrary to the reason, that does not necessarily entail that the agent was not responsive to the reason. It may simply be the case that other motives were stronger than that particular reason. In the case of addictive desires, the addict may have good reason to quit using the substance. The addict might be responsive to that reason and act in such a way that the addict yet still fails. This failure is not a result of a lack of reason-responsiveness but, rather, a result of the prevailing strength of the motives competing with the reason the addict has to quit. Although in this case it may appear that the addictive desire was literally irresistible, in fact, it was only effective due to the motivational state of the agent at that particular time. No addictive desires will always be irresistible, since their strength is always determined by the strength of other occurrent motives. To sum up, addictive desires are causally efficacious when they are the strongest motive. Reason-responsiveness is not the only factor that determines whether the addictive desire will prevail or not because reasons are only one kind of motive that will move an agent to act. Addicts may be responsive to reasons and still fail to act in accordance with them.

In the next section, I will explore one kind of motive that is relevantly different, but regardless of their strength, desires—although they play an important part in the motivational
framework—operate in the same way as most other motives. On my view, addictive desires should be understood just as other types of motives are in terms of relative motivational strength (where the motives in competition include the desire itself, the reasons one has for acting or not acting on the desire, and the other motives present in the motivational state of the agent). Which of these motives is causally effective depends on the sum equation of motivational power at a given time. It may appear that it is the desire alone that is pushing one to act, but in fact other motives, including reasons, are at play. More importantly, some of these motives belong to the agent in a relevant way.

**Reasons as Motives**

Desires come in different forms, from the most basic such as the visceral desire for food when we are hungry to more complex ones like the desire to flourish as a human being. The fact that motivating desires can only be satisfied through action is enough to generate a motive to perform the action(s) necessary to acquire the desire. The simple desire provides a straightforward motive. The given desire to satisfy hunger gives rise to a motive for action, in this case, consuming food. In the case of a more complex desire, the actions needed to satisfy that desire may also be more complex. The question is, does the fact that I think I need to do A in order to acquire x provide a motivation for A? Mele’s answer is a resounding “No.” He writes, “Al’s belief that his acquiring x requires his doing A does not constitute motivation to A.” Contra Mele, I would argue that although Al’s belief that he must do A to acquire x may not result in a corresponding desire for A, it does make A a motive. Because the desire for x requires A, it provides a reason to A and reasons, like desires, are also motives. In the

---

27 This view is also supported by Bennett Foddy and Julian Savulescu, "A Liberal Account.”

case of reflective endorsement, endorsement becomes the basis for intentions about how one will act. To form an intention is to decide on a course of action. Therefore, when one reflectively endorses a first-order desire, not only is the endorsement itself a motive, it also provides a motive in the form of a reason to perform actions that will make the endorsement effective. The act of deciding to act in a certain way generates other motives. Reflective endorsement, when paired with the appropriate belief gives rise to an intention to x. Self-control and/or willpower are often at least one of the As that are required to obtain x. For this reason, I will include in the set of motives both self-control and will-power.

Self-control should not be equated with autonomy. Autonomy, as I have defined it is being the source of some of one’s preferences and successfully acting upon the preferences in an attempt to realize them. Although it may be tempting to equate self-government and self-control, self-control should be understood as the means to achieve autonomy and not as autonomy itself. Because self-control is generated from one’s reflective endorsement, which is an expression of autonomy, it should also be understood as an expression of one’s autonomy. Since autonomy requires being the source of some of one’s preferences and acting on them, self-control is only part of the equation. Self-generated motives are also required. To illustrate my point, one may have a reflectively endorsed desire to quit smoking and successfully act upon it in an attempt to make the desire effective but that does not necessarily mean that the action was an act of self-control. I have also included willpower in the set of means to achieve one’s reflectively endorsed desires. It could be that the agent in question successfully abstains from cigarettes relying on willpower alone. Willpower can support self-control but is not synonymous with it. I have also left open the possibility that

there may be other means to achieve the same end that do not entail self-control. So while it may be the case that all exercises of self-control entail a self-generated motive and therefore are autonomous to some extent, one may act autonomously without relying on SCS. This point will be especially relevant in the later discussion of Wallace’s argument against the hydraulic model.

**Intentions Are Motives**

Here I will discuss the ideas of resolutions and intentions put forth by Richard Holton and Edmund Henden. According to Holton, to form an intention is to decide on a course of action. In order for intentions to fulfill their function, they need to be relatively resistant to reconsideration. Holton relies in part on the work of Michael Bratman to describe the characteristics of intentions. Future-directed intentions are described as "controlling" and have "stability." "Controlling" means that the intention is formed by deciding to perform the action and, unless the intention is revised, will lead the agent to do so. "Stability" refers to the tendency of intentions to persist; in other words, they are fairly immune to reconsideration. Intentions, according to Holton, have motivational power, and while he admits that some other tactics may be needed to support the decision, "very often, intention is enough."

In his article, "What Is Self-Control?" Henden proposes similar idea which he calls "resolutions." According to Henden, resolutions have the unique feature that they are 

---

30 Ibid., 241.
32 Holton, "Intention," 244.
33 Ibid., 246.
special kind of commitment, part of whose function is to remain firm in the face of competing motivation the agent may expect to arise. On his view, resolutions are formulated prior to the actual temptation so by their natures are never caused by desires. I would not go so far as to say that intentions are formed in the complete absence of the desires one wishes to act counter to. It seems more likely that the competing motives are still present but are less salient at the time the resolution is formed. A more modest version that I would like to adopt is to say that the resolution is formed when the competing desires are relatively weak. Henden also thinks that resolutions and commitments have the special power to exert a motivational influence on the actor in the face of temptation even in the absence of the desire that motivated the formulation of the commitment.

A motive is something that pulls or pushes us to act in a certain way. Intentions and resolutions as described here are formed specifically for this purpose. More importantly, they are formed to compete with motives that are expected to increase in motivational strength in the future. In anticipation of the difficulty one expects to encounter in the face of strong motives in the future, one forms an intention that is meant to persist over time, precisely because one expects the strength of competing motives to change over time. I agree that resolutions and intentions have motivational power. By virtue of having made the decision to act in a certain way, one has added a causally efficacious motive to the motivational framework. Later we will see how intentions figure into discussions of self-control.

---

34 Henden, “What is Self-Control?” 72.
HOW IS IT SUPPOSED TO WORK?

The MHM aims to explain our reasons for acting in the ways that we do. While the basic HM ultimately sees all actions as determined by preferences we have no control over, the MHM introduces various motivational factors apart from desires, leaving open the possibility that one may act against one's strongest preference. The MHM distinguishes between desires and motives, showing that not all desires act as motives. It also introduces reasons and intentions as motives that can move one to act. Motives exist as part of a motivational system. Motives have an absolute force, while their strength or power is a function of their force and the forces of other various motives occurrent in the motivational system.

WHY DOESN’T IT FAIL IN THE SAME WAY HM FAILS?

The reason that the basic HM fails is because it deprives the agent of autonomy. On the HM, one is subject to one's preferences, over which one does not have control, and whichever of these is strongest determines one's actions. Neither does the actor have control over the strength of these preferences. The actor is reduced to a passive bystander to his desires and to what he does.”35 The MHM is a vast improvement over this model. Not only does the MHM allow for the actor to be the source of some of his or her preferences and act on them to attempt to realize them, it also allows him or her to impact causally the motivational strength of motives within the motivational system. The MHM explains action in such a way that the actor becomes an active participant, and most importantly, can act autonomously. In the next section, I will explain the role autonomy plays in the MHM.

---

CHAPTER 4

MHM AND AUTONOMY

DEFINITION OF AUTONOMY

In this section, I will briefly outline the model of autonomy that I will rely on to explain how an agent can act autonomously in the MHM. In an attempt to rescue the Hobbesian account of the will from charges that it doesn’t leave room for autonomy, some influential thinkers have proposed an endorsement model of autonomy, chief among them being Harry Frankfurt. He advocates the view that the key ingredient needed to supplement Hobbes’ account is the reflective endorsement of desires. He does not deny that desires are what move us to act but rather sees autonomy as being expressed in decisions about whether we identify with the desire in a relevant way. We experience what he calls “first-order desires,” which have different strengths and can conflict with each other. Higher-order desires are characterized as a desire about a lower-order desire. One can have a lower-order desire for x, and the higher-order desire is a desire either to have that lower-order desire or not. When we desire that first-order desires move us to action or not, this is called second-order volition. Higher-order volitions constitute our reflective endorsement. On Frankfurt’s view, free will is cashed out in terms of whether the desire that is actually effective (moves the agent to action) is one the agent reflectively endorses or not. If it is the desire that one wants to be one’s will and it moves one to action, then the will is free. If, on the other hand, the desire that moves one to act is one that the agent does not endorse, then one’s will was
not free and the agent is merely a passive bystander to his desires and to what he does.”

For Frankfurt, endorsement is important because the act of endorsement or refusing to endorse a first-order desire makes it truly his own or instead turns it into an alien force.”

I am largely in agreement with Frankfurt’s model. However, I take it that autonomy is rooted in reflective endorsement itself. As I understand it, autonomy consists in the agent’s being the source of some of his or her preferences and successfully acting upon those preferences in an attempt to realize them. Whether the endorsed desire becomes effective or not does not matter for autonomy. Even in cases where the agent’s will was not free, the agent could still have acted autonomously. My reason for making this move is that in terms of motives, some of which are given and hence out of our control, what matters for autonomy is not which of these is causally efficacious in the end but rather whether the ones we act on are self-generated. An agent expresses his or her autonomy when he or she self-generates motives that causally influence the motivational structure. This is achieved through reflective endorsement. The agent who decides to quit smoking has reflectively endorsed the desire to abstain from cigarettes. In doing so, the agent creates a new motive that interacts with other motives present in the motivational system. This motive is not given but is one that is created by the agent and causally impacts the relative strength of the other motives present. The desire to abstain from cigarettes may impact the motivational strength of the desire to drink coffee, for example, if the agent believes that he or she will experience craving for a cigarette when drinking coffee. For the purposes of this paper, I will adopt the endorsement theory of

---

36 Ibid.

autonomy I have just outlined to show how autonomy cashes out in the hydraulic model I propose. In the next section, I will show how reflective endorsement expresses autonomy on this definition.

**The Nature of Reflective Endorsement**

Reflective endorsement is a kind of desire, namely the desire to have a first-order desire become effective or not. Endorsement, when understood as a desire, makes it a possible motive. I have argued that not all desires are motives. In order for reflective endorsement to be a motive, it, like all other desires, must be paired with the belief that the object of reflective endorsement might possibly be attained. What distinguishes reflective endorsement from other kinds of desires is that it is a desire that issues from one’s self.

Reflective endorsement is not a given desire but rather one that constitutes an expression of self. It has this quality because it comes from an active assessment and evaluation of one’s motivational state. Based on the evaluation of one’s motivational state, the agent chooses which motive the agent wishes to move him or her to action. This motive either, as Frankfurt says, makes the motive in question one’s own or into an alien force. When paired with the appropriate belief, reflective endorsement becomes a self-generated motive. While the agent may not be able to control all of the motives present, the agent can control the choices the agent makes about how the agent wants preferences to move him or her. Whether or not other forces causally influence the agent in such a way that his or her desires become effective determines whether the will was free, but it is in deciding how one wants to govern his or her self and attempting to do so that autonomy consists. To demonstrate, the agent who desires to quit smoking from the previous example may employ self-control skills to the best of the agent’s ability yet still fail. Imagine that the agent is
hypnotized and commanded to smoke or somehow tricked into doing so. The agent was autonomous but his or her will was not free since the agent was subject to forces beyond his or her control.

**Self-Generated Motives**

I have argued that there are two kinds of motives that agents have direct control over: self-generated motives and given desires that once reflectively endorsed become the agent’s own. Reflective endorsement is a desire that when paired with the appropriate belief is a motive that issues from the self in an important sense. When reflective endorsement becomes a motive, two important things happen. First, it gives rise to other motives in the form of reasons and intentions, and second, it interacts with the other concurrent motives, causally influencing them. Motives that spring from the motive of reflective endorsement are self-generated motives. This is because when the reflectively endorsed desire becomes a motive, it contributes to the positive motivational base of other desires. Because one’s motivational state includes all of the motives present in the motivational system, and the strength of motives is a function of the other motives within that system, autonomously acquired self-generated motives exert force within the motivational system and have the power to change one’s motivational state. In this way, agents can exert autonomous control over aspects of the total motivational system. One can imagine an agent with various motives including a desire to lose weight. The desire to lose weight will impact the motivational strength of other related motives. The agent’s desire to exercise may become stronger in light of the desire to lose weight, and other lifestyle-related motives that impact weight loss may be similarly affected. The desire to eat ice cream and pizza may decrease in motivational strength to some degree. Motives are interconnected in such a way that changing one will have a direct impact
on relevantly connected motives. Now that I have described how reflective endorsement provides the basis for autonomy and self-generated motives in the MHM, I will show how it relates to the notion of self-control.

**Self-Control**

In this section, I will look at the roles of self-control and willpower under the MHM. First it will benefit the discussion to clarify how the term “self-control” will be used. In common usage, “self-control” is used as a success term. One can be said to have little, none, or a lot of self-control. This is usually judged on the relative success, or lack thereof, of the individual in question to adhere to his or her stated goals about which he or she needs to exercise self-control. This interpretation could be misleading. I propose we adopt a conception of self-control that includes skills, techniques, strategies, and methods that may be exercised with more or less force and which I will refer to as Self-Control-Skills (SCS). My reason for making this distinction is that we can consider a case where the actor employed many strategies and methods (SCS) to achieve the goal but still failed. In this case, the common conception would say that the agent had no, or at least very little, self-control. On my account, the agent had a lot of self-control (in terms of skills exercised). This matters when we consider that SCS are an exercise of autonomy regardless of whether they achieve their purpose or not. From here on out, when I use the term “self-control,” it should be taken to mean Self-Control-Skills. A better way to understand this on the MHM is to say that the SCS employed produced a lot of motivational strength. Attempts to exercise self-control can be more or less successful; the outcome will rely on the total motivational equation. What this indicates is that self-generated SCS vary in motivational strength and hence their ability to influence the motive that ultimately moves the agent. What is important to note is that
regardless of whether they end in success, SCS are an expression of one’s autonomy and causally influence one’s motivational state.

**Synchronic and Diachronic Self-Control**

According to Stephen P. Garvey, self-control can be broken into 2 categories: synchronic and diachronic. Synchronic self-control is —the exercise of self-control from the moment in time at which, and during the interval of time over which, the strength of the actor’s desire to phi would, but for the exercise of self-control, exceed the strength of any countervailing desires.”38 This kind of self-control is usually exercised when one does not anticipate the situation. One can imagine an agent who is on a diet and finds him or herself in a situation where a favorite food is offered. The only defense the agent has against the sudden desire for the food present is the agent’s synchronic self-control. Diachronic self-control describes —self-control an actor exercises prior to the point in time at which the strength of the desire to phi would, but for the exercise of synchronic self-control, exceed the strength of his countervailing desires.”39 Exercises of diachronic self-control are in anticipation of a situation that the agent believes will make it difficult to adhere to the agent’s reflectively endorsed desires. In the case of the dieter, he or she might consciously choose to go only to restaurants that serve healthy food to avoid temptation. Garvey goes on to say, —The wayward desire is either not yet occurrent (present to the actor’s awareness), or if occurrent, not yet sufficiently strong to overtake countervailing desires.”40 Here I think it matters if the desire is present in the unconsciousness or not because even in the unconscious

---

39 Ibid.
40 Ibid.
state, it can affect the agent’s formulation of intentions and have motivational strength.

Contrary to Garvey, I take it that concurrent desires can be conscious or unconscious since in both cases the desire in question is an active part of the agent’s motivational state. Moreover, unconscious desires can be causally efficacious. This point will be particularly relevant in later discussions regarding the role of mindfulness in relation to autonomy.

According to Garvey, self control can be achieved in two ways. First Garvey describes the skilled or indirect method: “An exercise of skilled self-control consists in an actor’s choice to deploy any of a number of strategies or techniques the aim of which is to decrease the strength of the wayward desire, or alternatively, to increase the strength of the countervailing desires, or both.”41 This description embodies diachronic self-control. The second way he calls “Brute/direct.” He says only that this form of self-control is more difficult to conceptualize but that it is the type that is limited and subject to fatigue.42 I take it that what he is describing here is willpower. While willpower and self-control can be subsumed under the title of self-control, they have different origins. Willpower is the capacity of the will to make itself effective. This capacity is limited. It also lacks the power to change the motivational strength of the desire it is exerting force on. Willpower can support self-control, but because willpower is a limited resource and does not directly influence the strength of motives, it should be considered a last resort. Some forms of diachronic self-control do have the ability to alter the motivational strength of desires, and it is these, I will argue, that will ultimately support the strong will. Another notable difference between SCS and willpower is that skilled self-control entails a choice about how one will

41 Ibid., 4.
42 Ibid., 4-5.
make the reflectively endorsed desire effective, while willpower requires no such decision. Willpower’s aim is always to make the will effective, and there need be no plan for how to employ it.

**Types of Diachronic Self-Control Skills**

Diachronic SCS fall into different categories. George Loewenstein describes several variations such as avoidance, cognitive transformation of rewards or punishments, and pre-commitment devices. The latter provide incentives or disincentives and can function in four different ways. First, they can actually eliminate the possibility of non-compliance. For example, one might lock oneself in a room to avoid the temptation one wants to resist. Second, they can introduce new visceral factors that compete with the one the actor wants to suppress. Loewenstein uses the example of losing $100 if one succumbs to a temptation.43 Third, they can disarm rationalizations. –Thus, one can rationalize that ‘one dessert can’t hurt’, but it is more difficult to rationalize that losing $100 won’t hurt.‖44 Finally, they can mitigate the power of the visceral desire.45 He writes that many visceral factors ‘are stimulated by the availability of reward. If pre-commitment tactics decrease the individual’s subjective likelihood of succumbing, they may initiate a kind of self-fulfilling prophesy by reducing the hunger that served as the initial impetus for the indulgence.‖46 These descriptions demonstrate some concrete methods one might employ to impact causally the motivational power of competing motives. Loewenstein takes it that these types of self-

---

44 Ibid.
45 Ibid.
46 Ibid.
control do not require willpower.\textsuperscript{47} Though the variations will cash out differently in terms of efficacy, "common to all cases of self-control is that they involve an attempt to master actual or expected competing motivations that would otherwise move us."\textsuperscript{48}

In terms of agency and autonomy in the MHM, cognitive transformation can most positively contribute to efforts to improve autonomy. Cognitive transformation involves actions taken on the part of the agent to influence the causal strength of motives. The agent can either strengthen or weaken motives in an effort to make the agent’s reflectively endorsed desires effective. This will be further discussed in a later section.

\textsuperscript{47} I do not agree with Loewenstein on this point. It seems to me that while forming an intention might not require willpower, enacting SCS will. I will not develop this point in the paper, though whether all forms of SCS require willpower is a question worthy of further consideration.

\textsuperscript{48} Henden, "What is Self-Control?" 71.
CHAPTER 5

WILLPOWER AND THE SELF

I have proposed a Motivational Hydraulic model in which our strongest motive is what ultimately guides our action. I assert that these motives all act in a like way, varying in strength. One question that can be asked of such an account is how is self-control possible? How can the ‘self’ control the ‘self’? Self-control and willpower are almost always employed to manage visceral motives that compete with our reflectively endorsed desires. In this section, I will develop a description of a dual-processing model of decision making that helps explain why we experience simultaneously competing motives and how self-control is possible. I will also provide an account for how reflective endorsement can be seen as truly representing the agent’s will. I will finish by arguing that self-control is a means to help integrate the fragmented self.

People with simultaneously conflicting desires often feel divided in two. In the case of addicts who want to quit, they often feel as if they have a separate, ‘dreaded’ self, which arises in certain situations that moves them to continue drug-seeking behavior. Over the years, philosophers, psychologists, and scientists have proposed various theories to explain why addicts continue to seek out the object of addiction in spite of their claims that they would prefer not to. While most of these theories have been criticized as being incomplete,

---

they are not necessarily incompatible. Redish, Jensen, and Johnson have presented a theory that unifies these various explanations from within a decision-making framework.  

**The Nature of the Will**

I take the will to reflect one’s reflectively endorsed desires. The aim of the will is to make these effective. When the agent decides reflectively to endorse or to reject lower-order desires, he or she does so based on his or her values. The will can realize this end through the exertion of willpower, self-control-skills, or a combination of the two. The will is not synonymous with willpower. The will is the reflectively endorsed desire while willpower is the causal force that aims to achieve that end.

While some use the terms “self-control” and “willpower” interchangeably, these terms refer to different phenomena, and although closely related, they ought not to be conflated. “Self-control” should be understood as a set of skills employed by one with the aim of acting in accordance with one’s intentions. In short, it is one’s plan for self-governance. “Willpower” is another force one exerts to achieve the same end and may or may not be paired with self-control skills. Willpower is the force of the will that aims to make its self effective. Willpower and self-control are alike in that they are both motives. Common to all cases of self-control is the active engagement of the agent in autonomously implementing strategies that aim to exert causal force over competing motives with the intention of making the desire the agent reflectively endorses effective. I will next explain the nature of willpower.

---

THE NATURE OF WILLPOWER

In a previous section, I described how self-control and willpower act as motives in the MHM. I will now provide further explication of how willpower works to support SCS. Like SCS, willpower is exercised with the intention of causally influencing competing motives.\textsuperscript{51} Willpower is almost always employed to suppress or override some type of ‘visceral’ motivation.\textsuperscript{52} Roy Baumeister describes it as a ‘limited resource, akin to energy or strength that can become depleted through use.’\textsuperscript{53} He likens it to a muscle that becomes fatigued through exercise. Based on his studies, Baumeister determined that willpower as a resource is related not only to self-control-skills but also appears to be central to the self’s executive function more generally, including those aspects of choice, volition, active instead of passive responding, and taking responsibility.\textsuperscript{54} Willpower is inextricably tied to other motives as well, evidenced by the fact that it is enhanced or diminished by factors such as psychic preparation, internal and external cues, stress, and physiological conditions.\textsuperscript{55} Willpower’s tendency to depletion can be explained precisely because of the way it is interconnected with other motives. Studies such as Baumeister’s serve to demonstrate how the motivational strength of willpower is influenced by the factors mentioned. The way that willpower can be depleted is similar to the way in which gasoline is depleted in a vehicle. There is a limited

\textsuperscript{51} While willpower may be needed to ‘form, retain or execute an intention’ therefore supporting SCS, it is not synonymous with SCS. Henden, ‘What is Self-Control?’ 73. This is evidenced also by the fact that one can exert willpower in the absence of SCS. It is for this reason that I consider willpower a type of motive distinct from self-control.

\textsuperscript{52} By visceral, Loewenstein refers to: (1) drives, such as hunger and sexual desire, (2) emotions such as anger and fear, and (3) somatic sensations, such as pain. Loewenstein, ‘Willpower,’ 52.


\textsuperscript{54} Ibid., 134.

\textsuperscript{55} Loewenstein, ‘Willpower,’ 57.
supply of fuel in the gas tank. It is not only depletable but also influenced by various other factors such as, the speed the car is driving, whether or not the heat or air conditioning is being used, and the weight of the passengers in the car. Likewise, willpower is impacted by other motives that can deplete it at different rates. Although willpower can be influenced by various factors, and is in a sense flexible in terms of limits, its powers are depletable. Therefore, failures to realize the desire one has reflectively endorsed cannot be blamed on failures in willpower alone. They occur because although willpower is a motive that can help one push towards the action one wants to be effective, it is only one among many motives that exist within the wider motivational system.

**THE TWO DECISION-MAKING SYSTEMS**

According to Redish, Jensen, and Johnson’s framework, a decision can come from one of two learning systems that interact with each other. The first system is a flexible, cognitive, planning system,” and the second is a rigid, automatic, habit-based system.”

The planning system takes the situation, predicts, and then evaluates the outcome. The habit system takes the situation and determines the best course of action based on the remembered past. One common element among theories of addiction has to do with predicting the value of various outcomes. It is in part because addicts do not accurately predict the value of the repeated drug use that they continue to seek it out. According to Redish, Jensen, and Johnson, there are two ways that people predict the value of outcomes: forward-search and caching. The forward-search system is slow but flexible. It takes time because consequences for a course of action and the alternatives are considered, but for this same

---

57 Ibid., 417.
reason it is flexible. This type of learning leaves open different courses of action. For example, when one is driving to an unfamiliar location, one will have to plan out the route and based on the available options choose the most efficient route to the destination. By contrast, the habit system, which relies on caching, is fast but inflexible. It simply goes back in the memory and determines the choice to make based on past decisions. The decision is determined by the agent’s history and is restricted to that. This is a useful feature of our learning but can be problematic when we develop habits that we wish to change. In the previous example, the driving route is new, but if we imagine the agent driving a familiar route that has been committed to the habit system, the agent relies on memories to guide him or her to the destination. This can be problematic in this instance if the destination (consequence) changes. If the first part of the route is similar but then a different turn must be taken to get to the new location, unless the agent applies attention, he or she will be guided by the habit system and likely find his or herself driving to the familiar destination almost without realizing it. That is because caching does not identify consequences and so cannot change in response to changes in the value of the relevant consequences.\(^\text{58}\)

Despite the rigidity of the habit system, it is possible to change developed habits. Usually, we employ the planning system when learning a new behavior, –Behavior generally begins with flexible planning systems but, for repeated behaviors, can become driven by the less flexible […] habit system.”\(^\text{59}\) However, Redish, Jensen, and Johnson point out that given the right conditions, the planning system can override even well-developed automated

\(^{58}\) Ibid., 418.

\(^{59}\) Ibid., 419.
responses. When these two systems diverge in term of their outputs (decisions), system 2 (planning system) can exert regulatory control over system 1 (habit system). System 1 is characterized by emotions and urges, while system 2 has the planning/practical reasoning role. On this view, there can be two strongest desires: one dictated by system 1 and the other by system 2. One might have a well-developed habit to finish a meal with a cigarette. This desire is guided by the habit system. The agent might also have resolved to quit smoking. The resolution derives from the planning system. In this case, the agent will experience two simultaneous desires: the desire to smoke after the meal and the desire not to. This helps explain how we can have the experience of being divided or having simultaneous competing motives.

**THE FRAGMENTED SELF**

This view is supported by Neil Levy, who also conceives of the self as being fragmented. He writes that we are fragmented in the sense that our minds are built out of “sub-personal mechanisms” that interact. Some of these are cut off from other ones, or as he puts it, “informationally encapsulated.” This, he says, helps explain how “conscious perception of what we are doing and why can diverge, under some circumstances, from the real causes of our behavior.” His descriptions are consistent with the dual model proposed by Redish, Jensen, and Johnson. A unified self is not a given, it is an achievement.” Levy takes the unified self to be the result of negotiation, bargaining and strong-arm tactics

---

60 Redish, Jensen, and Johnson, do not go into detail as to what the appropriate conditions would be but instead refer to several studies. Ibid., 420.

61 Chandra Sekhar Sripada, “Philosophical Questions about the Nature of Willpower,” *Philosophy Compass* 5, no. 9 (2010), 796.


63 Ibid.
employed by sub-personal mechanisms as they attempt to achieve their aims.\textsuperscript{64} These internal struggles he describes take place between the planning system and the habit system. When there is divergence, i.e., conflicting outputs by the two systems, the planning system can try to override the habit system. It is when these two systems are in agreement that the self is unified. —The multiplicity of mechanisms, each pulling in its own direction, comes gradually to be replaced by a self, with a more or less consistent set of preferences, dispositions and desires—in short, a character.\textsuperscript{65} If an agent only had motives that accorded with the agent's reflectively endorsed desires, then the agent would not experience the conflict that springs from competing motives. This consistency is what Levy calls unification. I will argue that there are two main ways in which this can be achieved. There are (1) devices that change the circumstances so that the habit systems are not triggered, and (2) practices that aim directly to modify how one experiences the desire. This point will be addressed later in the paper.

\textbf{The Authority of the Will}

I have argued that reflectively endorsed desires are an expression of one's self and in that sense can causally influence other motives. Reflective endorsement can lead one to employ self-control-skills if one forms an intention to act on it. But this may not seem sufficient to account for how reflective endorsement is genuinely the agent's own. One major criticism of Frankfurt's view of freedom of the will brought up by Michael L. Corrado is that the higher-order desires, which are supposed to represent the agent's reflective endorsement, cannot be distinguished from other types of desires in a relevant way. —If the substance he (the willing addict) is addicted to creates in him not only a desire for the drug, but also a

\textsuperscript{64} Ibid.
\textsuperscript{65} Ibid., 438.
higher-order desire to be an addict, then the drug makes him a willing addict, but we might still deny that he is responsible for his behavior.”\textsuperscript{66} In other words, there is no way to guarantee that the higher-order desire is not simply generated by the addiction itself. The line of thought goes that if an addict has lower-order desires to continue to use the drug and a higher-order desire to continue to do so, the higher-order desire could be a result of the addiction. If this is the case, then we cannot say that the agent was really free. How can we distinguish between given desires and the agent’s will? Dual processing helps provide an account of how this is possible.

Self-control skills always derive from the planning system. The planning system is slow, consciously controlled, uses linguistic/logical representations, and has access to much larger and more global stores of information.\textsuperscript{67} It is the processing system that has access to the necessary components to make decisions based on all things considered. It is active, unlike the habit system, which is reactive. Reflective endorsement reflects the agent’s evaluation of alternatives and whether to endorse a lower-order desire or not. This endorsement depends on judging whether the lower-order desire is consistent with one’s values or not. If the agent judges that it is, then the desire is endorsed, and if not, it is rejected. When the desire is endorsed, the agent makes it the agent’s own. The will’s function is to make reflectively endorsed desires causally effective. Since the reflectively endorsed desires are formed in accordance with one’s values, and the will is striving to make effective desires that express the agent’s desires, the will has authority. The will does not strive to make desires that are alien to the agent effective. Therefore, Corrado’s criticism of

\textsuperscript{66} Michael L. Corrado, “Addiction and Responsibility,” Law and Philosophy 18, no. 6 (1999), 586.

\textsuperscript{67} Sripada, “Philosophical Questions,” 796.
Frankfurt's model does not apply to the MHM. The agent has some of his or her own preferences, and more importantly, the will strives to make these effective. It is because reflective endorsement originates in the planning system that reflectively endorsed desires can be considered one's own.

**WILLPOWER AND SELF-CONTROL: A MEANS TO SELF-UNIFICATION?**

I take it that self-unification is achieved when one brings one's lower-order desires in line with one's reflectively endorsed desires. I have said that there are two main ways in which unification can be improved. One is a form of SCS. While there are various forms of SCS, most of them do not support self-unification. The problem with techniques such as avoidance, for example, is that they don't really solve the problem. At some point, the addict will most likely, despite efforts to avoid it, encounter situations in which the addict is confronted with the strong urges to consume. In this case, the addict will have to rely on willpower to resist. The self-control technique that may very well best serve the goal of unification is cognitive transformation.

Cognitive transformation aims at altering the motivational power of certain desires and aversions. Loewenstein uses the examples of a dieter who fantasizes about having a thinner figure and the smoker who imagines the horrors of cancer to introduce competing motives.\(^{68}\) One can also impact motivational strength by imagining the object of desire in a negative light or something aversive in a positive light. Given that motives are what move us to act, and cognitive transformation has the power to alter the strength of motives, it is the skill that can most directly contribute to the goal of self-unification. Cognitive transformation

---

\(^{68}\) Loewenstein, "Willpower," 58.
can reduce the motivational strength of lower-order desires one does not wish to become effective or strengthen competing motives to the same end. If we are able to reduce the motivational strength of motives we do not wish to become effective, then we will have less need to rely upon willpower and self-control skills. Let’s imagine an agent who has a desire to smoke cigarettes and a competing desire to quit smoking. To assign an arbitrary value, let’s say the desire to quit is Ms (Motivational strength) 10. The agent can use cognitive transformation to reduce the motivational strength. Instead of seeing smoking in a positive light, as a harmless form of relaxation, for example, the agent might attempt to see smoking as a weakness that reflects poorly on his or her character or as potentially life threatening. In recasting the agent’s perception of the values of smoking, the motivational strength might be reduced from Ms10 to Ms 7. In this case, the agent will not have to rely so strongly on SCS and willpower since the motivational strength of the desire has been reduced. In the next section, I will look at applications of the MHM, beginning with meditation. This will include a discussion of the second way in which one can improve self-unification.
CHAPTER 6
MHM AND MEDITATION

In Buddhist philosophy, the three roots of human suffering are desire, aversion, and ignorance. In the first, we rationally pursue our welfare (desires); in the second, we rationally avoid pain (aversion); and in the last, we are irrational or have false beliefs (ignorance). While these are presented as contradictory views in the West, Buddhism sees the three as interrelated. One method Buddhists prescribe to undermine the three roots of suffering is meditation. Undermining these roots, I will argue, is one way to improve autonomy in addicts.

MEDITATION

There are many varieties of meditation, even within the Buddhist tradition. Meditation is a mental practice or discipline that aims to benefit the practitioner. Meditation can be practiced alone, in a group, with or without the guidance of a teacher and in secular or religious contexts. "[M]editation refers to a family of self-regulation practices that focus on training attention and awareness in order to bring mental processes under greater voluntary control and thereby foster general mental well-being and development and/or specific capacities such as calm, clarity, and concentration."69 The two major categories are mindfulness, or Vipassana meditation, and transcendental meditation. Vipassana is also called insight meditation. During mindfulness meditation, the subject is asked to sit quietly

and focus on a sound, word, or process such as breathing. Thoughts are allowed freely to arise, and as they do, the meditator is gently to turn attention back to the original focus (sometimes called the “anchor”). Thoughts are to be recognized as they arise but not judged or analyzed. Instead, by returning to the focus, one remains in the present experience.\(^70\) Transcendental meditation, on the other hand, requires a more narrow focus. It involves chanting of a mantra in an effort to quiet the mind. This is not to say that it is characterized by an absence of thought but rather on controlling one’s reaction to thought. Both are used to create calmness and concentration.\(^71\) Transcendental meditation practice has also been shown to improve mindfulness in practitioners.\(^72\)

**GOALS OF MEDITATION**

Meditation has been described as mental cultivation as well as a means to developing familiarity.” This familiarity applies to habits, ways of seeing, or ways of being. Through this type of meditation, one also becomes familiar with one’s psychic functioning. Thupten Jinpa writes, “The term *Gom* (meditation) is applied not only to the process of ‘cultivation’ or ‘development of familiarity,’ it is also applied to the resultant states achieved through such processes.”\(^73\) Implied in this comment is that not only does one become familiar with one’s psychic functioning but also that change occurs through the activity of meditation. This is why Jinpa writes, “In this sense, meditation can refer both to the practice of disciplined

---


\(^72\) Tanner et al., “The Effects of a Transcendental Meditation Program on Mindfulness,” *Journal of Clinical Psychology* 65, no. 6 (2009), 574.

\(^73\) Thupten Jinpa, “Is Meditation a Means of Knowing our Mental World?” (working paper, Institute of Tibetan Classics, McGill University, Montreal, Quebec, 2011).
cultivation as well as the cultivated result of such a discipline.” Mindfulness becomes a part of daily activity. Vipassana is used to generate insight through the bare attention applied to one’s mental activities. Jack Engler describes the goals of this type of mental training: “The aim is threefold: to come to know one’s own mental processes; in this way to begin to have the power to shape or control them; and finally to gain freedom from the condition where one’s psychic processes are unknown and uncontrolled.” In “Mindfulness: A Proposed Operational Definition,” the authors provide a two-component model of mindfulness that I will explore in the next section.

**DEFINING MINDFULNESS**

Scott R. Bishop et al. describe mindfulness approaches to enhancing mental health and emotional well being as “a form of mental training to reduce cognitive vulnerability to reactive modes of mind that may otherwise heighten stress and emotional distress or that may otherwise perpetuate psychopathology.” The two-component model of mindfulness that they have developed describes the respective components as “self-regulation of attention” and “adopting a particular orientation towards one’s experiences in the present moment.” I will refer to these as the “attention mode” and “orientation mode” respectively. I will begin with an explication of the attention mode and follow it with the orientation mode in the following sections.

---

74 Ibid.
77 Ibid.
78 Ibid., 232.
Mindfulness entails regulating the focus of attention to bring awareness to the current experience, enabling one to observe and to attend to feelings, sensations, and the changing field of thoughts. By consciously attending to one's breath, one sustains attention on the immediate experience so that "thoughts, feelings and sensations can be detected in the conscious awareness."\textsuperscript{79} By employing mindfulness in this way, one can experience them as they arise, without necessarily engaging in secondary elaborative processing of them. This means that one can experience them without getting caught up in thoughts about what they mean to the self. This aspect of attention also leads to increased availability of information that might normally occur outside of one's conscious awareness, resulting in a wider perspective on one's current experience. This is described as "beginner’s mind\textsuperscript{80}

\textsuperscript{80} Ibid., 233. and is distinguished from our normal mode of information processing in that "Rather than observing experience through the filter of our beliefs, assumptions, expectations, and desires, mindfulness involves a direct observation of various objects as if for the first time.\textsuperscript{81}"

\textsuperscript{79} Ibid.

\textsuperscript{81} Ibid.

Because of the increased ability to monitor and to control mental processes, Bishop \textit{et al.} describe mindfulness as a metacognitive skill.

The second component in this model is that of orientation to experience. This entails one's committing oneself to an attitude of curiosity about the objects of observation during mindfulness practice. The objects of observation include thoughts, feelings, and sensations. By maintaining this curiosity, one is able to achieve a position of acceptance of the current experience. Bishop \textit{et al.} predict that this stance should lead to reductions in the use of

\textsuperscript{79} Ibid.

\textsuperscript{80} Ibid., 233.

\textsuperscript{81} Ibid.
cognitive and behavioral strategies to avoid aspects of experience." They go on to say that this stance should change the way that painful or unpleasant thoughts are experienced within the psychological context. In essence, emotional distress would be experienced as less unpleasant and threatening since the context of acceptance changes their subjective meaning. This investigative approach helps one to see how one thought or experience gives rise to the next, enhancing one’s understanding of the relationships among thoughts, feelings, and actions and the underlying causes of subjective experience and behavior. Developing the skill of mindfulness as it is laid out here results in an increased ability for deautomatization of behaviors, which will be explained in the next section.

**Deautomatization**

The term “deautomatization” comes from a theory that predicts that meditation practice should minimize biasing or habitual conceptual activity through the interruption or undoing of automatic processes. Because mindfulness enables one to reduce secondary elaborative processing of experience, one is increasingly able to experience the world without cognitive bias, filters, or models built up from past experience. Deautomatization is proposed as the antidote to automatization, the process that occurs with the repetition of an action or behavior wherein the intermediate steps of the behavior disappear from conscious awareness. This theory says that by bringing attention back to these actions and experiences, this process can be reversed. Based on this, Wenk-Sormaz hypothesized that

---

82 Ibid.
83 Ibid.
84 Heidi Wenk-Sormaz, “Meditation Can Reduce Habitual Responding,” *Advances* 21, no. 3-4 (2005), 34.
85 Ibid., 33.
86 Ibid., 34.
this mindfulness should enable the practitioner to encounter a situation that he or she would normally experience as stressful with a new perspective or response different from the typical. Instead of being guided to action by the typical response dictated by the habit system, action becomes a choice for the agent. Wenk-Sormaz conducted two studies to test this hypothesis. The first aimed to determine whether meditation practice resulted in non-habitual responses in practitioners, and the second tested whether this ability to respond in non-habitual ways applied only in cases where this type of response was optimal or if it applied equally in all cases of habitual responses. Wenk-Sormaz's studies demonstrated that meditation does lead to a reduction in habitual responding, but only when the non-habitual response is optimal for the present situation. In the following section, I will argue that these results support my view that meditation can improve autonomy in addicted persons.

**Improved Autonomy through Meditation**

Autonomy as I have defined it has two main components: (1) having some of the preferences one wants to have and (2) successfully acting on them. Self-integration consists in the capacity to attempt to bring one's lower-order desires in line with one's reflectively endorsed ones. Self-integration improves autonomy by reducing the number of discrepancies between lower and higher order desires. This mainly pertains to the first component of autonomy. Meditation can contribute to increased autonomy through the development of mindfulness. Mindfulness skills will enhance one's ability to act on those preferences one

---

87 While the studies in question relied on transcendental meditation as opposed to Vipassana, based on the description of the skills given by the authors, I take it that the skills developed parallel those developed for mindfulness in Vipassana.

wishes to have. The primary way it achieves this is through increased activation of the planning system or a decrease in the activation of the habit system. In this section, I will first look at how mindfulness can lead to an increase in self-unification and follow that with an argument for meditation’s capacity to reduce reliance on the habit system.

I have argued that there are two mains ways one can improve self-unification. In a previous section, I discussed how cognitive transformation contributes to unifying the self. The second way is through mindfulness. Mindfulness can be applied when there is conflict between the habit system and the planning system. The self is considered fragmented when one’s lower-order desires and higher-order desires differ. Reflective endorsement derives from the planning system. If the reflectively endorsed desire conflicts with the lower-order desire, and the lower-order desire is controlled by the habit system, then mindfulness provides a way to modify the lower-order desire. Mindfulness has been shown to lead to deautomatization in cases where an alternative course of action is optimal for the situation.\(^\text{89}\)

If one reflectively endorses an action, then it must be the case that one sees this as the optimal outcome. Applying mindfulness to the habit one wishes to override can lead to a reduction in automatization of habits that conflict with one’s higher-order desires. For example, in Wenk-Sormaz’s study, participants who had a habitual response were asked to respond atypically. Those who were in the meditation group were able to provide the most atypical responses.\(^\text{90}\) It follows that if one were to apply this skill to respond atypically to situations where the habitual response is undesirable, one could override the habitual response in favor of a more desirable course of action. Therefore, if the lower-order desire is

\(^{89}\) Ibid., 44.

\(^{90}\) Ibid.
under the guidance of the habit system, deautomatization provides a means to narrow or even to close the gap between conflicting desires. This supports the goal of self-unification. I will now look at how mindfulness can reduce reliance on the habit system.

One way to influence the motivational strength of desires comes from mindfulness. In a state of mindfulness, thoughts and feelings are observed as events in the mind, without overidentifying with them and reacting to them in an automatic habitual pattern of reactivity. This dispassionate state of self-observation is thought to introduce a ‘space’ between one’s perception and response.91 Instead of feeling a need to act on the lower-order desire, one can reflect on it. If the desire is not reflectively endorsed, then the agent can decide to employ self-control-skills, use willpower, or a combination of the two in order to attempt to make the will effective. G. Alan Marlatt has developed a relapse prevention model that sees the way addicts deal with high-stress situations as crucial to their likelihood for relapse.92 On this model, the addict works to develop coping skills to deal with scenarios that would usually incite the addict to resume use of the object of addiction. Coping skills can be understood as a kind of SCS. One coping skill he recommends is using mindfulness to reduce reactivity to cognitive urges to seek out the object of addiction.93 He describes the process and benefits, the addition of mindfulness provides clients with a new way of processing situational cues and monitoring environmental contingencies. Clients are taught to observe pleasant and unpleasant sensations, thoughts or feelings, and they are encouraged to accept them without

judgment." Addicts who encounter a high-risk situation that previously might have led to seek out the object of addiction now have a kind of early warning system. Because they become more aware of situations that might lead to a relapse, they can better prepare for them and deal with them actively, engaging the planning system instead of reactively, or habitually.

Mindfulness also impacts the motivational power of unpleasant emotions: adopting a state of acceptance toward painful or unpleasant thoughts and feelings would be expected to change the psychological context in which those objects are experienced. In essence, emotional distress would be experienced as less unpleasant and threatening since the context of acceptance changes their subjective meaning. This works because repeated exposure to being mindful in high-risk situations without giving in to the temptation to engage in substance use or acting impulsively in the presence of substance-related cues will lead to increased self-efficacy and the counterconditioning of the positive and negative reinforcement previously associated with an addictive substance. So one still experiences desire and aversion but reduces the need to act upon them. An alcoholic who is confronted with the sight of someone enjoying a cold martini (a situational cue that would normally lead to a relapse) will still experience a strong desire to drink but armed with mindfulness the alcoholic can recognize the desire and accept it. Simply because the alcoholic has the desire doesn’t mean he or she must act on it. Instead of changing the content of one’s thoughts, one

---

95 Bishop et al., "Mindfulness," 233.
changes his or her *relationship* to the thoughts.\(^97\) Desires and aversions’ motivational strength may be reduced in this way. This is another example of cognitive transformation.

Cognitive transformation can make a reward seem less desirable or make an alternative course of action more desirable. However, if the habit system is still active, it will not change the course of action in response to the different valuation of outcomes. That is because the habit system does not take changes in consequences into account when deciding on a course of action. Imagine a person who travels to a foreign country. In the person’s home country, she had developed the habit of whistling while working. In the country the person is visiting, whistling is considered highly inappropriate and offensive behavior. Although she is aware of this change in consequences and does not wish to offend those around her, she finds herself whistling without realizing it. So, even if we can change the way we perceive the rejected lower-order desire, if it is guided by the habit system, we will necessarily have to engage the planning system to change the choices we make. On the other hand, if the planning system is already active, then altering the motivational strength of the lower-order desires and competing motives may be all that is necessary to change the outcome of behavior. In the whistling example, if the woman actively has to decide whether to whistle or not, she can choose not to based on her evaluation of the consequences. Therefore, increased activation of the planning system will result in increased control over actions.

Mindfulness can reduce reliance on the habit system because mindfulness involves a direct observation of events as if for the first time, a quality that is often referred to as

The habit system relies on memories of past experience to choose a course of action. If the experience is perceived as a novel one, then the habit system will not be activated. If this is true, then mindfulness should lead to reduced activation of the habit system and a corresponding activation of the planning system. Because the habit system relies on caching, it should be very predictable. An addict who has an established pattern can use mindfulness to disrupt the pattern precisely because of how predictable it is. Imagine the case of a gambling addict I will call “Lisa.” Lisa tends to gamble when she encounters high stress situations, particularly when she has a fight with her husband. The pattern looks something like this: Lisa has an argument with her husband. She storms out of the house and decides to drive around to calm down. She finds herself driving in the neighborhood of her favorite casino. She thinks that a drink in the cool casino is just what she needs, and she could chat with her favorite bartender. Lisa goes to the casino, has a drink, chats with the bartender, and the next thing she knows she is sitting at the blackjack table with a handful of cards. Lisa might even feel surprised, after all, this was not her intention when she left the house, but there she is all the same. If Lisa were able to apply mindfulness in this situation, it might benefit her in various ways. First, having acted out this scenario many times before, she might be able to recognize the pattern and know what the most likely outcome will be. She can perceive the experience as a novel one and choose to do something different this time. In this case, she engages the planning system to overcome the habit system. Alternately, she might end up in the casino and recognize a strong desire to gamble. Applying mindfulness, she accepts the desire but does not feel as compelled to act upon it. Increased awareness of her feelings and situational clues through mindfulness gives her

---

increased power over them. In the next section, I will explore different approaches to understanding the phenomenon of addiction and how they relate to the MHM.
CHAPTER 7

MHM AND ADDICTION

In this chapter, I will explore the various ways of conceptualizing the phenomenon of addiction. I will begin with current definitions of addiction and follow with an overview of the disease view and the willpower view. I will incorporate Bennett Foddy and Julian Savulescu’s “A Liberal Account of Addiction” to show how addictive desires parallel other “natural” desires but argue that the desires in cases of addiction have more motivational strength due to the way we experience them in our motivational landscape. In short, I will attempt to explain how it is that addiction impairs autonomy.

DEFINING ADDICTION

Definitions of substance dependence have traditionally come from the field of medicine. In order to provide a more substantial exposition of the nature of addiction, in this chapter I will look at addiction from three perspectives. Because addiction can be conceptualized in different ways by different disciplines, there is a danger of their talking past each other. It will be beneficial therefore to provide some guidelines for what is being referred to when we use the term. As Lennert Nordenfelt points out, “if these theories [of addiction] are flawed and if the authors propose a theory that is more correct, then for such a proposal to be meaningful all the theories must concern the same group of people.”99 To this end, the first section will explore the definitions of and features of addiction. Section two will

provide an overview of addiction theories from the field of psychology as well as how each of these explain or propose to deal with the various features. The final section will look at the main philosophical theories of addiction.

Terms such as “addiction” and “dependence” are often used interchangeably but in fact refer to different, though sometimes overlapping, phenomena. In the *Diagnostic Statistical Manual of Mental Disorders* (DSM), “substance abuse” is defined thusly: “The essential feature of Substance Dependence is a cluster of cognitive, behavioral, and physiologic symptoms indicating that the individual continues use of the substance despite significant substance-related problems.” The DSM unfortunately does not provide a separate definition of “addiction.” Since there is no consensus on how to define “addiction,” I will try to give some basic guidelines, relying on the work of Howard J. Schaffer.

While some might restrict the concept of addiction to include only addictions to substances that cause physical dependence, I think it is prudent to allow for a broad enough interpretation to include “process” addictions that do not involve the use of psychoactive substances. My reason for taking this approach is that physical dependence is not a necessary or sufficient condition for addiction. For example, upon stopping, pathological gamblers who do not use alcohol or other psychoactive drugs often show physical symptoms that appear to be very similar to either narcotics, stimulant, or poly-substance withdrawal.”

---


101 Howard J. Schaffer is an associate professor and the Director of the Division on Addictions at Harvard Medical School. In addition to research and private practice activities, he consults internationally to a variety of organizations in business, education, human services, and government. His major research interests include the social perception of addiction and disease, the philosophy of science, impulse control, as well as compulsive behaviors, adolescent gambling, and the natural history of addictive behaviors.

Because the objects of addiction are pursued by individuals to different degrees and with varying consequences, the concept of addiction can be understood as existing at the far end of a spectrum. This creates difficulty in distinguishing cases of addiction. Schaffer offers some broad guidelines that I think will serve the purposes of this paper. There are three main aspects he thinks can help one identify possible cases of addiction: the person in question (1) demonstrates —Behavior that is motivated by emotions ranging along the Craving to Compulsion spectrum; craving can range from a mild desire to an overwhelming impulse to act. A compulsive behavior is a powerful repeating pattern of action,” (2) —Continued involvement with the drug or activity in spite of adverse social, psychological, or biological consequences,” and (3) —Loss of Control, that is, a subjective sense that one no longer can control one's behavior.”

In discussing the features of addiction, I will not consider the feature of physical withdrawal for two reasons. The first reason for this omission is that while this can be an important feature to consider at a certain stage of addiction, for the purposes of this paper, I would like to restrict the discussion to the features that persist even after withdrawal. Secondly, I would prefer to focus on features of addiction that include addiction not only to substances that cause physical dependence but also to other addictions that do not have this feature. It is for the same reasons that I will only give a brief treatment of the etiology of addiction. In the present discussion, I will not focus on why or how someone becomes addicted as much as what happens once someone is. Therefore, while I will recognize the importance of these key features, they will not be given a thorough treatment here.

\[103\] Ibid.
The Disease View

The disease view states that drug-seeking behavior in addiction is perpetuated by changes in the brain resulting from chronic drug use. On this view, the addict is no longer rational, and drug use is fundamentally due to non-voluntary processes. It is called the "disease view" because it conceptualizes substance addiction as a chronic physical illness. This model also assumes that drug abusers are predisposed, perhaps genetically, to seek out and use substances to excess once coming into contact with them. The key features of addiction under the disease model are that the disease is incurable, and as such, requires total abstinence. Moreover, it is believed that if left untreated, the disease will progress and, ultimately, lead to death. This is the model that is used by popular 12-step programs such as Alcoholics Anonymous and Narcotics Anonymous. It is considered a medical condition that requires formal treatment.

Part of the argument for addiction as a brain disease is that practically all drugs of abuse have common effects, either directly or indirectly, on a single pathway in the brain. Activation of this system is seen to be the common element that drives continued drug use. This is the main thrust of the arguments put forth by Foddy and Savulescu in "A Liberal Account of Addiction." They argue that the disease view of addiction is wrong and that

---

104 Foddy and Savulescu, "A Liberal Account," 1.
109 Foddy and Savulescu, "A Liberal Account."
although drugs do work in this way on the brain, they act no differently than desires for other substances or activities that similarly activate the reward system in the brain. Current neurological evidence gives us no reason to think that addictive desires are formed in a different way to regular desires. [...] In fact, any pleasurable experience causes dopamine to be released within the brain, activating these ‘reward’ pathways.”

Whatever caused the reward pathways to be activated strengthens the corresponding neural pathways associated with the activity and makes it more likely that the person will repeat the behavior. Not only that, but it has been shown that reward centers in the addicted brain change in such a way as to make the drug more rewarding through use. Some researchers have posited “supraphysiological” neural connections that act as a unique motivational mechanism that only exists in the brains of addicted individuals. While some argue that it is as if drugs have hijacked the brain’s motivational control circuits, resulting in drug use becoming the sole, or at least the top, motivational priority for the individual,” others argue that these patterns are not due to the particular chemistry of the drugs but rather by repeated presence of any type of brain reward.

While there is significant evidence to support the disease view, it has some serious shortcomings. First of all, as Foddy and Savulescu point out, it does not show how the effects in the brain caused by activation of the reward pathways differ relevantly from substances and activities that activate the pathways in a like way. They provide several examples of non-drug substances, such as saccharin, that provide even greater activation of

110 Ibid., 3.
111 Leshner, “Addiction Is a Brain Disease,” 1.
112 Foddy and Savulescu, “A Liberal Account,” 4.
reward pathways in the brains of addicted rats than does cocaine.\textsuperscript{113} I am inclined to agree with their assertion that drugs activate the reward pathways in the brain in an analogous way to non-drug reward pathway activating substances and activities. This understanding also allows for addiction to include behavioral addictions, such as gambling, sex, and food, which have been shown to exhibit the same type of brain activation and dopamine release as drugs.\textsuperscript{114}

Another objection to this view is that it fails to account for what utilitarian writers in economics refer to as —temporal preferences.\textsuperscript{115} These writers espouse the view that addicts are simply trying to maximize their satisfactions. These theories of —rational” addiction say that your options compete based on a motivational quality—such as reward—and that the addict chooses the option with the biggest reward.\textsuperscript{116} I will return to this point later. The competing theory of addiction is the willpower view, which will be explained in the following section.

**The Willpower View**

In this section, I will examine the competing sides of the debate over whether or not addiction can be classified as a defect of the will. I will first look at what the theory entails, including the views of some proponents of the view and follow that with an overview of those who deny defects of the will are possible.

\textsuperscript{113} Ibid.

\textsuperscript{114} Ibid.

\textsuperscript{115} Ainslie, "Research-Based Theory," 81.

\textsuperscript{116} Ibid.
**Defect of Will?**

First of all, it is necessary to clarify what it would mean to say that addicts experience a defect of will versus a defect of willpower. I will argue that neither of these are appropriate and, moreover, that neither are possible. I conceptualize the will as a reflection of the reflectively endorsed desire that the agent wants to become effective. On this conception, what would it mean to say that one has a defect in will? Some would say that it would mean that we are unable to make our will effective.

This doesn’t seem correct, because even in the instance that one does not succeed in making his or her will effective, the will may still have the same, though unrealized, goal. This seems to be the case with an addict whose will is to abstain, yet in spite of that, seeks out the object of addiction again. This is not a defect of the will; this is an inability to comply with the will. That inability doesn’t have anything to do with the will itself but rather indicates a failure in making the will effective. So, perhaps in this case, is it a failure in willpower?

**Defect of Willpower?**

The willpower view of addiction is one that has a long history among philosophers and psychologists alike. This view holds that while there may be a desire on the part of that addict to abstain from the object of addiction, his or her will is not strong enough to overcome his or her desire for it. The addict is seen to lack control over the addiction, and the lack of control is attributed to weakness of the will. A failure in willpower would mean that one did not have enough or did not employ willpower in the appropriate way to ensure compliance with one’s will.
Since willpower is by definition a depletable resource, it seems illogical to say that an insufficient amount is a defect. Willpower is depleted in relation to the relative strength of that with which it competes. If willpower is a depletable resource, then every person’s willpower will be defective when faced with a motive that overpowers it. If this is the universal nature of willpower, then we cannot describe this feature as a defect.

**Defect of Desire?**

Based on the previous discussion, it seems, then, that the problem lies not in the will or in willpower itself but rather in the relative strength of the desire that competes with them. Should we then say that the defect lies in the desire? To say that the defect lies in the desire would be to attribute an absolute value to the desire. That is to say that the motivational strength or power of the object of desire would impact all agents who come into contact with it in a like way. If that were the case, then in the case of alcohol consumption, for example, all people who came into contact with it would become alcoholics. This is simply not the case. My desire for chocolate is experienced significantly differently from my best friend and even differs over time. As human beings, we are subject to a wide range of desires that we experience to lesser or stronger degrees. In cases where one experiences unusually strong desires that may out-muscle willpower, it seems that instead of describing this as a type of defect, it might be more appropriate to describe it in terms of a phenomenological disorder. A disorder indicates an abnormality in physical or mental functioning. In this sense, the way that the addict phenomenologically experiences desires for the object of desire is abnormal in that it is experienced as much stronger than the average person. While addictive substances may by nature cause these types of abnormally strong desires, it is the case that two people may ingest the same substance and one will experience it as significantly stronger than the
other. So it would seem we cannot chalk it up exclusively to the effect of the drug itself; what matters is how the person in question experiences it.

**Denying Defects of the Will**

On the side of those who deny addiction as a defect of the will, there are three main approaches. Here I will rely on Michael L. Corrado’s summary of the views provided in *Addiction and Responsibility: An Introduction*. The first objection comes from those who would deny that addictive behavior is no different than any other sort of behavior. Supporters of this view say that the addict could stop if he or she wanted to. Included in this camp are economists, who say that addicts are rationally pursuing their own welfare. Corrado calls this approach “rational addiction.” Next, there are those who agree that addictive behavior is rational in the context of a coercive situation. Coercion may come in the form of withdrawal, cravings, or strong desires. Corrado calls this view “addiction as duress.” Finally, there are those who conceive of addiction as essentially irrational. On this view, it is not the will that is defective but, rather, one’s rationality. They say that the addict possesses false beliefs that cause him or her to act in ways that bring about consequences that he or she did not want and/or fail to bring about the desired consequences. The addict’s beliefs do not respond to evidence appropriately. Corrado calls this the “addiction as distortion” view.

**STRENGTH AND WEAKNESS OF WILL**

In order to explain how it is that addiction might impair autonomy, I will begin with a discussion of the notions of strength and weakness of will. Henden references Holton’s

---

118 Ibid., 584.
119 Ibid., 584-5.
theory of strength of will. On Holton’s view, strength of will is the opposite of weakness of will, which consists in an over-readiness to abandon one’s resolutions. A strong will, on the other hand, means a capacity to persist in one’s resolutions by refusal to reconsider them. Holton says that this stick-to-itness requires mental effort. The faculty that enables one to make the mental effort is willpower. Willpower is employed actively by the agent when there is a perceived struggle, characterized as a resistance to one’s own inclinations.\textsuperscript{120} So, willpower supports the strong will but is a limited resource. –Agents exercise will-power,‖ he claims, –to avoid focusing on and developing thoughts that might lead to a reconsideration of a resolution.‖\textsuperscript{121} Henden goes on to criticize the distinction that willpower is an ability to make a mental effort to maintain one’s prior resolution by blocking reconsideration (after all, sometimes it is prudent to reconsider) and comes up with an improved version of the theory that instead of separating self-control and will-power, it conceptualizes will-power as an essential component of synchronic self-control. –The self-controlled agent uses will-power to direct her attention away from a rebellious desire in order to form, retain or execute an intention to perform the act she takes it she has the most reason to perform. Since the exercise of willpower to achieve self-control implies strength of will, it follows that strength of will is an essential component of self-control.‖\textsuperscript{122} Therefore, Henden says, akratic behavior, as he refers to it, may originate in the agent’s voluntary decisions and intentions in an autonomous way. One decides to employ self-control, decides not to employ it, or makes no decision either way.

\textsuperscript{120} Henden, –What is Self-Control?‖ 82-3.
\textsuperscript{121} Ibid., 83.
\textsuperscript{122} Ibid., 84.
According to Holton, to form an intention is to decide on a course of action. *Akrasia* on this view is not the same as weakness of will since one may form the intention to do something against one’s better judgment: “*central cases of weakness of the will are best characterized not as cases in which people act against their better judgment, but as cases in which they fail to act on their intentions.*”¹²³ Weakness of the will arises […] when agents are too ready to reconsider their intentions.”¹²⁴ He contrasts this with the traditional account of *akrasia*, which refers to voluntary actions taken against one’s better judgments.¹²⁵

There are times, Holton says, when it is reasonable to revise one’s intentions. It is reasonable to revise intentions —if one believes that circumstances have changed in such a way that they defeat the purpose of having the intention […], if one believes that they can no longer be carried out,” or —if one believes that they will lead one to great suffering when that suffering was not envisioned at the time of forming the intention.”¹²⁶ It is reasonable not to reconsider intentions —in circumstances that prevent clear thought, if those intentions were made in circumstances that allow clear thought,” or if they —were expressly made in order to get over one’s later reluctance to act.”¹²⁷ This last point is especially important because Holton proposes that one reason for forming intentions may be temporal: —it could be that now is a more propitious moment to reason. Tomorrow, there will be some

---

¹²⁴ Ibid.
¹²⁵ Ibid., 243.
¹²⁶ Ibid., 249.
¹²⁷ Ibid.
factor that will stop us from reasoning well.” So, on his view, intentions serve to protect us from future selves faced with temptation.

**MOTIVES CHANGE OVER TIME**

In a previous section, I outlined how intentions act as motives. While some have argued that there are times when a well-formed intention will possess the requisite strength to win out over competing motives, often times it will not. Weakness of will has been described as an over willingness to abandon or to overturn one’s intentions. Neil Levy goes one step further in saying that agency consists essentially in one’s capacity for extended agency. He says that addiction undermines this capacity and impedes addicts from integrating their lives.

Unwilling addicts’ failure to comply with their own reflectively endorsed desires and intentions can be explained in part by the temporal aspect of the motivational strength of motives. The problem for most addicts is not quitting but staying quit. Some addicts will willingly abstain from drugs for periods of time, even going through physical withdrawal, in order to lower their tolerance to the substance, thereby getting more effect from less of the substance. The likelihood of relapse is estimated at 40-60% in drug addicts who have undergone supposedly successful treatment. The problem, it seems, is a temporal one. The unwilling addict is frequently able to quit but all too often will revert to the drug use over

---

128 Ibid., 244.
130 Ainslie, Research-Based Theory,” 82.
time. In the next section, I will explore some possible explanations that may account for this phenomenon.

What accounts for what Levy describes as the “oscillation of preferences” in an addict? George Ainslie, among others, cites an innate tendency in people to form temporary preferences. Utility theorists such as Ainslie assume that people generally seek to maximize their utility. They also discount their utility over time, much in the same ways that banks do. Ainslie provides the example of a car worth $10,000 today being discounted at a rate of 10% per year. That car would have been worth $9,000 to him a year ago, $8,100 two years ago, and so on. Utility theorists say that people have the same tendency to discount about internal rewards. This type of discounting is considered “rational.” Until recently, it was assumed that all people discounted this way, but addicts do not show the same consistency. They alternate between preferences, and preference is correlated with the proximity of the perceived reward. As opposed to rational exponential discounting, addicts exhibit tendencies to discount hyperbolically. In essence, this means that they devalue future rewards proportionate to their delay so that rewards at very short and very long delays would be valued more than those who discount exponentially. What this means is that the person will choose a smaller reward at a closer time over a larger reward at a further time. This view is supported by Herrnstein’s matching law, which states that the proportion of rewards chosen matches their relative immediacies (as well as their amounts and frequencies

---

133 Ainslie, “Research-Based Theory,” 86.
134 Ibid.
135 Ibid., 87-88.
136 Ibid., 88.
of occurrence).”\textsuperscript{137} For addicts, even if they prefer abstinence over continued use at one time, as the perceived rewards associated with drug use become more immediate, they will change their preference to drug use. According to Richard Holton, this leads to a “judgment shift” in the addict.\textsuperscript{138}

**ADDICTIVE DESIRES**

I have argued that addictive desires are formed in a similar way to other desires and act as motives in a like way as other desires. There is, however, a way in which addictive desires can differ importantly from other desires that directly impacts autonomy. The problem for addicts is not just that they have difficulty translating intentions into action but also that they have trouble doing it consistently over time in the face of the competing motives. When the competing motive is less immediate, the strongest motive might be to abstain, but the closer the reward becomes, its power as a motive increases and can succeed in overpowering one’s intentions. This problem can be partly accounted for by the activation of the habit system over the planning system. Since the habit system deals primarily with urges and impulses, as the reward becomes more immediate, it may activate this system and becomes a stronger motive in spite of the intention in the planning system. If the addictive desires are being guided by the habit system, then the agent’s ability to struggle against those desires is directly compromised. It is this tendency for the planning system to be overcome by the habit system that amounts to an impairment of autonomy in the addict. If reflective


\textsuperscript{138} Holton, “Intention,” 241-62.
endorsement and intentions derive from the planning system and the competing motivation lies in the habit system, then activation of the habit system impairs one of the capacities to struggle against the competing motivation. Only when the planning system is activated can it override the competing motive in the habit system.

In this chapter, I have explored numerous ways to conceptualize the phenomenon of addiction. I have argued that autonomy impairment occurs when the addictive urges generated in the habit system lead an agent to act upon them. Since the habit system guides one to action in an automatic way, it impairs the agent's ability to choose to struggle against it. Extended agency depends not just upon intentions but also on the planning system, which gives us the ability to choose our actions in accordance with those intentions. In the following section, I will look at how the unwilling addict might fail to make his or her reflectively endorsed desires effective.

**THE UNWILLING ADDICT**

If we accept Frankfurt's definition of the will, anytime agents fail to make their reflectively endorsed desires effective, the will is not free. What I am concerned with here is not whether the agent's will is free or not but, rather, whether the agent acts autonomously. I will argue that in the case of the unwilling addict, his or her failure to conform to the reflectively endorsed desire may either be autonomous or may constitute an instance of impaired autonomy. In spite of the failure on the part of the agent, there are scenarios in which although the agent's will was not free, it was still autonomous.

The unwilling addict, according to Frankfurt, is one who has a lower-order desire to pursue the object of addiction and a higher-order desire not to have the lower desire and a higher-order volition that the higher-order desire be effective. Since autonomy consists in (1)
having some of the preferences one wishes to have, and (2) acting on them, impairment in either of these areas will constitute impaired autonomy.

The unwilling addict may fail to make the reflectively endorsed desire not to pursue the object of addiction effective in several ways. Because the addict has the reflectively endorsed desire to abstain from the object of addiction, he or she fulfills the first requirement for autonomy; having some of the preferences one wants to have. What is at stake, then, is the second requirement; acting upon that preference. I will look at several scenarios, all of which assume the reflectively endorsed desire to abstain. (1) He or she may fail to form the intention. This could be a result of a lack of the appropriate belief regarding the reflectively endorsed desire. For example, the addict does not believe it is within his or her power to abstain. In this case, although the addict has the preference he or she wants, the desire is not a motive. Since there is no motive, there is no reason to act on the reflectively endorsed desire. In this case, the addict’s capacity to act is frustrated by a lack of appropriate belief necessary to act. This instance constitutes an impairment of autonomy. (2) He or she may form the intention but fail to create an action plan for how to carry out the intention. The addict may decide to rely on willpower alone. Since willpower is a depletable resource, if willpower fails and seeking the object of addiction is the strongest motive at the time willpower fails, the addict will be moved to act on that motive. In this case, the addict did not act on the reflectively endorsed desire. If it is the case that the addict could not act on it as in the previous case—due to a false belief (perhaps that willpower would be sufficient)—then his or her autonomy was impaired. (3) He or she may form the intention and have an action plan but still fail. There are two possibilities to consider here. It might be the case that the addict employs SCS but does not create the requisite motivational power to overcome the
competing lower-order desire. The addict is then forced to rely on willpower, which when depleted leads him or her to succumb to the lower-order desire. The addict has the wanted preference and acts on it but fails due to the relative strength of the competing motive. In these cases, the addict acted autonomously. This is an example of one’s acting autonomously but not freely. The second possibility is that the addict fails to form the action plan well. François Schroeter takes it that addictive desires interfere with one’s capacity to develop action plans. —The crucial point is that compulsive action tendencies [brute action tendencies such as those generated by drug-related stimuli] place abnormal constraints on the addict’s capacities to select action plans and thus prevent these capacities from functioning properly.”¹³⁹ While it is debatable whether or not compulsive action tendencies have this effect or not, I take it that if this is the case, then this also constitutes an impairment of autonomy. Any circumstance where the agent’s capacity to form action plans is compromised will likewise constitute an impairment of autonomy.

To summarize, failures to make reflectively endorsed desires effective can happen in 2 ways. Agents can fail autonomously, such as when they fail to comply with their autonomously formed plans and intentions through a failure in willpower. They can also fail due, at least in part, to impaired autonomy. This may occur when the addict fails to form intentions or action plans or fails to do so well through an impaired capacity to do so.

The MHM I have outlined in this thesis provides a possible foundation for a model of action that is based on motives as the causal vectors of force that move one to act. In this model, one expresses autonomy when one has some of the desires one wants to have and acts upon them in such a way as to attempt to make them effective. I have attempted to show that

one might accomplish this by impacting motives in such a way that they contribute to the positive motivational base of the desire that moves one to act. In the section, I will explore the role of addictive desires and the implications for reflective agency.

**IMPULSES AND THE IMPAIRMENT OF REFLECTIVE AGENCY**

Schroeter is not alone in thinking that addictive desires diminish the addict’s capacities to act in certain ways. Wallace calls these types of desires A-Impulses or A-Desires. According to him, desires can influence the processes in which agency consists in a negative way by inhibiting the orderly progression from deliberative reflection through to corresponding action.¹⁴⁰ This can happen because not all of our desires respond to practical reasoning about what we should do. A-Desires can present obstacles to the processes in which deliberative agency consists.¹⁴¹ Wallace says this is what happens with A-Desires. They are resilient in light of the fact that they do not respond to deliberative reflection. He goes on to discuss the relation between desires and pleasure and pain. He thinks the right approach to take here is to think of desires in terms of perceptions rather than sensations. He says that desires to obtain something pleasurable or avoid something painful can be thought of as quasi-perceptual modes of presentations of these anticipated situations.¹⁴² They resemble perceptions in that they have a conceptual structure but are not necessarily judgments or rest upon them. So, one can have an irrational desire to obtain some pleasure in spite of the fact that one, after reflecting, judges that obtaining the thing in question would probably not provide the pleasure one seeks. This describes a case in which the desire consists in part in

¹⁴¹ Ibid.
¹⁴² Ibid., 641.
one’s thinking about the potential course of action or experience in terms of some evaluative category, its persistently presenting itself to one as, say, pleasant.”\textsuperscript{143} Wallace goes on to say that the way we categorize these is not necessarily fully under one’s voluntary control. This explains why we continue to desire them in spite of our judgment that they will probably not be that pleasant. Another distinctive feature of these kinds of desires is that one’s attention is especially focused on the possibility of the action’s leading to pleasure. This hyper-focus also helps explain resilience of these types of desires. Wallace goes on to say that these desires have a particular force or intensity not because of their causal force but rather because of the way the person in their grip experiences them. He calls this the “Phenomenological conception” of strength of desire.\textsuperscript{144}

Wallace supposes that A-desires might interfere with one’s capacity to deliberate rationally about the proposed course of action, though, he says, it would in no case completely remove one’s ability to appreciate basic moral, legal, and factual considerations. Impairment, says Wallace, would likely be in terms of one’s capacity to weight those normative factors accurately and judiciously, in reflection leading to a verdict about one has most reason to do on the whole.”\textsuperscript{145} This is not to say that addicts are completely impaired but that making the right decisions is very difficult and requires — effort, concentration, strength of will, and so on” and as such is best understood as being at least a partial impairment of our capacities for reflective agency.\textsuperscript{146}

\textsuperscript{143} Ibid., 642.
\textsuperscript{144} Ibid., 643.
\textsuperscript{145} Ibid., 645.
\textsuperscript{146} Ibid., 646-7.
According to Wallace, the description thus far does not constitute a defect of the will, which by contrast would be a form of interference with the processes of reflective agency that goes beyond, and is independent from, deficiencies in respect to rationality alone.\textsuperscript{147} A-desires can amount to volitional defects of the will in light of the fact that even when the person in their grip reasons well about what he or she ought to do, A-desires persist and make it very difficult to choose to comply with the rational verdict one arrived at.\textsuperscript{148} So, Wallace also characterizes A-desires as impairments of our powers of choice. \textquoteleft\textquoteleft By making one vividly aware of alternatives for action that promise an immediate and visceral pleasure, A-impulses bring it about that compliance with one\textsc{'}s better judgment would require effort, concentration, strength of Will.\textquoteright\textquoteright \textsuperscript{149}

Because intentions and action plans rely on the decision-making processes, any impairment in these processes will result in reduced autonomy. The impairments described by Wallace here reflect those described by Redish, Jensen, and Johnson.\textsuperscript{150}

\begin{flushleft}
\textsuperscript{147} Ibid., 647.
\textsuperscript{148} Ibid., 648.
\textsuperscript{149} Ibid., 649.
\textsuperscript{150} Redish, Jensen, and Johnson, \textquoteleft\textquoteleft Unified Framework,\textquoteright\textquoteright 436-9.
\end{flushleft}
CHAPTER 8

CONCLUSION

This thesis set out to answer several questions related to autonomy and addiction. In Part 1, I explored the hydraulic model of action. Based on my research, I concluded that in light of its failure to account for an agent’s will, the HM fails to account accurately for how agents choose to act. Instead, I proposed the motivational hydraulic model. This model interprets human action as issuing from various motives, with the strongest motive prevailing. On this model, desires are possible motives but must be paired with the correct beliefs in order to become motives. Apart from motives, I have shown that other factors such as reasons, intentions, self-control, and willpower also act as motives. Most importantly, I have demonstrated that agents can control some of these motives. While the HM deprives the agent of the capacity to struggle against desires the agent does not wish to have, the MHM preserves autonomy even in the face of some given motives. We do not, as the HM suggests, do what we most want to do but rather what we are most motivated to do.

In the first section, I also discussed the notion of willpower and distinguished between willpower and self-control. I concluded that willpower can be used to support self-control, but since it is a depletable resource, it should be considered a last resort in terms of trying to control one’s actions. Based on research on addiction, I endorsed a dual-system model of decision-making. In this system, choices are made in either the planning system or the habit system. While the planning system involves an active choice based on a variety of factors, the habit system is largely automatic. The planning system has been shown to have the capacity to override the habit system under certain circumstances. Relying on this model,
I argued that the will can have authority because it aims to make our reflectively endorsed desires, which derive from the planning system, effective.

In Part 2, I explored a possible application of the MHM. I offered working definitions for relevant notions such as meditation, mindfulness, and addiction. I allowed for a broad interpretation of addiction leaving open the possibility for addictions that are not substance-related. Based on research on mindfulness and addiction, I concluded that in some cases, addiction can impair autonomy. I argued that mindfulness is a skill that can contribute to improved autonomy in addicts. Mindfulness can improve autonomy by enhancing activation of the planning system and decreasing the role of the habit system. When addictive desires are controlled by the habit system, one can apply mindfulness to override these desires and have increased access to alternative choices.

The MHM I propose provides a more comprehensive way to conceptualize autonomy. It contributes to the theory of action by offering a way to explain human action that still allows for agency and autonomy and accurately reflects research from interrelated fields. This thesis also contributes to an improved way of understanding the phenomenon of addiction. There are currently no widely accepted or agreed upon standards by which to measure addiction, which creates problems in diagnosing and treating it. It also makes it difficult to answer questions about how addicts ought to be treated by the law. The extent to which we hold an addict (or any agent) responsible for his or her actions should correspond to how much control he or she had over those actions and if the ability to realize the goals those actions aimed at were in his or her control or not. While it is far from a definition that could be adopted for these purposes, it does shed light on some important aspects of the phenomenon that can help us move in that direction. The most important contribution this
thesis makes is in offering possible methods for improving autonomy. If we accept that addiction impairs autonomy, then developing methods to improve addict’s ability to control his or her choices and actions may improve relapse rates and help mitigate this widespread problem.

While I chose to focus on addiction in this thesis, I believe that the methods I put forth here can be effectively applied to various situations where an agent’s autonomy might be compromised. In order to identify other possible applications for improving autonomy, I see two major areas that would benefit from further research. First, we need to identify instances of impaired autonomy. I have described autonomy as a spectrum with varying degrees. Diagnostic tests that aim to measure the severity of autonomy impairment would be one way to identify other possible candidates for these kinds of treatments. The other area of research should aim to determine methods for developing mindfulness. Meditation is the method I refer to here, but it is not the only one. Even in terms of meditation, it is unclear how much or how often one must practice in order to develop these skills to the point where they will be useful in terms of improving autonomy. Hopefully, continued research and interdisciplinary work in these areas can resolve the various unanswered questions surrounding why we act the way we do and how we can better control our actions.
REFERENCES


