

# Florida State University Libraries

---

Electronic Theses, Treatises and Dissertations

The Graduate School

---

2014

## Are We Free?: Psychology's Challenges to Free Will

Zachary T. Martin



FLORIDA STATE UNIVERSITY  
COLLEGE OF ARTS AND SCIENCES

ARE WE FREE?  
PSYCHOLOGY'S CHALLENGES TO FREE WILL

By  
ZACHARY T. MARTIN

A Dissertation submitted to the  
Department of Philosophy  
in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

Degree Awarded:  
Summer Semester, 2014

Zachary T. Martin defended this dissertation on July 18<sup>th</sup>, 2014.

The members of the supervisory committee were:

Randolph Clarke  
Professor Directing Dissertation

Michael Kaschak  
University Representative

Alfred R. Mele  
Committee Member

Michael Bishop  
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the thesis has been approved in accordance with university requirements.

I dedicate this to myself.

## ACKNOWLEDGEMENTS

It's probably impossible for me to acknowledge everyone that deserves to be thanked for their help on support on this project, but I'll do my best to recognize those that I can remember.

Above all, I would like to thank Randolph Clarke for all his help in guiding me through the writing of this dissertation and looking at numerous drafts of each chapter. I would also like to thank the rest of my dissertation committee: Al Mele, Michael Bishop, and Michael Kaschak. I have learned quite a bit from each of these individuals, both in and outside of seminars. Other faculty members at FSU that I am honored to have had the privilege of learning from include Stephen Kearns, David McNaughton, Russell Dancy, Piers Rawling, John Roberts, Michael Ruse, Jack Justus, and Michael McKenna.

Though they did not help directly with my dissertation, I am greatly indebted to many professors from my undergraduate career at Bowling Green State University without whom I might never have gone to graduate school. In particular, I would like to thank Janice Dowell, Don Callen, and Michael Bradie for writing me letters of recommendation and being wonderful mentors. I would also like to recognize Sean Foran, Stephen Wall, Marvin Belzer, and the late Raymond Fry.

I have also had the privilege of discussing many of the nascent ideas in this dissertation with graduate students at FSU. In particular, I would like to thank Aron Vadakin, Clifford Sosis, Steve McFarlane, Joshua Shepherd, Heather Perez, Megan McGrew, Rich Cordero, Dan Miller, Kyle Fritz, Michael Robinson, and Robyn Waller. I have learned quite a bit from each of these individuals and was lucky to have them as colleagues.

For their love and support, I would like to thank my parents, Tom and Mary Martin, and my brother, Alec, and sister, Kiley. And last of all, but certainly not least, I would like to thank my soon-to-be wife, Christine Weissglass, for the myriad of ways my life is better because of her. Through countless discussions (both philosophical and non-philosophical), I am grateful for everything she has helped me to see more clearly. Without her encouragement, support, wisdom, love, and patience none of this would be possible.

# TABLE OF CONTENTS

LIST OF TABLES .....	vii
ABSTRACT .....	viii
1. INTRODUCTION .....	1
1.1 Free Will and Psychology? .....	1
1.2 Preview of Things to Come .....	2
2. FREE WILL AND AWARENESS .....	5
2.1 Awareness of What? .....	5
2.2 Types of Awareness .....	12
2.3 Awareness Requirements .....	15
2.4 Situationism and Its Implications .....	20
2.5 Conclusion .....	22
3. FREE WILL AND CAPACITIES .....	23
3.1 Condition (MR) .....	23
3.2 Evaluating the Evidence .....	30
3.3 Extending the Argument .....	36
4. FREE WILL AND SELF-CONTROL .....	38
4.1 Philosophy and Self-Control .....	38
4.2 Psychology and Self-Control .....	41
4.3 Evaluating the Evidence .....	48
4.4 Concluding Remarks .....	55
5. FREE WILL AND AUTOMATICITY .....	57
5.1 What is Automaticity? .....	57
5.2 Automaticity and Lack of Control .....	59
5.3 The Threat of Global Automaticity .....	67
5.4 Conclusion .....	73
6. FREE WILL AND THE ALLEGED ILLUSION OF CONSCIOUS WILL .....	74
6.1 Conceptual Investment in Consciousness .....	74

6.2 What is the Illusion of Conscious Will?.....	76
6.3 Wegner on Consciousness, the Will, and the Self.....	85
6.4 The Apparent Mental Causation Argument .....	87
6.5 Conclusion.....	95
7. CONCLUSION .....	96
APPENDIX: COPYRIGHT PERMISSION FOR FIGURE 6.1.....	98
REFERENCES .....	99
BIOGRAPHICAL SKETCH .....	109

## **LIST OF TABLES**

2.1	Objects of Awareness .....	16
2.2	Types of Awareness .....	16



## ABSTRACT

Do we really have as much control over our behavior as we think we do? Might we be mistaken that our actions are ultimately up to us? Some philosophers argue that recent scientific research shows that we don't have the control required to act freely (even when we think we do) and, therefore, we should not be held responsible for what we do. This dissertation examines the scientific research that has been marshaled in support of such claims, which I argue does not threaten our belief in our freedom and the control we have over our actions.

This work proceeds by looking at four broadly construed areas of research in psychology: situationism, self-control, automaticity, and conscious control. I identify ways in which each research program has been used in an attempt to demonstrate that we do not have the control required to act freely. One research program that has been of interest to the free will community in philosophy and psychology is situationism. The results of situationism, which include demonstrating the behavioral effects of seemingly insubstantial situational features on our thoughts and actions, have been cited as a threat to our freedom of action on two fronts. The first is that situationism indicates a widespread absence of awareness of the reasons for which we act. The second has it that situationism actually disturbs some of our capacities to respond and react to reasons that are crucial for acting freely. I make a contribution to the debate involving the first sort of threat by developing a taxonomy of types of awareness along two dimensions: objects of awareness in acting freely and ways of being aware in acting freely. This allows me to evaluate what situationism demonstrates we are unaware of compared to what is required for free action. I defend the claim that the lack of awareness involved in situationism doesn't appear to threaten our prospects of acting freely.

I make a contribution to the debate involving the second sort of threat by defending the claim that situationism hasn't demonstrated that the general or specific capacities required for acting freely have been hampered to such an extent that we are rendered unfree even if most participants in the situationist experiments fail to exercise such capacities.

A second research program that has claimed to indicate a threat to our acting freely involves the study of self-control. Psychologists have demonstrated that our capacity for self-control relies on a limited resource the depletion of which impairs subsequent attempts to regulate behavior. I carefully examine this research and argue that while it may indicate an impairment in the degree to which we control our behavior, it does not show that our free will is

diminished to such an extent that we are rendered unfree in a sense of lacking moral responsibility.

I then look at research on automaticity that has been used to argue that we are not free and morally responsible. Such research maintains that automatic processes can perform tasks once thought to be unique to controlled conscious processes. This has led some psychologists and philosophers to conclude that there is little explanatory room left for conscious control and no need to posit free will. I take a critical look at such experiments and argue that they do not show that consciousness plays no role in acting and that we nonetheless have a significant amount of control over our behavior.

Lastly, I look at Daniel Wegner's research on conscious control which is used to support the idea that our conscious will is an illusion. I argue not only that his claims rest on a mistaken notion of what is required for free will, but that the scientific experiments do not warrant the conclusion that our conscious will is entirely inefficacious.

# CHAPTER ONE

## INTRODUCTION

Most of us tend to think that we have free will. Most of us think that, most of the time, when we act we act freely. We, of course, acknowledge that there are many things that shape who we are and the actions we perform, but overall we think that the things we do we do freely and can be held responsible for them. Until recently, free will has, primarily, been the subject of investigation by philosophers. But over the past couple of decades, psychologists have begun to see free will as a topic to which they can contribute. Like philosophers, some psychologists think that we have free will and some think that we do not. However, psychologists who have claimed to have demonstrated through experimentation that we do not have free will have generally garnered the most attention. Given that most of us have a strong belief in our own free will, challenges to our freedom appear troublesome not only for our own sense of agency, but for many of our social practices and institutions. This dissertation evaluates the veracity of the claims presented by psychologists that we are not free or that we have diminished free will. I hope to determine whether there is in fact evidence that we lack the control over our lives to be free and morally responsible.

In the rest of this introductory chapter, I provide more background about the free will debate and why, in particular, I am looking at challenges from psychology. I will also give a brief summary of the rest of the chapters in this dissertation.

### 1.1 Free Will and Psychology?

The philosophic debate concerning free action and moral responsibility has traditionally been focused on a conceptual issue, *viz.*, whether or not the control required for free action and moral responsibility is compatible with determinism—determinism being the thesis that “there is at any instant exactly one physically possible future (van Inwagen, 1983, p. 3). Philosophers have mainly debated the compatibility question and compatibilists and libertarians have focused their attention on developing a set of sufficient conditions the satisfaction of which allows for

free action and moral responsibility. There is, however, another issue concerning free action which has not received as much attention, *viz.*, to what extent actual human agents can satisfy the control required for free will. It is this latter question that is the focus of this dissertation. But notice that this issue is an empirical one. Determining to what extent actual human agents are free requires some hard, physical evidence.

Luckily, some psychologists have turned their attention to investigating free will and related issues. However, many psychologists doing this sort of research claim to have empirical evidence demonstrating that we lack free will. Benjamin Libet has claimed to have shown that our unconscious makes decisions before we are even aware of them. Daniel Wegner purports to have shown that free will is nothing but an illusion. And John Bargh argues that all human behavior is automatic, obviating the need to posit anything like free will. Many of these and related ideas have also been packaged and sold as popular psychology books, like Eliezer Sternberg's *My Brain Made Me Do It* (2010), Trick Slattery's *Breaking the Free Will Illusion for the Betterment of Humankind* (2014), and Sam Harris's *Free Will* (2012). Given the prevalence of such declarations of our lack of freedom, investigating whether the empirical evidence appealed to actually supports these claims seems like a reasonable place to start in order to determine whether we have free will.

## **1.2 Preview of Things to Come**

In the next five chapters I look at four different research programs in psychology that have purported to show that we lack or have a diminished amount of free will. The first two chapters deal with an area of research that has come to be known as situationism. Experiments in this domain have been taken to show that seemingly irrelevant factors of one's situation or environment have a profound impact on peoples' behavior unbeknownst to them. I'll report many of these experiments later on, but, for now, consider the following experiment a preview of things to come. Experimenters have found that hospital patients whose room had a view of a natural setting tended to recover from surgery much quicker than patients whose room had a view of a brick wall (Ulrich 1984). This experiment seems to show that something as seemingly irrelevant as the view from one's window can have a profound impact on the rate at which one's

body recovers from surgery. Although this example isn't one that demonstrates how situational features impact peoples' behavior, I think it illustrates how subtle features of the environment or situation can drastically affect an individual.

As interesting as this and other studies like it are, they have been employed by some in order to make the case that we have a diminished amount of free will or less free will than we think we do. One way in which situationism has been used in this regard is by arguing that we lack the awareness required to act freely. In chapter 2 I evaluate whether this is actually so by identifying several plausible candidate awareness requirements on acting freely, and then determining whether situationism shows that we do not satisfy these requirements. I argue that it's unclear whether the most plausible awareness requirements for acting freely are shown not to be satisfied by agents in virtue of situationism primarily because the conditions at issue involve agents' abilities. Whether situationism shows that we lack the required abilities or capacities for acting freely is taken up in chapter 3.

I begin chapter 3 by looking at an argument put forth by Eddy Nahmias. Nahmias distills two conditions he claims are shared by most free will theorists and then argues that situationism shows that we often do not satisfy those conditions. I argue, however, that Nahmias's argument moves too quickly and lacks precision. Once his argument has been recast to rid it of ambiguity, I argue that situationism does not clearly indicate that we lack the relevant abilities to act freely. In particular, I think that situationism does not indicate that we lack the required abilities at issue as opposed to not exercising the required abilities at issue.

Chapter 4 moves away from situationism and looks instead at Roy F. Baumeister's research program involving self-control. Baumeister's work shows that the capacity for self-control relies on resources that are limited. Once our self-control resources have been depleted, we are less likely to succeed in exercising self-control. Baumeister's influential work on self-control has been used by him and others to argue that free will is only an occasional phenomenon or a "sometimes thing" (2008a). In this chapter, I closely examine the self-control experiments conducted by Baumeister and attempt to determine whether they support his claims about free will. I argue that they do not, primarily because the capacity for self-control is only part of what constitutes our capacity to act freely. However, I do think that this research program gives us valuable information about why we behave the way we do and how we can exercise more control over our behavior.

Chapter 5 of this dissertation looks at the literature on automaticity. Such research in psychology tends to show that there are processes which operate outside the scope of our conscious awareness. John Bargh, among others, has used the results of this research program to argue that we have very little (if any) free will. In order to determine the veracity of these claims I attempt to identify what automaticity is and how it might contribute to a lack of control of the sort that would diminish our free will. Even though I identify a kind of automatic process that I think could potentially diminish our freedom, I argue that most automatic processes aren't of this sort, and automatic processes in general do not diminish our freedom just in virtue of being automatic processes. I also consider two arguments that make use of the pervasive nature of automatic processes to argue that we do not have free will. I find these arguments wanting and not supported by the empirical evidence.

The last main chapter of this dissertation considers Daniel Wegner's research on the illusion of conscious will. Wegner and many of his colleagues have conducted research which they take to indicate that the conscious will is actually an illusion, and the control we think we consciously exercise is merely apparent. I closely examine many of Wegner's claims and find that not only do many of them lack conceptual clarity, but also that Wegner's theories are far from supported by the empirical evidence he presents.

I conclude this dissertation by briefly discussing some of my main conclusions and what I think this dissertation shows. Briefly put, much of the empirical evidence that has been used to support the claim that we lack or have diminished free will has been overblown. Though many of the research programs I have looked at provide us with interesting information about the control we exercise over our own behavior, they do not tend to support the negative claims made about free will.

## CHAPTER TWO

### FREE WILL AND AWARENESS

What does it take to act freely? Philosophers focusing on free will have attempted to articulate the various conditions under which an agent can be appropriately said to act freely. Among these conditions is the requirement that an agent possess some degree of awareness or self-knowledge.<sup>1</sup> But recent experimental data—particularly, from situationism—seems to illustrate that we have a significant lack of self-knowledge or awareness, and it has been suggested that this could be troubling for our prospects of acting freely.<sup>2</sup> The aim of this chapter is to determine whether the experimental data do in fact show that we lack the self-knowledge or awareness required for free action.

In order to do this, we need to determine what kind of awareness is required for free action, and, secondly, what we are required to be aware of in order to act freely. In the first two sections of this chapter, I will attempt to develop candidate conditions for these two aspects of an awareness requirement. In the third section I will combine these two aspects to generate a list of possible awareness requirements for acting freely and attempt to determine how plausible they are. In the last section I will explain situationism and evaluate whether any of the plausible requirements are shown by situationism to often not be satisfied. If it can be determined that any plausible requirements on acting freely are often not satisfied in virtue of situationism, this might be evidence that we aren't as free as we think we are.

#### 2.1 Awareness of What?

First, a terminological note is in order. Self-knowledge is a term that is used to refer to a variety of things. It is taken to be a type of knowledge which has special characteristics that differentiate it from other types of knowledge an agent may have.<sup>3</sup> Self-knowledge can also be used to refer to knowledge of what the self is, what kind of persisting thing the agent happens to

---

<sup>1</sup> For example, Double (1991), Ekstrom (2002), Fischer and Ravizza (2000), Frankfurt (1971), Mele (1995), Nahmias (2007), and O'Connor (2002).

<sup>2</sup> Nelkin (2005) and Nahmias (forthcoming; 2007).

<sup>3</sup> For example, Moran (2001) suggests the characteristics of “immediacy” and “authority.”

be.<sup>4</sup> Neither of these types of self-knowledge will concern me here. The type of self-knowledge I am interested in has to do with knowledge the agent has of her mental states and surroundings. In this sense, awareness can be used synonymously with self-knowledge, and this is often how the terms are used in the literature. Fischer and Ravizza (2000) write that it would be inappropriate to judge a person morally responsible for a kitten's death if he was "unaware that he was driving over the little kitten" (p. 12). "This type of ignorance," they claim, "is merely one way in which a person can lack the particular knowledge requisite for being responsible" (p. 12). It is natural to say that when an agent is aware of some feature of herself, she can be said to have knowledge of that feature of herself. An agent who is aware of her volatile temperament can be said to have some self-knowledge of her temperament. From here on out, I will be using the terms self-knowledge and awareness interchangeably in the senses I have indicated.

Many things an agent can be said to be aware of or know might be pertinent to her agency and pertinent to her free agency. Aristotle in his *Nicomachean Ethics* discusses several things an agent must have knowledge of in order to act voluntarily.<sup>5</sup> Aristotle writes that the agent must have:

[Knowledge of] who is doing it; what he is doing; about what or to what he is doing it; sometimes also what he is doing it with—with what instrument, for example; for what result, for example, safety; in what way, for example, gently or hard (1111a3-5).

Roderick Chisholm expresses a similar requirement on acting freely when he writes, "If I reach for the staff and pick it up...[a]nd if it is something that I do, then there is a very clear sense in which it may be said to be something that I know that I do" (2001, 131). Additionally, in discussing why the results of some social psychology research programs such as situationism seem threatening to our freedom, Dana Nelkin (2005) acknowledges that one explanation might be that they seem to show a lack of general knowledge including features of one's environment. But what appears to be more important is the knowledge or awareness an agent has of her own mental states including beliefs, desires, intentions, and motivations.

---

<sup>4</sup> See Neisser (1988) who argues that self-knowledge is based on several disparate forms of information and that each one establishes a different "self."

<sup>5</sup> *Nicomachean Ethics* (1985) (trans. Terence Irwin).



Eddy Nahmias (2007) argues that our freedom or autonomy<sup>6</sup> is threatened by a body of psychological data indicating that we lack the relevant degree of self-knowledge. “Autonomy,” he writes, “depends on the agent’s ability to know her principles and to know how to act on them,” meaning that an autonomous agent “can articulate to herself and others her principles for action” and can “recognize whether she is acting on reasons she would accept were she to consider them” (2007, 3). Similarly, Dana Nelkin (2005) discusses the importance of this type of self-knowledge to our freedom and responsibility. Nelkin’s discussion involves a different strategy. Her goal is to attempt to uncover why it is that we find situationism threatening to our freedom and responsibility, and she surveys a number of different routes to such a threat. One consideration she investigates is that the situationist literature suggests we lack, to a certain extent, some degree of self-knowledge. She writes, “For one might think that we need to know about our motivations in general in order to learn how to make them effective and develop habits that make them effective” (2005, 197). Even though, as she writes, this posits a contingent relationship between self-knowledge and, perhaps, freedom and responsibility, it could still explain why the situationist literature appears threatening.

To further motivate the idea that some type of awareness of these mental items is required for acting freely, consider an agent who has absolutely no knowledge or awareness of his reasons for action, the intentions he forms, or of what is motivating him. Even if such an agent can perform actions, it doesn’t seem plausible that such an agent can perform actions freely. Timothy O’Connor, for example, claims that he is “unable to conceive an agent’s [freely] controlling his own activity without any awareness of what is motivating him” (2002, 88). On the other hand, if we consider an ideally free agent, an agent with an ideal amount of freedom and self-control, it seems that such an agent would have a great deal of awareness or self-knowledge about their mental states including reasons, intentions, and motivations. Richard Double, for example, writes that ideally exercising free will requires that the agent “knows the nature of [his or her] beliefs, desires and other mental states that bring about [his or her] choice” (1991, 48). So it seems clear that the minimal amount of self-knowledge required for acting freely must fall somewhere between these two extremes.

As we have seen, one obvious suggestion for why awareness of mental items is important is that awareness or self-knowledge of certain mental items allows us to exercise some control

---

<sup>6</sup> Nahmias uses these terms synonymously.

over our behavior. It provides us with an opportunity to ensure that we are acting for the reasons we want to be acting on and not for some other reasons we wouldn't want to be moved by. An agent wanting to quit smoking, for instance, doesn't want to be moved by his desire to have a cigarette. Awareness of this desire and its influence on his behavior provides him with an opportunity to attempt to resist smoking a cigarette, perhaps, by chewing nicotine gum instead. Acting freely can typically involve deliberating about what to do based on the desires we have, weighing reasons for various alternatives open to us, and deciding on a course of action, then acting. An agent may act unfreely when she fails to do this. One who is hypnotized and then given the instruction to cluck like a chicken when the telephone rings presumably doesn't deliberate about clucking like a chicken once she hears the phone; the suggestion has been given to her and has bypassed her deliberative process. She may not even know why she clucked like a chicken after the fact and become embarrassed. There is a sense in which it wasn't up to her to act as she did and so she hasn't acted freely. It seems that her lack of awareness has prevented her from guiding her behavior in the way that she would otherwise want to. But what was she unaware of? The answer seems to be the reason for which she clucked like a chicken.

Reasons for action have developed into a topic of philosophical interest in their own right. Here, I will be using reasons for action in line with the Davidsonian tradition to refer to combinations of beliefs and desires or intentions, the types of things that can non-deviantly cause the agent to act. However, an agent can have reasons for which she doesn't act. On the current suggestion, an agent must be aware of the reasons for which she performed a certain action; that is, she must be aware of the beliefs and desires or the intention for which she is performing the action at issue. But matters can quickly become complicated. Often times, there is not just one reason for which a person performs an action. An action may be performed and caused by a multitude of reasons. A young man may decide to apply to medical school because he wants to become a doctor and believes that entering into and completing medical school is necessary to fulfill his desire. But he may also apply to medical school because he wants to make his parents proud and believes that this would be a good way of ensuring their pride. Or he may want to make a great deal of money in the future and believe that becoming a doctor is a good way to do this. If the young man is motivated by all of these reasons to apply to medical school, need he be aware of all of them or only some of them? I have already mentioned that it would be too

extreme a standard to require that agents be aware of all of their reasons for action, but perhaps it will be useful to state this requirement in order to juxtapose it with others.

R (all): In order to act freely, an agent must be aware of all of the reasons for which she acts.

Another suggestion would be that the agent only needs to be aware of some reason or other which rationalizes his action. So perhaps a different requirement along these lines can be developed which allows for more ignorance with respect to one's reasons for action but not so much so that the action is rendered unintentional.

R (rationalizing): In order to act freely an agent must be aware of at least one reason for which she acts that rationalizes her action.

While a requirement of this sort would allow for more ignorance of the reasons for which one acts, one may suspect that it may not capture what is required for free action; free actions are often seen as requiring more than what is required for intentional actions and, so, one may think that a greater type of awareness of one's reasons is required.

A third possibility for which reasons an agent must be aware of in order to act freely would be the reasons (if there are any) for which an agent acts that she does not identify with. For example, Frankfurt writes:

It is in securing conformity of his will to his second-order volitions, then, that a person exercises freedom of the will. And it is in the discrepancy between his will and his second-order volitions, or in his awareness that their coincidence is not his own doing but only a happy chance, that a person who does not have this freedom feels its lack (1971, 15).

Watson (1975) also writes that “[t]he possibility of unfree action consists in the fact that an agent's valuational system and motivational system may not completely coincide” (1975, 215). These accounts of acting freely indicate that what is important about one's own awareness of her

reasons for acting and motivations has to do with ensuring that we in some way approve of what drives us to act in certain ways. Absence of this sort of awareness could be problematic for acting freely. This condition could be expressed as follows:

R (unidentified): In order to act freely, an agent must be aware of reasons motivating her to act in a way that she does not identify with (if there are any such reasons).

A few parts of this condition need further explanation. The condition is phrased in terms of the reasons that motivate one to act rather than the reasons for which one acts. This is because being aware of reasons for which one acts that the agent doesn't identify with would not help secure more control for the agent over her behavior since she has already acted. It would seem to be important to be aware of these sorts of reasons before one actually acts on them so that she may have the opportunity to refrain from acting on such reasons and, instead, act in a way she wants to act. Also, the notion of what it means to identify with one's motives or desires has been the subject of scrutiny.<sup>7</sup> It's unclear what identifying in this sense amounts to. For my purposes, I'll just give a rough characterization of what I take identifying here to mean. Identification with one's reasons or desires in this sense seems to involve something like accepting being motivated by such a reason or desire because it is in accordance with the majority of one's set of values that one takes to be overriding.

So far I have identified a few types of reasons for action that it might be important to be aware of in order to act freely. But maybe it's not our lack of awareness of our reasons for action that is threatening to our freedom, maybe it's something else. Consider the following study. Nisbett and Wilson (1977) present results from several studies arguing that people generally lack knowledge or awareness of their reasons for action. Bertram Malle (2006), on the other hand, has criticized Nisbett and Wilson's conclusions. In one of Nisbett and Wilson's studies, they presented passersby with a series of stockings and asked them to choose which one they thought was the best quality. Unbeknownst to the participants, all of the articles of clothing were exactly the same. However, people systematically responded that they thought the last item in the row of articles of clothing was of the best quality. Nisbett and Wilson hypothesized that it

---

<sup>7</sup> Frankfurt's (1971) paper stirred up much debate about the issue of identification. For some responses, see Bratman (1996), Moran (2002), and Velleman (2002).

was the position of the articles which influenced people's judgments about which had the best quality, in part, because people never mentioned the order of the clothing items having an impact on their judgment. This study has generally been taken to show that people don't know why they judge or act the way they do. But Malle argues that the participants in the study did offer reasons for why they chose a particular article of clothing over the others and insofar as they did this, they have offered a reasons explanation. What the experimenters were expecting was instead a causal history of reason explanation. Causal history of reason explanations "cite factors that lay in the background of those reasons and brought them about (e.g., context, culture, personality, habits, unconscious mental states)" (2006, 212). With this distinction in mind, Malle concludes that the experiments of the sort reported by Nisbett and Wilson cannot be used to show that people do not generally know their reasons for action. While it may be true that people are generally unaware of certain features that influence the reasons they eventually act on, lack of awareness of these features doesn't imply that people do not know why they judge and act the way they do.

Whether or not Malle's argument is correct, it does seem as though Nahmias and (to some extent) Nelkin make use of the fact that people are unaware of the causal history of their reasons as something that could potentially be threatening to freedom. So, in addition to being aware of one's reasons for action, one may also require being aware of one's causal history of their reasons for acting.

R (causal history factors): In order to act freely, one must be aware of the causal history factors that bring about one's reasons for acting.

This may seem like an extremely strong requirement—and in fact, much stronger than requiring one to be aware of all their reasons for action—so perhaps a modification could make this requirement less demanding. Given that the reason awareness seems to be required for acting freely is because it provides us with an opportunity to ensure that we are acting in ways that are in line with other ways we want to conduct our behavior and goals we have, we could simply require that agents be aware of causal history factors that influence their reasons for acting which frustrate other attitudes they would rather act in accordance with.

R (unidentified causal history factors): In order to act freely, an agent must be aware of causal history factors that bring about one's reasons for acting (if there are any such reasons) which motivate her to act in a way that she does not identify with.

The proposals I have enumerated thus far are not exhaustive and have only tackled one part of the awareness problem *viz.*, what it is one must be aware of. The other part of the problem that deals with the way in which one must be aware of these features will most likely influence the plausibility of the requirements discussed so far. I now turn to addressing the other part of the awareness problem: in what way must one be aware of the mental items that affect their behavior.

## 2.2 Types of Awareness

The type of knowledge the agent has at the time of acting could be understood in a variety of ways. In what follows, I will start by proposing some of the more demanding ways this could be understood and work towards proposing some of the less demanding ways an agent's mode of awareness could be understood. In developing the following types of awareness that may be required, I will simply state what one must be aware of as their "reasons for action." Following the discussion of the various types of awareness an agent may have, I will then consider the plausibility of the possibilities generated by the various types of awareness that may be required and the various things one must be aware of. The first, and most demanding, way an agent could be required to be aware or have knowledge of their own mental items is being directly occurrently aware of what reasons and motives they are acting on when they act. This can be characterized as follows:

OA: In order to act freely, an agent must be directly occurrently aware of the reasons and motives for which she is acting.

On the face of it, this requirement seems implausible for a number of reasons. First of all, an agent could form a standing intention to act on some principle or other so that they do not have to waste precious cognitive resources considering various reasons about whether and how

they should act. For example, one could have the standing intention to call for help when they see that someone else is in danger because they have a general concern for the well-being of others. Rather than wasting time deliberating about their reasons for acting in such situations, the standing intention acts as a sort of pre-packaged plan for action the agent need not consider every time they act on it. But if the standing intention for acting in certain situations was formed freely, it's plausible that acting on such an intention was also a free action.<sup>8</sup>

Second, some actions occur much too quickly for the agent to be directly occurrently aware of their reasons and motives for the performance of the action. Even if one doesn't have a standing intention or principle to act on a certain set of reasons, an agent may nonetheless decide in an emergency situation that help should be called for and this is what they shall do. Execution of such an action may require finding the first available telephone in order to call emergency services. Having concerns such as these occurrently in mind, the agent calling for help may not have the reasons for which he is requesting help occurrently in mind. Nonetheless, the agent's action being free does not seem to be ruled out by this fact.

Third, an action performed by an agent may require a good amount of the agent's conscious attention to be performed such that there isn't enough conscious attention to dedicate to awareness of their reasons for acting. A surgeon performing a risky operation to remove a brain tumor would have to dedicate a large amount of their conscious attention to the precise placement of instruments on living tissue rather than allocating attention to the reasons for which they are performing the action of removing the tumor.

Fourth, and last, it's important to point out that when we act our attention seems to be much more focused on the world in which we are interacting rather than on our inner mental world.<sup>9</sup> In the example above, it's unlikely that the surgeon is paying attention to his mental states and reasons he has for the actions he performs rather than the actions themselves. After the surgeon has settled on a particular course of action to remove the tumor it doesn't make much sense to keep in the forefront of his mind his reasons for doing so; his attention would be better directed at how he is performing his particular actions so that he may better guide his behavior or alter his plan if he deems it necessary.

---

<sup>8</sup> Peter Gollwitzer (1999) describes a kind of standing intention he calls implementation intentions—intentions that link anticipated situations to goal directed action (p. 493).

<sup>9</sup> Elisabeth Pacherie (2006) makes this point. She writes, "Typically, when we are acting, our attentional focus [is] on the outside world we are acting on rather than on the acting self" (p. 158).

Another possible type of awareness is dispositional awareness of at least some of the agent's reasons for acting.<sup>10</sup> To be dispositionally aware of some X, one must have the disposition to become occurrently aware of X. For example, to be dispositionally aware that one's anger at a friend for forgetting their birthday is motivating them to retract an invitation to their party, there must be some triggering or stimulus conditions the occurrence of which would, at least some of the time, trigger the agent to become occurrently aware of their anger. But this kind of awareness seems to be too broad in scope to capture the type of awareness required for free action. The agent considered above who has been hypnotized to cluck like a chicken upon the telephone ringing may not be directly aware of her reasons for acting when she clucks like a chicken, but she could satisfy this sort of dispositional awareness by having the disposition to recall the hypnotic suggestion under certain circumstances. Perhaps, in addition to being hypnotized to forget the suggestion that's been implanted in her, the hypnotist also implanted in her the suggestion to recall her particular reasons for clucking like a chicken upon hearing an actual chicken cluck. In this case, she would be dispositionally aware of her reasons for action, but she would become occurrently aware of them only under circumstances she is unaware of and unlikely to encounter. Nonetheless, her action of clucking like a chicken upon hearing the phone ring seems to remain unfree.

The problem seems to result from the fact that the stimulus condition triggering her disposition to be occurrently aware isn't well specified. Another suggestion is an agent's reasons for action being reportably available to them such that they would be able to report to themselves or others their reasons for action. This proposal gains further support by considering the fact that Nahmias claims, "At a minimum, [the agent] can recognize whether she is acting on reasons she would accept were she to consider them" (2007, 3). This kind of reportable awareness can be characterized as follows:

RA: In order to act freely, an agent must be able to recognize and report (to herself or others) her reasons for action.

A similar requirement is discussed by Neil Levy (2011). Levy considers an awareness requirement on responsible action he calls personal availability. Levy characterizes this type of

---

<sup>10</sup> Neil Levy (2011) considers this requirement as a condition on moral responsibility.



awareness as follows: An agent's reasons for action are personally available when "the information that guides [the agent's] behavior...is available for easy and relatively effortless recall (no special prompting or cues are needed, and...it is constantly available to him" (2011, 246). The agent at issue whose reasons for action are personally available to him would be able to respond when asked why he is doing what he's doing, accurately and without hesitation. This would be characterized the following way:

PA: In order to act freely, an agent must be able to effortlessly and without hesitation recognize and report (to herself or others) why she is doing what she is doing.

While these two requirements are similar, they differ in the amount of effort one needs to exert in order to bring her reasons for action occurrently to mind. Reportable awareness leaves it open that agents may need to exercise a good deal of reflection in order to accurately report the reasons for which they are acting. On the other hand, Levy's personally available requirement has it built in that an agent's awareness of her reasons for action must be able to be reported without effortful reflection; she must be able to accurately express why she is acting as she is without trying a great deal.

### **2.3 Awareness Requirements**

Thus far I have developed several proposals for two aspects of awareness required for free action: what the agent must be aware of and what type of awareness the agent must have. These proposals can be combined in various ways to generate several alternative requirements on free action. The next section of this chapter will consider whether any of these various alternatives are plausible requirements for acting freely. Then I will look more closely at the situationist literature and attempt to identify what type of awareness is shown to be lacking by the oft cited experiments. If some of the more plausibly required kinds of awareness for acting freely are shown by the situationist experiments to be lacking, then this might be grounds for claiming that the situationist literature demonstrates that we aren't free or that we aren't as free as we think.

The proposals for what an agent must be aware of are listed in table 2.1.

**Table 2.1:** Objects of Awareness

R (all)	In order to act freely, an agent must be aware of all of the reasons for which she acts that causally contribute to her action
R (rationalize)	In order to act freely, an agent must be aware of at least one reason for which she acts that rationalizes her action.
R (unidentified)	In order to act freely, an agent must be aware of reasons motivating her to act in a way that she does not identify with (if there are any such reasons).
R (causal history factors)	In order to act freely, one must be aware of the causal history factors that bring about one's reasons for acting.
R (unidentified causal history factors)	In order to act freely, an agent must be aware of causal history factors that bring about one's reasons for acting which motivate her to act in a way that she does not identify with (if there are any such reasons).

And the proposals for the types of awareness required for free action are listed in table 2.2.

**Table 2.2:** Types of Awareness

OA	In order to act freely, an agent must be directly occurrently aware of the reasons and motives for which she is acting.
RA	In order to act freely, an agent must be able to recognize and report (to herself or others) her reasons for action.
PA	In order to act freely, an agent must be able to recognize and report (to herself or others) effortlessly and without hesitation her reasons for action.

Combining these two categories creates a list of potential awareness requirements on acting freely. However, we need not consider all of the fifteen requirements that can be generated from these two lists; some of them can be ruled out. For example, I have already mentioned that it is an implausible requirement that an agent be directly occurrently aware of any of their reasons for action. As mentioned earlier, it has been noted by some that when acting agents often focus their attention outward on the world rather than inward on their own mental states. Similarly, it seems unrealistic to require that an agent have awareness (of any type) of all of the reasons that causally contribute to her action. There are so many factors that contribute to an agent's action that it would be unrealistic to expect her to be able to find out about them let alone be aware of them in any sense. And if this is a good rationale for rejecting R (all) then it seems to be equally reasonable to reject R (causal history factors) and R (unidentified causal history factors). Causal history factors are factors which structure one's reasons for action, according to Malle. For example, an agent's low blood sugar may cause him to desire a snack before lunch, but his lacking awareness of his low blood sugar doesn't impede on the control he exercises when he acts on this desire. It still seems that he is able to act not only intentionally but freely as well despite his ignorance of the causal history factor that affects his reasons for action. Our lack of awareness of these features seems not to be problematic for acting freely.

That leaves four possible awareness requirements on the table. I'll look at each of them in turn. The first awareness requirement is the combination of reportable awareness (RA) and R (rationalize).

**Awareness Type A – RA & R (rationalize):** In order to act freely, an agent must be able to recognize and report (to herself or others) at least one reason for which she acts that rationalizes her action.

And the second type of awareness which is very similar is the combination of personally available awareness (PA) and R (rationalize).

**Awareness Type B – PA & R (rationalize):** In order to act freely, an agent must be able to recognize and report (to herself or others) effortlessly and without hesitation at least one reason for which she acts that rationalizes her action.

On the face of it, these each seem like reasonable requirements for acting freely. After all, if an agent is unable to report to herself or others at least one reason that would help someone see why she is doing what she is doing, we would begin to question the amount of control she had over her behavior. Say a woman, call her Marge, is standing in her front lawn in the middle of summer with her feet strapped into a set of skis and holding ski poles. No doubt, this would look a bit strange. But we would think her even stranger if she was unable to report to us why she was wearing skis during this time of year. If she replied, "I'm not sure" or "I don't know," we might think that she had gone mad and is unable to properly control her behavior. But if the woman responds to us by saying that she wanted to make sure the skis were the right size or that she is performing some sort of performance art we could make sense of what she is doing.

On the other hand, there are considerations that might suggest this requirement is too strong. It is sometimes the case that we don't know the reasons for which we act. This can happen especially if there are several reasons in favor of acting in some particular way. Perhaps Marge wants to show off her new skis and make her neighbor jealous, but she also wants to make sure they fit properly and her front lawn is the best space in which to do this. It could happen that Marge is not able to report to herself or others her desire to make her neighbor jealous by showing off her new skis yet this is the reason that caused her to put on the skis in her front yard. According to awareness type A and B, Marge would not be acting freely. But is this the correct judgment? At the very least, it's not clear that it is. Marge's behavior appears to be intentional and she isn't compelled or manipulated into acting the way she did. Perhaps the reason it seems unclear is because there are often many desires that constitute an agent's motivational base for acting. And just because an agent can't report a particular one that seems to have caused her to act in a certain way doesn't mean she is unable to report at least one reason that caused her to act. In this case, she would satisfy awareness requirements A and B. Awareness types A and B, then, seem to be plausible candidates as requirements for acting freely.

What about the other two types of awareness?

**Awareness Type C – RA & R (unidentified):** In order to act freely, an agent must be able to recognize and report (to herself or others) her reasons motivating her to act in a way that she does not identify with (if there are any such reasons).

**Awareness Type D – PA & R (unidentified):** In order to act freely, an agent must be able to recognize and report (to herself or others) effortlessly and without hesitation her reasons motivating her to act in a way that she does not identify with (if there are any such reasons).

Again, I think there are several initial considerations in favor of these two conditions. It seems intuitive that if an agent is motivated by and acts on reasons she does not agree with, she might not be acting freely. After all, this is a feature in many manipulation cases which are meant to generate the intuitive result that the agent is not free. But we can also consider an example. Say a male boss has just conducted a series of interviews and is deliberating about whom to hire. Unbeknownst to him, he is motivated by reasons that are sexist: he tends to think that women aren't as good workers as men. If the boss doesn't identify with making decisions in this way and doesn't want to be motivated by such considerations, is his making a decision based on such considerations free? According to awareness types C and D, the boss would only have made his decision freely if he was able to recognize and report to himself or others his reasons motivating him to act in a sexist way. And whether or not he has this ability might be a point of contention, but assuming he is a normally functioning agent with reflective capacities, he appears to be able to consider his own biases in judgment and decision making. Perhaps if we found out that he genuinely lacked the ability to recognize his sexist attitude we would think differently. Another reason in favor of judging the boss to be acting freely is that we would most likely hold him morally responsible. Many agree that one can only appropriately be held morally responsible for things which one does freely. Therefore, it would appear that awareness types C and D are also plausible requirements.

Awareness requirements A through D seem to be tracking something we care about in our intuitive judgments of acting freely. Even if they are not as precisely specified as they can be, they are plausible candidates for awareness requirements on free action. But what matters for

my purposes is whether these requirements are shown to be lacking given the situationist literature.

## 2.4 Situationism and Its Implications

The type of evidence Nahmias (forthcoming; 2007) considers comes from social psychology, in particular, a research tradition that has come to be known as “situationism.” Let me illustrate this with a paradigmatic example of such a psychological experiment referred to as the “Good Samaritan” study. Darley and Batson (1973) asked Princeton Theological Seminary students to participate in a study on “religious education and vocations.” The first part of the study involved the seminary students filling out a questionnaire in one building and the second part involved reporting to a second building to give a short verbal presentation. Before leaving the first building, the participants were told one of three things: they were running late, they were right on time, or they were a little early. These statements were aimed at exerting different degrees of time pressure on the subjects. On the walk between buildings the seminarians passed an experimental confederate who appeared in some sort of distress. The seminarians’ helping behavior in this experiment varied significantly according to the degree of time pressure exerted on them. Ten percent of those who were told they were running late stopped to help, 45% of those told they were on time helped, and 63% of the participants who were told they were early stopped to help. The concluding thought is that being a few minutes late is a significant situational factor which can impact whether someone ends up stopping to help someone in distress.

Another paradigmatic example often discussed was carried out by Latane and Rodin. Latane and Rodin solicited participation from Columbia University students who were told they were involved in market research. At the experimental site an attractive “market research representative” provided the students with a questionnaire to complete, then withdrew behind a curtain dividing the room and staged a distressing fall. Seventy percent of the students offered help when they were alone, whereas only 7% offered help when they were seated next to a confederate who lacked concern.<sup>11</sup> This experiment seems to demonstrate the purportedly

---

<sup>11</sup> Only 5% of the participants reported afterward that they suspected the distressing fall was faked (Latané and Rodin 1969).

overwhelming influence that features of a situation can have on a person's behavior—in this case, whether a person helps someone who appears to need assistance.

The literature is rife with other interesting experiments with similar findings. For example, psychologists have found that members of a group are more likely to dismiss evidence of their senses if subjected to patently false claims by a majority of others in the circumstance (Asch 1951); judgments of likeability can be affected by smells one is unaware of (Li et al. 2007); one's name can play a large role in the important life choices one makes (Anseel and Duyck 2008; Anseel and Duyck 2009; Pelham, Mirenberg, and Jones 2002); and there are many more.<sup>12</sup>

The general lesson from situationism seems to be that we underestimate the influence of certain factors on our behavior. Seemingly irrelevant features tend to have a profound influence over substantial types of behavior, including ethical behavior. And importantly, such seemingly irrelevant features influence our behavior without our awareness (in any sense discussed so far) of being so influenced. This shows that we tend to not be aware of causal history factors, unidentified causal history factors, and motives we don't identify with that influence our reasons for action. Occasionally, we might also lack awareness of some of the reasons for which we have acted on a given occasion even if we can report others. Does this show that we often do not satisfy awareness types A through D? No. It only shows that we often may not satisfy part of awareness requirements A through D. Even though we might lack awareness of certain things on a given occasion, requirements A through D state that agents must be able to recognize and report things like one's reason for action. Whether the situationist experiments rule out this sort of ability or capacity is a separate question and one in which I consider in the next chapter. But as things stand so far, it's not clear that situationism shows that an awareness requirement on acting freely is often not satisfied by most normally functioning humans.

---

<sup>12</sup> Others include Baumeister, DeWall, Ciarocco, & Twenge (2005), Haney, Banks, & Zimbardo (1973), Jones, Pelham, Mirenberg, & Hetts (2002), Jones, Pelham, Carvallo, & Mirenberg (2004), Milgram (2009), and Nuttin Jr. (1985).

## 2.5 Conclusion

In this chapter I have attempted to identify some intuitively plausible awareness requirements on acting freely. This involved specifying what it is that one ought to be aware of and what type of awareness one ought to have of those things. After considering several different types of awareness, I argued that only four of them were plausible candidates for a requirement on acting freely. And after looking at the situationist literature, I argued that though it might appear that situationism shows awareness of certain things to be lacking, it doesn't show that the most plausible requirements I have identified are not satisfied. I don't claim to have definitively resolved this issue, however. There may be a more plausible type of awareness that I have not discussed that is ruled out by situationism. But until free will theorists more precisely specify what is required for free will it remains difficult to see just exactly what such a requirement would be. Or, alternatively, it might turn out that situationism shows that we lack the ability to recognize and report our reasons for action. And it is this question I'll address in the next chapter. At the very least I hope to have brought attention to more particular ways in which this discussion may continue.



## CHAPTER THREE

### FREE WILL AND CAPACITIES

In this chapter my starting point for addressing whether we actually have free will is an argument put forth by Eddy Nahmias (forthcoming) concluding that we have less free will than we generally think we have. In section 3.1, I introduce Nahmias's argument and raise problems for it. In particular, I think his argument is ambiguous in several respects. I attempt to clear up such ambiguities and recast his argument in terms of agents' capacities. In section 3.2 I attempt to more carefully evaluate whether the psychological evidence Nahmias considers shows that we have less free will than we generally think we have.

#### 3.1 Condition (MR)

Eddy Nahmias has argued that empirical evidence provided by social psychology shows that we have “significantly less free will than we think we have” (forthcoming, p. 16). He arrives at this conclusion by distilling necessary conditions of free will that theorists—compatibilist, libertarian, and free will skeptics—and the folk (i.e., non-philosophers) agree upon, and then argues that psychological evidence gives us good reason to think that these conditions are often not met. In what follows, I will critically evaluate one of the conditions Nahmias constructs.

Nahmias begins to develop a partial conception of free will on two fronts: what most people think is involved in free will and what free will theorists require for free will. According to Nahmias, most people think that what allows us as opposed to infants and animals to act freely is that we have certain cognitive capacities, e.g., the capacity to consciously deliberate, the capacity to reflect on our reasons for action, and the capacity to make choices on the basis of reasons.<sup>13</sup> In addition to what most people think allows us to act freely, Nahmias points out that most free will theorists, irrespective of what particular theoretical position they advocate, also agree that free agents must have certain cognitive capacities. First, he cites two compatibilists. Harry Frankfurt says that a free agent has “the capacity for reflective self-evaluation” (1971, p. 7) and

---

<sup>13</sup> Recently, Stillman et al. (2011) have shown that laypersons associate high levels of conscious thought and deliberation among the features essential to acting freely.

is “prepared to endorse or repudiate the motives from which he acts . . . to guide his conduct in accordance with what he really cares about” (1999, pp. 113-14). R. Jay Wallace mentions the capacity of “reflective self-control,” described as “the power to grasp and apply moral reasons, and the power to control or regulate one’s behavior by the light of such reasons” (1994, p. 157). Next, Nahmias cites two libertarian views. Laura Ekstrom claims that “[a]n agent enjoys freedom of action only if the agent’s act results from a preference—that is, a desire formed by a process of critical evaluation with respect to one’s conception of the good” (2002, p. 108). Timothy O’Connor claims that free agents are “conscious, intelligent agents, capable of representing diverse, sophisticated plans of action,” and he is “unable to conceive an agent’s [freely] controlling his own activity without any awareness of what is motivating him” (2002, pp. 121, 88). And lastly, Nahmias cites Richard Double—a free will skeptic—who says free will requires self-knowledge, i.e., the agent “knows the nature of [her] beliefs, desires and other mental states that bring about [her choice]” (1991, p. 48).<sup>14</sup>

These quotations illustrate the large degree of overlap among free will theorists with respect to requiring cognitive capacities for free will even though philosophers may disagree about exactly how to specify the capacities at issue. Despite the subtle differences in each account, Nahmias constructs two conditions which he claims capture the overlap among the theorists’ requirements. My concern here is only with one of the conditions.<sup>15</sup> It is as follows:

(MR) *motivation by (potentially) endorsed reasons*: agents’ free will is diminished to the extent their actions are motivated by factors that they are both unaware of and would reject were they to consciously consider them (forthcoming, 11).

There are several difficulties in interpreting this condition. What needs to be spelled out first is what ‘diminished free will’ means. Nahmias is not entirely clear what he means by this

---

<sup>14</sup> These examples are quite representative of the free will literature. Authors who mention similar cognitive capacities as a requirement for free will include Fischer and Ravizza (2000), Robert Kane (1996), Daniel Dennett (1984), Susan Wolf (1993), Alfred Mele (1995), and Galen Strawson (1986).

<sup>15</sup> The other condition distilled by Nahmias is the following: “(CR) *conscious reflection*: agents have free will only if they have the capacity for conscious deliberation and intention-formation and that capacity has some influence on their actions” (forthcoming, 11). Nahmias considers psychological evidence which purports to show that (CR) is generally not met by human agents. Such evidence comes primarily from Daniel Wegner (2002) and Benjamin Libet (1999), but examining the evidence in greater detail leads Nahmias to conclude that neither Wegner nor Libet has shown that conscious deliberation and intention-formation do not have an influence on our behavior. Mele (2010) gives a very detailed treatment of the evidence provided by Wegner and Libet.

phrase. Having diminished free will could mean that one acts freely less frequently, or it could mean that free will comes in degrees and when one acts he exercises a lesser degree of free will. Nahmias does seem to be aware of both of these senses of ‘diminished free will’. He says, when discussing evidence that would be contra (MR), that his question is “whether these challenges to our free will are more pervasive than we realize,” which seems to indicate the first notion of diminished free will (forthcoming, 11). Elsewhere, he writes, “cognitive capacities are possessed and exercised in varying degrees. . . and individuals may exercise more or less freedom in particular actions,” which indicates the second sense of having diminished free will (forthcoming, 3). More often, Nahmias seems to intend ‘diminished free will’ in the second of these senses. Of course, Nahmias could have in mind a combination of these two notions. On this reading, having diminished free will would mean that we act with a lesser degree of free will than we generally think we do and this phenomenon is more prevalent than we realize. But these are very different interpretations that need to be separated and evaluated independently. We could do this by specifying a condition on the frequency with which one acts freely and specifying a condition on the degree to which one acts freely.

(MR1) *frequency*: agents act freely less frequently to the extent their actions are motivated by factors that they are both unaware of and would reject were they to consciously consider them.

(MR2) *degree*: agents act with a lesser degree of freedom with respect to a particular action to the extent their actions are motivated by factors that they are both unaware of and would reject were they to consciously consider them.

But what is meant by acting with greater or lesser degrees of free will? Many free will theorists take free will to be a degreed notion, but not all specify what makes a particular act less free than it otherwise would be on a given occasion.<sup>16</sup> “Less free will” is a relative notion; it leads one to ask, “relative to what?” One obvious answer is “less than ideal free will.” Acting with a “lesser degree of free will” according to this answer would be relative to acting with ideal

---

<sup>16</sup> For example, Double (1991), Dennett (1984), Pereboom (2001), and O’Connor (2009), all acknowledge that free will comes in degrees and discuss in what sense one could act with a lesser degree of free will.

free will. Another answer is “less free will than what we think.” On this reading acting with a “lesser degree of free will” is meant to be relative to how much free will we generally think we have. But these are very different claims that would require different evidence in order to confirm or undermine. So again, we can further specify a condition on acting freely by modifying (MR) to accommodate these two interpretations.

(MR1.1) agents act freely less frequently than we think they do to the extent their actions are motivated by factors that they are both unaware of and would reject were they to consciously consider them.

(MR2.1) agents act with less than ideal freedom with respect to a particular action to the extent their actions are motivated by factors that they are both unaware of and would reject were they to consciously consider them.

(MR2.2) agents act with less freedom than we think they do with respect to a particular action to the extent their actions are motivated by factors that they are both unaware of and would reject were they to consciously consider them.

Since Nahmias’s argument is meant to establish a conclusion about how much freedom we have relative to the freedom we think we have, our focus will be on (MR1.1) and (MR2.2). Comparing these conditions with a few things Nahmias says may inspire us to make further modifications to (MR). Nahmias argues that psychological evidence showing our actions to be motivated by factors that we are both unaware of and would reject were we to consciously consider them also shows our “capacities to be limited.”<sup>17</sup> He says that “*to the extent* that our ignorance of the influence of situational factors limits our capacities to act on reasons we accept, it thereby limits the scope of our free will” (forthcoming, 15). Two things require clarification. The first is minor and the second is more substantial.

---

<sup>17</sup> This is Nahmias’s main reason in an earlier paper in which he argues that autonomy (which he takes to be a broader notion than free will) “is compromised” because social psychological evidence indicates that our capacity to act in line with our reasons is diminished (Nahmias 2007).

First, the idea of rejecting factors which motivate one may be unclear. The idea seems to be that the agent would reject the claim that these factors should motivate them, and Nahmias gives evidence in support of the claim that the participants in the situationist experiments in fact reject the claim that the seemingly insubstantial situational factors do and should motivate them. Similarly, the notion of “acting on reasons we accept” could be clarified as “acting on reasons we want to be motivated by.” Second, condition (MR) isn’t a condition on an agent’s capacities at all. The condition which Nahmias develops omits capacities and powers but instead expresses a condition on what agents must actually be aware of. This is quite problematic for at least two reasons. First, Nahmias’s aim is to capture a condition that free will theorists generally agree upon. Insofar as free will theorists require an agent to have a capacity and do not specify to what extent the capacity must be exercised, it’s unclear that a condition on an agent’s actually exercising the capacity can be used to accurately gauge the agent’s free will.<sup>18</sup> Secondly, and perhaps more importantly, Nahmias’s argument for the claim that we have “significantly less free will than we think we have” hinges on what the psychological evidence he examines shows about capacities (forthcoming, 16). So some thesis is needed connecting failure to act in a certain way with a relative lack of capacity to act in that way.

There are two clear ways in which Nahmias thinks that the situationist experiments reveal something about our capacities: the extent to which we possess certain capacities and the extent to which we exercise certain capacities. He claims that “it may be possible to convince people that we do not possess these capacities to the extent we think we do—that humans have *less* free will than we tend to think we have,” and he notes that there is evidence which “challenges the extent to which we possess and exercise the relevant cognitive capacities [for free will]” (forthcoming, 10). An agent’s possession of a general capacity seems to indicate that in certain circumstances the agent is able to exercise their capacity. In other words, possession seems to indicate the ability to exercise. However, exercising a capacity less than one thinks doesn’t itself show that the agent has less of that capacity than they think they do. The free will theorists mentioned above seem to agree upon requiring the agent possess certain capacities. Possessing the relevant capacities, according to most theorists, indicates that the agent can, at least in some

---

<sup>18</sup> Mele (1995) and Double (1991) each develop sufficient conditions on certain capacities and require that agents regularly exercise these capacities. But since the conditions discussed by Mele and Double are sufficient conditions, their requirements cannot be taken to be representative of what other theorists generally require.

situations, exercise those capacities. Failure to exercise their capacities alone doesn't indicate that the agent couldn't have exercised their capacities.

Mele (2003a) provides us with a distinction that is useful in this context between *specific* practical abilities and *general* practical abilities.<sup>19</sup> Specific practical abilities are abilities “an agent has at a time to *A* then or to *A* on some specified later occasion” (p. 447). A general practical ability, on the other hand, “is the kind of ability to *A* that we attribute to agents even though we know they have no opportunity to *A* at the time of the attribution and we have no specific occasion for the *A*-ing in mind” (p. 447). The ability to exercise one's capacity in a given situation is an example of a specific practical ability. It's also important to note that philosophers theorizing about free will disagree about the nature of capacities or abilities that are required to act freely. In particular, libertarians and classical compatibilists<sup>20</sup> often require that the agent have certain specific practical abilities, for example, the specific practical ability to choose to do otherwise in a given situation. Compatibilists who do not think free will requires the ability to do otherwise, on the other hand, often attempt to give sufficient conditions for free will that focus on general practical abilities, for example, the general practical ability to be receptive and reactive to reasons.<sup>21</sup> However, even agents who meet such compatibilists sufficient conditions must also, at least sometimes, have the specific practical ability to exercise their capacities. Given this, the condition we are refining should include both the ability to exercise one's capacity understood as a specific practical ability and a general practical ability. This suggests the following reformulations on condition (MR).

(MR 1.2) agents act freely less frequently than we think they do to the extent that their general capacity to act on reasons they want to be motivated by, or would want to be motivated by were they to consciously consider them, is possessed less often when they act than we think it is.

(MR 1.3) agents act freely less frequently than we think they do to the extent that their specific capacity to act on reasons they want to be motivated by, or would want to be

---

<sup>19</sup> For all intents and purposes, I am using 'capacity' and 'ability' interchangeably.

<sup>20</sup> Classical compatibilism is the position that determinism does not rule out the ability to do otherwise.

<sup>21</sup> See, for example, Fischer and Ravizza (2000) and Whittle (2010).

motivated by were they to consciously consider them, is possessed less often when they act than we think it is.

(MR 2.3) agents act with less freedom than we think they do with respect to a particular action to the extent their general capacity to act on reasons they want to be motivated by, or would want to be motivated by were they to consciously consider them, is possessed to a lesser extent than we think it is.

(MR 2.4) agents act with less freedom than we think they do with respect to a particular action to the extent their specific capacity to act on reasons they want to be motivated by, or would want to be motivated by were they to consciously consider them, is possessed to a lesser extent than we think it is.

These modifications and refinements to Nahmias's condition put us in a better position to evaluate the extent to which the psychological evidence Nahmias appeals to reveals limitations to our freedom,

I won't spend much time repeating what situationism is since I described it and some paradigmatic experiments in the previous chapter. But just as a reminder, the general lesson from situationism seems to be that we underestimate the influence of certain seemingly irrelevant factors on our behavior. Seemingly insubstantial features tend to have a profound influence over substantial types of behavior, including ethical behavior. And importantly, such insubstantial features influence our behavior without our awareness of being so influenced. Recall, in particular, the Latane and Rodin study which seems to demonstrate the bystander effect—people are much less likely to offer help to someone in need when they are not alone. Given this evidence, Nahmias asks us to suppose that “you and a group of friends do not stop to help a woman who needs help and she turns out to suffer great harm that you would have prevented . . . [a]nd . . . as the experimental results suggest, you would have helped the woman had you been alone” (forthcoming, 15). Nahmias argues that if you are like most people and do not know the effects a group has on an individual member with respect to helping behavior<sup>22</sup> and you reject the idea that it should influence you, then your free will is threatened (forthcoming,

---

<sup>22</sup> Importantly, participants of the experiments deny situational factors actually did influence them, says Nahmias.

15).<sup>23</sup> Nahmias claims that the evidence provided by this body of literature in conjunction with condition (MR) implies that free will is often diminished because we're acting with a lesser degree of free will more of the time than we think. In the next section, I will evaluate Nahmias's argument for this conclusion.

### 3.2 Evaluating the Evidence

Does this literature in conjunction with (MR) indicate we have “significantly less free will than we think we have” (forthcoming, 16)? Recall our modified (MR) conditions:

(MR 1.2) agents act freely less frequently than we think they do to the extent that their general capacity to act on reasons they want to be motivated by, or would want to be motivated by were they to consciously consider them, is possessed less often when they act than we think it is.

(MR 1.3) agents act freely less frequently than we think they do to the extent that their specific capacity to act on reasons they want to be motivated by, or would want to be motivated by were they to consciously consider them, is possessed less often when they act than we think it is.

(MR 2.3) agents act with less freedom than we think they do with respect to a particular action to the extent their general capacity to act on reasons they want to be motivated by, or would want to be motivated by were they to consciously consider them, is possessed to a lesser extent than we think it is.

(MR 2.4) agents act with less freedom than we think they do with respect to a particular action to the extent their specific capacity to act on reasons they want to be motivated by, or would want to be motivated by were they to consciously consider them, is possessed to a lesser extent than we think it is.

---

<sup>23</sup> Jokingly, Nahmias apologizes for increasing the reader's responsibility by informing them of the group effect.



Answering the above question requires that we evaluate whether the situationist experiments show that in general agents' general and specific capacities to act on reasons they want to be motivated by, or would want to be motivated by were they to consciously consider them, is possessed less often than we think or to a lesser degree than we think. Nahmias doesn't provide an argument for this claim and it's not entirely clear why he thinks that the situationist experiments do show this. Further, there are at least two ways in which agents' capacities could be possessed less often or to a lesser degree than we think in such a way that the freedom they have or exercise is impacted.<sup>24</sup> Nahmias could mean to argue one of two things: agents' capacities are reduced or eliminated temporarily in virtue of the situation they find themselves in, or he could mean that their capacities are diminished chronically or are generally possessed by the agent more weakly than we think.<sup>25</sup>

*Chronic* degradation of a capacity is when an agent is impaired due to lacking a sufficiently developed capacity in general. Many disorders afflicting agents are cases of chronic degradation of a capacity. For example, Tourette's syndrome is an impairment of an agent's capacity for motor and verbal control. Such people experience involuntary movements of the body and involuntary utterances (sometimes profane). This neural disorder, though it can wax and wane in severity, is a chronic condition for which there is no known cure. Tourette's syndrome patients suffer from a diminished capacity to control their behavior in ways they would want. The second way an agent's capacity could be possessed to a lesser extent is by *transient* capacity disruptions. These are disruptions which temporarily reduce an agent's ability to exercise a given capacity. Suffering from extreme stress due to the loss of one's job may temporarily impair an agent's capacity to control his mood. His capacity to cheer himself up or to not dwell on negative thoughts hasn't been permanently eliminated (one would hope); he may be able to effectively boost his mood once his stressful situation passes.

Each of these types of impairment would be problematic for agents' freedom. If our cognitive capacities were temporarily disrupted in ways we are unaware of in everyday

---

<sup>24</sup> Nahmias isn't quite clear about how he thinks our capacities are affected. On the one hand, he writes that situational "influences limit your ability to do what you think you should do" and that "our ignorance of the influence of situational factors limits our capacities to act on reasons we accept" (forthcoming, 15). This seems to indicate that he is thinking that our capacities are temporarily disrupted. On the other hand, he also writes that "it may be possible to convince people that we do not possess these capacities to the extent we think we do—that humans have *less* free will than we tend to think we have" (forthcoming, 10). This sounds like he is thinking that our capacities are weak in general.

<sup>25</sup> The following two types of capacity impairment are discussed by Doris and Murphy (2007).

situations, we would seem to be more limited in the control we can exercise over our behavior, since the disruptions would impede our capacities integral to our control. Loud music in the background may disrupt our capacity to focus our attention on reading a novel. Similarly, if our capacities were shown to be generally impaired or weaker than we think they are, then again we would seem to be able to exercise less control over our behavior than we think. If our capacity to reflect on what reasons we have to perform an action were constantly disrupted by a neurological disorder we might have less control than we think we do.

An argument that we act less freely with respect to particular actions might go something like this. The capacity that seems to be most affected by the situationist literature is our capacity to recognize or be aware of what factors are motivating us. If the evidence were to show that this capacity is temporarily disrupted such that sometimes we possess it less than we think, it might show that our capacity to ensure that we are motivated only by factors we want to be motivated by is temporarily possessed less than we think it is. Then, according to (MR 1.2), (MR 1.3), (MR 2.3) and (MR 2.4), this would show that we both act freely less frequently than we think and act less freely with respect to particular actions than we think we do.

The evidence in general seems to show that unbeknownst to agents insubstantial features of one's situation have an impact on the way that person behaves. The question is whether the situation they are in results in reducing the participants' general and specific capacities to recognize which factors are actually motivating them. These capacities would seem to be reduced if the participants were to attempt to recognize what features are motivating them but were unable to do so. But the experiments don't seem to show this. The situations in which people are being tested don't appear to be akin to the types of impairments discussed earlier. They appear to be the type of ordinary situations that people find themselves in everyday such as walking out of a telephone booth (if these are still around) or walking across a campus.<sup>26</sup> These types of situations do not seem to disrupt one's capacity to recognize what is motivating them in the way one's capacity to focus their attention could be disrupted by loud music. Simply because the participants are unaware or fail to recognize the situational features which influence their action doesn't seem to indicate that they are incapable of recognizing those features. Lack of awareness alone does not diminish or reduce one's general capacity.

---

<sup>26</sup> The Milgram experiments are an exception.

To make this point clearer, consider an analogous case. A weightlifter with the capacity to bench press a significant amount of weight would, at a later time, possess his capacity to a lesser extent if the amount of weight he could press was less than before. The extent to which the weightlifter possesses his capacity to bench press a significant amount of weight is determined by the amount of weight he presses. If the weightlifter attempted to press 300 pounds and failed, this would show that he lacks the specific practical ability to press 300 pounds. And if the weightlifter could press 300 pounds on a later occasion, but not now because he is too tired, then he would seem to possess the general practical ability to press 300 pounds. If not, then he might lack the general ability. It hasn't been shown that in the situationist experiments our specific or general capacity is reduced or diminished in virtue of being in that situation. It might be that, were we to investigate the factors that are motivating us we would recognize some of the many seemingly insubstantial features which affect our behavior. The things we are and are not aware of do not seem to determine the extent to which we possess the capacity to act on reasons we want to be motivated by. So the experiments don't seem to show that our general capacity to act on reasons we want to be motivated by, or would want to be motivated by were we to become aware of them, has been compromised or reduced. Nor do the experiments seem to show that our specific capacity to act on reasons we want to be motivated by, or would want to be motivated by were we to become aware of them, has been eliminated.<sup>27</sup> Widespread failure to exercise one's capacity might be evidence for the fact that we generally lack the ability to exercise our general capacity, one could argue. But the situationist experiments are consistent with the claim that the participants nonetheless possess the specific ability at issue so it cannot be claimed that the experiments show such an ability to be lacking.

How about chronic capacity degradation? Do the situationist experiments show that our freedom-relevant capacities are possessed to a lesser extent than we think in this sense? Again, the capacities that seem to be most affected by the situationist literature are our capacities to recognize or be aware of what factors are motivating us. If the evidence were to show that these capacities are generally weaker than or possessed less than we think they are, it might show that in general our capacities to ensure that we are motivated only by factors we want to be motivated

---

<sup>27</sup> John Doris (2002) argues that situationism may problematize attributions of moral responsibility (rather than freedom) in virtue of disrupting our normative competence. Although he seems to have in mind transient disruptions, Doris is not clear about whether he takes situationism to show that our capacities have been diminished in strength or that some situations disallow the exercise of our capacities.

by are generally weaker than or possessed less than we think they are. Then, according to (MR 1.2), (MR 1.3), (MR 2.3), and (MR 2.4), this might show that we both act freely less frequently than we think and act less freely with respect to particular actions than we think we do.<sup>28</sup> We can evaluate each step of this argument in turn.

Is our general capacity to recognize or be aware of what factors are motivating us shown to be weaker than we think it is? I think the answer to this question is ‘no’ for the following reason: it’s not even clear that the situationist experiments show anything about the strength of our capacity (e.g., that it is weak), therefore, *a fortiori*, it is not clear that the evidence shows it’s weaker than we think it is. To reiterate, the situationist experiments haven’t shown that our lack of awareness in and of itself indicates that we lack the capacities requisite for free will. Similarly, the same point can be made with respect to our specific capacity to recognize and be aware of what factors are motivating us.

While it may be the case that the situationist experiments show that there are factors which motivate us yet we are unaware of, a positive answer to the above question requires knowledge of how strong we generally think these capacities of ours to be. To my knowledge, no such evidence is available. In addition, I claim it’s not clear the experiments show anything about the strength of our general or specific capacity to recognize or be aware of the factors that are motivating us, because it’s not clear that the participants in the experiments are unable to become aware of what is motivating them.

One might be inclined to think that the situationist experiments show that our capacity to recognize the factors which motivate us is chronically weak because we rarely if ever recognize the myriad of insubstantial situational factors which appear to be influencing our behavior. We seem to be systematically blind to recognizing such factors as relevant for present deliberation. Daniel Dennett (1984) suggests that there is a plausible evolutionary explanation for why we might fail to recognize such features. He argues that “it is (higher-order) rational of us to cede a bit of our (lower-order) rationality in the interests of efficient, speedy decision making” (1984, p. 69).

In general, situationism may seem to establish the fact that it is perhaps more prevalent than we realize that we fail to recognize what factors influence our behavior. Insofar as awareness of the factors that motivate us could afford us further control over our behavior, one

---

<sup>28</sup> I thank Randolph Clarke for bringing this interpretation of Nahmias’s argument to my attention.

might think that we seem to fail to exercise some control that we could exercise. Dennett (1984) offers an analysis of control in which he highlights the importance of awareness and knowledge of factors affecting us. Dennett writes:

Foreknowledge is what permits control. Circumstances that are clearly beyond our control may nevertheless not disrupt or prevent our control of the events we wish to control—if we know about those circumstances in time. For if we know about them in time, we can plan in light of our expectations, and take steps to prevent, avoid, preempt, avert, harness, exploit, or accommodate ourselves to those circumstances (1984, pp. 54–5).

One example Dennett gives is that of a bobsledder who is on a familiar course. Since the bobsledder is accustomed to riding on this course, he is “much more in control, thanks to foreknowledge of conditions, than one who is on a maiden run” (1984, p. 54). This type of control is often discussed with respect to self-control. By being more aware of features that affect one’s behavior one is much more likely to effectively guide their behavior such that it is in line with reasons they consciously endorse as good reasons to act upon, or would endorse were they to reflect on such reasons, and judgments they make about what they ought to do. Being aware of the influence being in a group has on our behavior may prompt one to reconsider their decision to not stop and help someone who seems to be in distress. In this way, we may be able to better ensure that we act for reasons we want to motivate us.<sup>29</sup> So it does seem as though there is a type of control that we do not exercise when insubstantial features of our environment influence us to act in certain ways that we would not act were we to become aware of the influence of such features.

But, it’s not clear that even if the situationist experiments show our capacity to recognize factors which motivate us is limited that it is limited with respect to the strength of capacity we think we have. And it is this last claim that needs to be shown in order for Nahmias to conclude that we have less free will than we think we do. Nahmias’s argument seems to show that there is

---

<sup>29</sup> In fact, Beaman et al. (1978) conducted an experiment to test whether knowledge of the group effect had an impact on helping behavior and concluded that such knowledge significantly increased the rate at which people decided to help.

some further control that we could exercise or that if we exercised our capacities ideally that we would have more control over our behavior, but this doesn't indicate that we have less free will than we think unless we generally think that we act ideally freely.

### 3.3 Extending the Argument

Dana Nelkin (2005) and Manuel Vargas (forthcoming) advance arguments similar to Nahmias's argument just considered. Nelkin's strategy, however, is unique. She first offers a survey of the situationist literature and then explores some of the possible reasons this body of evidence may seem to pose a threat to freedom and responsibility. After exploring a number of possible reasons the situationist experiments may seem to threaten our freedom and responsibility, Nelkin argues that the most potent threat seems to be that the literature suggests "we aren't—and are perhaps systematically blocked from—responding to good reasons" (2005, 200). Elsewhere, she says that the evidence "might suggest that we have trouble translating our commitments into actions" and that "for many there is a failure to apply their general commitments in the particular situation" (2005, 201). In general, the situationist literature seems to indicate that we lack the required capacities of reasons-responsiveness in order to act freely and morally responsible. Similarly, Vargas (forthcoming) argues that one might take the threat of situationism to our freedom and moral responsibility to be that our rational capacities are "very fragile and bounded" (forthcoming, 15). Although Vargas doesn't spend much time elaborating precisely what he takes this threat to be he does seem to have in mind the idea that we lack the reasons-responsive capacities that many take to be required for free will and morally responsible agency.

These threats just identified are similar to the threat discussed by Nahmias in that it deals with challenging the idea that we have the requisite capacities to be free and morally responsible. However, care needs to be taken in precisely identifying exactly what the threat is. For example, Nelkin seems to be suggesting, perhaps, three possible threats: (i) that we do not respond to good reasons, (ii) that we are systematically blocked from responding to good reasons, and (iii) that we have trouble translating our general commitments into action. It is unclear to me whether Nelkin construes these to be threats to our general or specific capacities, but, either way, it seems that one can respond to Nelkin and Vargas in the same way I have responded to Nahmias. In

particular, once we recognize that the situationist experiments don't demonstrate that we lack the abilities at issue as opposed to showing that such abilities are not exercised, it becomes clear that claims about such experiments diminishing or eliminating or freedom are not warranted. I think that Nelkin, in fact, recognizes this. She actually concludes from her investigation that "situationism does not provide a monolithic threat to either freedom or responsibility" (2005, 204).

Even if we set this reason aside, one may still want to press the idea that despite the absence of a monolithic or global threat to our freedom and responsibility, there does seem to be at least wide-spread failure to exercise certain capacities. I have argued that this wide-spread failure to exercise one's general and specific capacities doesn't indicate that they are weakened or eliminated. But it could also be evidence for weakened capacities. In this respect, I think the arguments I have considered are tracking something important about what the situationist literature indicates about our capacities and abilities. But what that is remains elusive.

## CHAPTER FOUR

### FREE WILL AND SELF-CONTROL

Nearly all free will theorists agree that capacities for self-control and conscious decision making are necessary in order to act freely. But recently, psychologists have conducted experiments demonstrating that our capacity for self-control and conscious decision making rely on limited resources the depletion of which impairs the capacity to regulate our behavior and make reasonable decisions.<sup>30</sup> This research has led some to draw implications about our prospects for acting freely, claiming that “free will is limited” and “free will is at best an occasional phenomenon” (Baumeister, 2008a, p. 17). The aim of this chapter is to evaluate whether the experiments warrant these skeptical claims about our ability to act freely. I’ll begin this chapter by giving some philosophical background on the nature of self-control and explain to what extent self-control is important for free will. Then I will explain the self-control experiments conducted primarily by Baumeister and evaluate whether the claims he makes about free will are supported by said experiments. I will then argue that although the self-control experiments may give us some insight into the nature of the control we have, they don’t warrant the skeptical claims made about free will.

#### 4.1 Philosophy and Self-Control

What is self-control and what is it to consciously make a decision? How are these to be understood? Here I would like to give some background on the philosophical work on self-control and conscious decision making. Then I will demonstrate how commonly these notions are employed as requirements for acting freely.

##### 4.1.1

Mele (1987) develops a thorough account of self-control and understands it as the opposite of *akrasia* or weakness of will (p. 50). On his account, having self-control means having the ability to do what you judge it best to do even when you are more motivated to do

---

<sup>30</sup> For example, Baumeister et al. (2007), Baumeister (2008a), Baumeister (2008b), and Vohs et al. (2008).



something else instead. When you yield to such motivation and fail to do what you judge it best to do, you act akratically. And it is important to note that akratic actions are commonly conceived of as intentional and free actions; according to Mele, this is what distinguishes compelled actions from akratic actions (p. 4).

There are several other dimensions to Mele's account of self-control (and akrasia).<sup>31</sup> Self-control can be either global or local. A highly self-controlled agent could be globally self-controlled, exercising self-control in all domains of her life. Another agent may have imperfect self-control and only exercise this capacity in certain domains of her life but not others. She may exercise self-control in overcoming temptation to not stick to her diet, but she may not exercise self-control when it comes to resisting smoking cigarettes.

Also, self-control is often discussed as exercised in the service of one's better or best judgment about a specific action or type of action, a "judgment that is based on (and supported by) a consideration of the totality of the agent's pertinent beliefs and values (including desires, fears, principles, etc.)" or "judgments made from the perspective of a segment of this totality which the agent takes to be overriding" (1987, 6).<sup>32</sup> Greg, who has made the judgment given all his beliefs and values that it's best to put in a full day's work rather than leave early to go golfing, acts with strength of will when he successfully resist the temptation to sneak out of the office. Since there is quite a lot involved in the making of such a judgment, one may note that these judgments may be a rare phenomenon in the real world. But this doesn't mean no one ever exercises self-control; self-control can also be exercised in service of more general judgments and better or best judgments the agent would have made (Mele 1987, 6). Greg may have instead made a more general judgment that he shouldn't unnecessarily jeopardize his job. Given this judgment, his resisting the temptation to leave work early may still count as an exercise of self-control. Similarly, self-control can be exercised in the service of judgments Greg would have made had he spent more time reflecting rather than focusing on the importance of not risking getting caught leaving early by his boss.

And lastly, self-control on this account extends beyond the sphere of actions to include self-control with respect to beliefs, emotions, principles, and values. So, self-controlled agents

---

<sup>31</sup> Since Mele takes these concepts to be two sides of the same coin, the dimensions I mention about self-control often have an akratic analogue. See Mele (1987; 1995).

<sup>32</sup> Mele also describes non-standard cases of self-control that don't exactly fit this description.

possess significant motivation to conduct themselves as they judge best, and they also possess the capacity to do what it takes to conduct themselves in that moment in the face of competing motivation.

#### 4.1.2

Decisions can be theoretical or practical. Theoretical decisions are decisions that something is the case. Sherlock Holmes might decide that all the evidence available to him indicates that Moriarty is the villain. The decisions that are of interest to me are practical decisions, decisions about what to do. One common understanding of practical decisions is that they are mental actions of intention formation (Clarke 2003, 22–27; Frankfurt 1988, 170–174; Kane 1996, 24; Mele 2003b). Generally, the practical decisions are conscious and we are aware of making such decisions. The intention involved in practical decisions is simply a plan of action that settles the question about what one shall do. Note that this is distinct from settling the question of what one should do.

#### 4.1.3

How important are self-control and conscious decision making for a theory of free will? Daniel Dennett (1984) claims that the concept of self-control is “utterly central to the questions of free will and determinism” (p. 19). Much of Dennett’s theorizing about free will involves giving an account of self-control and why we think it is important. And R. Jay Wallace (1994) argues that acting freely requires what he calls reflective self-control (p. 157). Others describe free will as requiring capacities that are closely related to self-control. Robert Kane (1996) argues that free will is associated with autonomy, self-governance, and self-creation (p. 80). And as we have seen, self-control is intimately related (if not identical to) these notions. Randolph Clarke (2003) claims that free agents must be able to make judgments about which course of action is better or best and make such judgments effective in action; this includes the capacity to reflect on one’s reason-states and the capacity to modify them (p. 16). And finally, Alfred R. Mele (1995) writes that “[h]uman beings *wholly* lacking self-control are at the mercy of whatever desires happen to be strongest, even when the desires clash with their better judgments” (p. 5). And so it seems the capacity for self-control is crucial to acting freely.

Similarly, much of the theorizing about free will focuses on what it means for one to act freely. This must naturally include making decisions freely since many theorists take deciding to be an action. Kane writes that free will includes “the powers to deliberate, or to reason practically, to choose or decide, to make practical judgments, to form intentions...” and so on (1996, 22). Fischer and Ravizza include in their analysis of the control required for free will that the agent have the capacity to recognize and react to reasons, which they take to naturally include decision making. Clarke (1992) writes that he takes free will to be a type of self-determination which involves “coming to have the particular intention with which one acts” (p. 54). And so it seems, conscious decision making—not surprisingly—is a crucial element required for one’s acting freely.

## 4.2 Psychology and Self-Control

Many psychologists investigate our capacities for self-control, developing theories meant to explain their data.<sup>33</sup> But since psychologists are involved in an endeavor distinct from philosophy, it is worth asking whether when psychologists use the term ‘self-control’ they are referring to the same thing as philosophers.

Psychologists investigating self-control use the terms ‘self-control’ and ‘self-regulation’ to refer to behavior that involves exercising control over oneself “especially with regard to bringing the self into line with preferred (thus, regular) standards” (Roy F. Baumeister 2004, 2). Compare this with a type of regulation found to occur naturally in our bodies, such as the endocrine system’s regulation of hormone levels. This latter type of regulation isn’t done by the self—it’s not done by us—in at least some sense. The type of self-regulation psychologists seem to be interested in is regulation that the psychological self does. As Baumeister notes, “one definition of “self-regulation” encompasses any efforts by the human self to alter any of its own inner states or responses” (2004, 2). This would include thoughts, emotions, impulses, appetites, moods, actions, and many other things.

One question that arises is whether self-control must be a conscious process. It seems that conscious regulation of the self is the focus of most psychological research, but, Baumeister

---

<sup>33</sup> See, for example Baumeister and Vohs (2004).

claims, evidence is accumulating which suggests the importance of non-conscious processes to self-regulation. In any case, this matter need not be settled for my purposes here.

Now one can see how the term 'self-control' in psychology compares with 'self-control' as it's used by philosophers. Self-control as understood by psychologists seems to encapsulate a broader spectrum of behavior, since the particular attitude or standard the self aims at bringing behavior into accordance with isn't precisely specified. For psychologists, a natural understanding of standards suffices. Again, this doesn't seem problematic. While the psychological use of the term 'self-control' may not correspond to strict self-control the way Mele has analyzed it, it seems to refer to the same type of capacity of an agent. This is true and can be seen by reminding ourselves that Mele also allows for exercises of self-control in accordance with not just decisive best judgments but judgments one would have made and judgments that are general, not specific.

Psychologists claim to have demonstrated that our capacities for self-control rely on resources which can become depleted and, once depleted, decrease the likelihood that we will subsequently successfully exercise our capacities for self-control. This phenomenon has been termed *ego depletion*. One of the initial experiments that demonstrated this effect had participants sit near a batch of freshly baked cookies with the pleasant fresh-baked smell still in the air. The subjects were instructed to refrain from eating for at least 3 hours prior to the experiment so they would be sufficiently tempted to eat the cookies. However, they were told they were not allowed to eat the cookies and were given radishes instead. The performance of these participants on an unrelated self-control task was inferior to two other control groups: those who were allowed to eat the cookies and those who were not tempted by cookies at all. The self-control task participants were given was an unsolvable geometry problem. Those who were subjected to the tempting cookies but were not allowed to eat them gave up more quickly on the geometry problem than the participants in the other two groups, who persisted much longer. The experimenters concluded that resisting the tempting cookies used up valuable resources which were also used during the unrelated persistence task.

Ego depletion has also been demonstrated to affect one's physical endurance. Subjects were asked to control their emotions while viewing a very sad film clip, while participants in the control group were instructed to watch the clip naturally. Afterwards, subjects were given a handgrip device and were asked to squeeze it for as long as they could. This task required self-

control in the form of confronting physical discomfort and resisting relaxing one's hand muscles, according to the experimenters. As in the previous experiment, those who were required to perform a prior act of self-control (in this case, controlling one's emotional state while watching a sad film clip) systematically gave up squeezing the handgrip device much sooner than those who had not been required to attempt to control their emotional states. Therefore, it seems as though the prior task of controlling one's emotional states in an effortful manner depleted resources needed in order to persist on a physical endurance task that is unrelated.

Not surprisingly, a similar experiment showed that controlling one's thoughts depleted resources such that subsequently one's capacity to control one's emotions was impaired. Subjects were instructed to suppress a forbidden thought, whereas those in the control group were asked to simply list their thoughts. After this initial task the subjects watched a comedic video clip and were instructed to suppress laughter and any signs of amusement. Here, as in the other experiments, the participants in the control group, those who did not attempt an exercise of self-control and thereby had available the resources to suppress laughter and other signs of amusement, were much more able to make successful attempts at regulating their behavior in the way instructed. Those who prior to the second task made an effortful attempt at self-control were much more likely to be unable to muster the resources in order to regulate their behavior.

This research led psychologists to consider what other behaviors may deplete the resources required for self-control. Baumeister speculated that since conscious decision making and self-control were both executive functions, it may be likely that conscious decision making would reduce the resources required to make effortful exercises of self-control. And this is precisely what Vohs et al. (2008) found. In one experiment, a group of undergraduates were divided into two groups: a choice condition and non-choice condition. Those who were in the choice condition were asked to make choices about the psychology class they were enrolled in that would affect the way the course was taught that semester and for future semesters. This was done by giving them a packet of material which they were instructed to read and then choose one of the options available. For example, one of the questions gave two descriptions of video clips, and the participants were to decide which clip they would rather watch in class as part of a lecture. Another question asked the students to choose between two different styles of test question. In total, the participants made 35 choices before the self-control task. Participants in

the no-choice condition were given the same packet of material but were instructed only to read the questions and not to make a choice regarding the questions.

Afterwards, both groups moved on to the subsequent self-control task they were told was unrelated. This task involved unsolvable geometric tracing problems in which participants were given a packet containing two geometric figures and a stack of blank paper. They were instructed to attempt to trace the geometric figures on the blank pieces of paper without ever lifting their pencils. They could attempt this task for as long as they wanted and could make as many attempts as they wanted. Their persistence on the task was measured and recorded. As with the other self-control studies, the experimenters found that those in the no-choice group tended to persist for much longer on the self-control task than those that were in the choice group. This study, along with other follow-up studies, led the experimenters to conclude that effortful decision making tends to deplete the resources required for self-control.

It's important to note that the experimenters' hypothesis was that it was choice itself that depletes self-control resources rather than associated processes such as deliberating or actually implementing the choice that has been made. On the face of it, it would seem as if deliberating would lead to a greater depletion of resources. After all, decision making is simply momentary whereas deliberating can take quite a bit of time. An analogy the experimenters use in order to illustrate their hypothesis involves a computer. A computer requires energy to do a bit of calculating which is similar to deliberating about what choice to make, but actually making a choice would be analogous to the computer writing the plan of action to perform in the future on a disc. In this light, according to the experimenters, actually making the decision uses up quite a bit of resources as well. In order to adequately test this hypothesis, another experiment was designed and executed.

In this experiment, participants were instructed to interact with the dell.com website, at which you can go through a series of steps to design your own desktop computer. Participants were broken up into three groups: the decision group, the deliberation group, and the implementation group. In the implementation group, participants were given instructions on which decisions to make at each step in the process of designing the computer and were asked to simply make the selections on the website that had been given to them in hardcopy. In the decision group, participants were instructed to carefully think about each available option at each step in the process and choose which one they most preferred, and then select that option. Those

in the deliberation condition were told to “form an opinion of the information, thinking about what [they] would prefer” (p. 892). However, they never actually made the selection on the computer screen. After the initial phase of the experiment, the participants moved on to the self-control phase in which they were given 80 five letter anagrams that were all solvable. They were instructed to attempt the anagrams until they solved them all, wanted to stop, or decided to give up. Their efforts on this task were timed by the experimenters.

The results of this experiment were in accordance with the experimenters’ hypothesis: those in the implementation and deliberation condition persisted much longer than those in the decision condition. And surprisingly, there seemed to be no difference between the persistence on the self-control task between those in the implementation condition compared to those in the deliberation condition. This experiment seems to show that actually making a decision depletes resources more so than implementing an already made decision and more so than just simply deliberating. However, it is worth noting that it isn’t clear that the participants in the deliberation group were actually deliberating. Their instructions were to just form an opinion of the information. It isn’t clear that the same cognitive capacity used in deliberating before making a decision would be used by participants in the deliberation group.

In Baumeister (2008b), instead of speaking in terms of conscious decision making, he speaks in terms of rational choices. He describes rational choice as “an evolutionarily new kind of decision making (new with humans) that relies on evaluating a potential course of action using logical reasoning, most commonly with cost-benefit analyses that mentally simulate the various behavioral options and their likely consequences” (p. 71). For my purposes here, it’s unproblematic to speak of conscious decision making and rational choice as synonymous.

A natural question that arises in light of these experiments is the following: what exactly are the resources that self-control relies on? A set of experiments designed to answer just this question found that the capacity for self-control has been associated with blood glucose levels (Gailliot et al. 2007). Gailliot et al. found that in self-control experiments similar to those in which Baumeister et al. found subjects’ capacity for self-control depleted, subjects’ blood glucose levels were significantly lowered as a result of effort put forth during the task. Blood glucose levels, therefore, served as an adequate predictor of future success on a subsequent task. In addition, Gailliot et al. found that when subjects with low glucose levels from exerting effort on a task consumed a glucose beverage the effects of ego depletion were not present. In short,

once the participants' glucose level was increased, they replenished the resources which their capacity for self-control relied on, and they were able to make more successful attempts at self-control even after exerting effort on a prior task.

According to Baumeister, the research just described serves as a basis and evidence for the idea that free will is "limited" and that it is a "sometimes thing." In explaining the usefulness of the capacity for self-control and the capacity for conscious decision making, Baumeister writes:

This line of thought fits the view of free will as a sometime thing. People are incompletely rational and self-controlled. They have the capacity for acting rationally and exerting self-control, but they only use it sometimes. This suggests the capacity is limited (2008a, 17).

Similarly, he claims that "[o]ur research on ego depletion provides one way to understand why free will is at best an occasional phenomenon" (2008a, 17).

But what does it mean for free will to be an occasional phenomenon or a sometimes thing? And what does it mean for free will to be limited? These claims require some elucidation. The block quote above indicates that the fact that we only exercise our rationality and self-control on occasion is evidence that free will is a sometimes thing. Not every action we perform is free because not every action of ours requires the exercise of self-control or even rationally deciding. On this view, free will is identical to exercises of self-control and rational decision making. But on the face of it, rational decision making is something we do quite frequently. Many times a day we are faced with a choice that requires cognitive effort and deliberation. Why, then, does Baumeister seem to think that free will is nonetheless an occasional phenomenon?

We can gain a better understanding of these claims by looking at Baumeister's view of free will in general. Baumeister (2008b) expands on this idea. He writes:

...the approach I am advocating is to compare different (forgive the expression) *degrees* of freedom and to learn what distinguishes relatively free from relatively unfree acts. Under that assumption, free will is at best a sometimes thing. Much, perhaps the



majority, of human action could be fully and simply determined by simple, explicable causal processes, including brain dynamics, reinforcement-based learning, and ingrained or acquired responses to stimuli. Free will would represent only an occasional opportunity to suspend or override those causal processes, so as to allow a different process to take control (2008b, 68).

At least two points about this quote need to be explained: the idea of degrees of freedom and the idea that the majority of our behavior is automatic. I'll begin with the second of these notions.

Why think that the majority of human behavior is automatic? Baumeister explains that the majority of human behavior is automatic in the sense that an "inner choice" is not being made (2008b, 68). Baumeister distinguishes between two different types of choice: an internal and an external view of choice. On the external view of choice different courses of action are technically possible but in this case "choice is defined by the situation" (2008b, 68). Baumeister seems to mean that even though there are different courses of action open to an agent, it is the situation that determines which of those actions is chosen. Internal choice, on the other hand, places an emphasis on the inner processes of making a choice. "If the person does not go through an inner process of choosing," he writes, "then the fact that in principle he or she could have done so is irrelevant" (2008b, 68). To illustrate, on Baumeister's view someone who sleeps with the same sexual partner or eats the same breakfast everyday is not making a choice. According to him, such behavior does not involve an inner process that can be called choice.<sup>34</sup> It would seem that Baumeister thinks the majority of human behavior is automatic because the majority of "choices" that humans make fall into the external choice category.

In addition, Baumeister's view seems to allow for some actions to be freer than others as indicated by the block quote above. Baumeister et al. (2008) claim that people experience some of their own actions as freer than others and we make similar distinctions while perceiving the actions of other people. As an example, he cites acting reluctantly at gunpoint as less free than following one's own carefully made plan. But Baumeister doesn't directly say what makes one

---

<sup>34</sup> Baumeister et al. (2008) claim that their understanding of free will is informed by research that distinguishes between System 1 and System 2 processing. System 2 processing is associated with slower deliberate thought whereas System 1 processing is effortless and automatic. Baumeister's picture is that most of our behavior is a result of System 1 processing and System 2 can occasionally step in and override System 1. This framework ties free will very closely to System 2 processing.

action freer than another. Given the example above, he seems to take actions done out of coercion to be less free than carefully planned actions. Presumably, the more carefully planned an action is the freer it is (insofar as it qualifies as an internal type of choice). But how does self-control relate to degrees of free will? Baumeister doesn't say explicitly, but one plausible explanation given the picture sketched thus far is that successful exercises of self-control performed when an agent's resources for self-control are fully stocked would be more free than a successful exercise of self-control when an agent has diminished resources (thus making his success less likely). Baumeister might seem to think this is so because in the latter case the agent's capacity for self-control is weaker. Additionally, successful exercises of self-control would be more free than unsuccessful exercises of self-control. And unsuccessful exercises of self-control would be more free than acting akratically without any attempt to exercise self-control. But this wouldn't mean that akratic actions are unfree, it would just mean that they are less free than actions involving attempted exercises of self-control. Akratic actions by definition are free actions. Understanding more and less free actions in this way would account for why we are incompletely self-controlled and thus are incompletely free. And it would also account for why our capacity to exercise self-control and thus free will is limited.<sup>35</sup>

### **4.3 Evaluating the Evidence**

Now that I have sketched Baumeister's account of free action and at least attempted to make sense of the claims he makes regarding free will, I'm in a better position to evaluate those claims. While my main concern is to scrutinize his claims which appeal to the research on self-control, I will begin by making a few remarks with respect to his view of free will in general. Then I will address the issue of what we can infer about our free will from the experiments and research involving self-control and conscious decision making.

---

<sup>35</sup> This is just a sketch of a view that attempts to make sense of how Baumeister might understand self-control and degrees of freedom. There are still some puzzles that would remain though. For example, does having a high degree of capacity for self-control make one's act free even if it is not used on some occasion? And would an act of this sort be freer than the act of an agent who behaves in the same way but who has a lesser degree of capacity for self-control?

### 4.3.1

One important thing to note about Baumeister's identifying free will with self-controlled actions and the ability to override impulses is how much this differs from philosophical conceptions of acting freely. As noted in section 3.1, self-control is an important ability which is required for acting freely, but it is hardly regarded as constituting the ability to act freely. This is so even among philosophers who situate free action in the natural order. One prima facie piece of evidence against the view that free actions are identical to self-controlled actions is our intuitions regarding actions for which we are morally responsible. Many philosophers take free will to just be the control required for moral responsibility. If free will were identical to self-controlled actions and actions which override impulses then we could only be directly morally responsible for self-controlled actions and actions which override impulses.<sup>36</sup> But this narrows the spectrum of actions for which we are morally responsible quite a bit from what we intuitively think, and this, I take it, counts against Baumeister's view of free will as identical to self-controlled actions.

Perhaps one may want to say that we are directly free only when we exercise self-control. Other actions we perform could still be free if they are appropriately connected to directly free actions, but only exercises of self-control are directly free. This may be an improvement on Baumeister's view but it still seems to implausibly narrow the scope of actions we think of as free actions too much.

A second piece of evidence that I think speaks against this identification of free actions with self-controlled actions is that in ordinary circumstances<sup>37</sup> the intentional actions we perform are often seen as free actions. Mele (1995) expresses something similar. He writes, "Autonomous A-ing (normally) is intentional A-ing" (p. 11).<sup>38</sup> But not all intentional actions are actions in which self-control is exercised or in which an impulse or temptation is resisted. So, on Baumeister's view, a wide range of behavior, including many intentional actions, would not

---

<sup>36</sup> We may still be able to be morally responsible but in an indirect sense, for example, for actions which result from free actions we performed in the past. This could occur via a type of tracing. See Fischer and Ravizza (2000).

<sup>37</sup> By ordinary circumstances I mean circumstances not involving manipulation, coercion, compulsion, and irresistible desires. The latter kinds of circumstances may disallow intentional actions from sufficing as free actions.

<sup>38</sup> In a footnote to this claim Mele describes a case in which one can coherently claim that the agent at issue acted autonomously but not intentionally. However, this seems not to be a normal situation. See Mele (1995) p. 11, N11. Also, autonomous action for Mele is synonymous with free action.

count as free actions despite the fact that we ordinarily think of intentional action as free action and we ordinarily hold each other morally responsible for our intentional actions.

Another contentious aspect of Baumeister's view of free will, and in particular his view of free will as occasional, has to do with its roots in what Baumeister calls the difference between internal and external choice (2008b, 68). Recall, Baumeister makes a distinction between external choice, which allows for different courses of action but is "defined by the situation," and internal choice, in which the inner process is "emphasized" (2008b, 68). Baumeister doesn't go into much detail explaining these two notions of choice, but he does illustrate them with examples. For instance, eating the same thing for breakfast everyday doesn't count as making an internal choice on this view because the choice is "defined by the situation" and, according to Baumeister, does not involve an internal process. And on Baumeister's view, "[f]ree will is relevant only when there is an actual inner process" (2008b, 69). But this seems plainly false. As mentioned earlier, a common understanding of phenomena such as choice or decision is that they are momentary actions of intention formation.<sup>39</sup> Forming an intention, loosely speaking, is simply forming a plan. And we form plans all the time even when it's the same plan we formed the day before and the day before that. This still seems to be a process that we engage in. Habitual actions, however, might not involve choice or active intention formation. Perhaps there is a decision only when practical uncertainty is settled. And since habitual actions don't involve practical uncertainty, they don't require a decision. But habitual actions could involve non-actively acquiring an intention in virtue of having a standing intention or some sort of plan or policy. But even if this is the case, there still seems to be an internal process occurring, and the agent's behavior isn't merely defined by the situation.

Something that might speak in favor of Baumeister's view of external choice is over-learned behavior. Take, as an example, learning to type on a computer keyboard. Before having acquired the skill of typing on a computer keyboard we have to practice slowly and make each keystroke with effort until we have become so accustomed to this task that it is effortless. While we were learning to type we may have had to form intentions to place our finger here or there, but after acquiring the skill it seems that we don't have to do any intention forming about finger movements at all. Perhaps Baumeister is thinking that actions performed everyday (like eating the same breakfast) are like acquired skills.

---

<sup>39</sup> Mele (2003b)

This does seem to be one possible understanding of Baumeister's view of choice, but it doesn't mean that free will is only relevant when there is an internal choice. For one thing, even with over-learned behavior like typing, there is still an inner process, some intention at work guiding one's behavior e.g., the intention to type a particular sentence. So even if you don't have to put as much effort into the individual keystrokes as you did when you were learning to type, you are still performing an effortful action in the scope of which individual keystrokes fall. Secondly, picking out and eating breakfast is a complex action that requires all sorts of bio-sensory feedback. Typically, we form a plan which requires some cognitive effort, but even if we don't form a plan we could still non-actively acquire an intention. This could be done by having a standing intention to eat a certain breakfast each morning or some sort of policy. Even if this is the case, this is still an internal process. And this internal process is one which characterizes a lot of our behavior, including that for which we are morally responsible. Given this, it doesn't seem warranted to discount such behavior, at least on the grounds Baumeister is suggesting, as unfree behavior or behavior that is not relevant to free will. If what I have been arguing is correct, then free actions are plausibly more frequent and more common than Baumeister has acknowledged.

#### **4.3.2**

What about Baumeister's research program involving self-control and conscious decision making? What claims is this research aimed at supporting and does it in fact support them? The main result that this research program has demonstrated is that activities that involve exercises of self-control and conscious decision making deplete resources. Once these resources have been depleted, our capacity for self-control is impaired such that future attempts at exercising self-control are much more likely to be unsuccessful. Baumeister's studies seem to adequately support these claims. But he concludes from them that one might suggest that, when our resources are depleted we are less free than we would be were we to retain our resources.

There seems to be good reason to doubt that this follows. For one thing, the experiments which conducted research on self-control capacities only generate a probabilistic result: exercises of self-control are much more likely to be unsuccessful or self-control is less likely to be exercised. In other words, even when one's resources may be depleted, this doesn't guarantee that a subsequent exercise of self-control will be unsuccessful, only that the likelihood that it will

be unsuccessful goes up or that the likelihood that an agent will exercise self-control goes down. In any case, it is possible for one's resources to have been depleted and one still make a successful exercise of self-control. Consider the following scenario. Jake has had a rough day at work. He spent a lot of time during the day resisting the urge to express his anger at coworkers that had been slacking off, and though he was constantly tempted by the unhealthy desserts that were left in the break room for everyone to enjoy, Jake resisted indulging in them and resolved to stick to his diet. Upon coming home from work, his resources depleted, we might be able to infer that Jake is less likely to exercise his self-control and overcome his temptation to relax on the couch rather than head out for a jog as he had previously planned to do. But does this mean that any hope Jake has at resisting his temptation is doomed to fail? Does this mean that without precious resources the capacity for self-control relies on Jake can't make a successful attempt to override his impulses to forgo jogging? No, not at all. It could be true that Jake is more likely to make an unsuccessful attempt at exercising self-control yet he still successfully overcomes his temptation. Nothing in the experiments indicates that Jake can't resist his temptation.

Let's say that Jake does in fact make an unsuccessful attempt at resisting his temptation. Jake gives in to his urge to relax on the couch instead of heading out for a jog. What are we to say about Jake's behavior? Was it less free? It's not clear that this is the case. If Jake had made a decisive better judgment to the effect that it would be better to go jogging than to relax on the couch yet he decided to relax on the couch anyway, his behavior might be akratic. Jake might suffer from weakness of will in this respect. However, one characteristic feature of akratic behavior is that it is free behavior.<sup>40</sup> Jake's weakness of will must be a free action in the sense that it was possible for him to overcome his temptation. If it were not possible then Jake's behavior would be compelled; he would have no other alternative but to do what he was most motivated to do. So if Jake's acting akratically in this respect is less free than a successful exercise of self-control would be, it's not clear that it is less free in a way that would make him not morally responsible.

Perhaps one might suggest that when our resources are depleted and our capacity for self-control is impaired, in the sense that we are less likely to make a successful attempt at self-control, and we are less likely to make rational decisions, trying to exercise self-control or attempting to make a decision is less free than it otherwise would be. This is an important

---

<sup>40</sup> Mele (1987, 4).

suggestion, but it's not clear what the argument for it is (or would be). The thought seems to be that since the capacity for self-control and conscious decision making are necessary for the ability to act freely, and the capacities for self-control and conscious decision making are impaired in these circumstances, so too is our ability to act free impaired. However, this relies on the assumption that the degree to which one possess the capacities which constitute free will indicates the degree to which one is free. This assumption isn't obviously true. One thing it does indicate is that it's unclear to what extent one must possess the capacities for self-control and conscious decision making in order to have the ability to act freely. Many free will theorists simply require the general capacity and do not specify to what degree. One possibility is that free will is a threshold concept: some degree of capacity is required, and an agent is either above the relevant degree or not.<sup>41</sup> But in any case, it seems that the general capacities for self-control and conscious decision making, even when the resources they rely on are depleted, are not impaired to such an extent that we are rendered completely helpless and entirely unfree. This can be seen by focusing on certain aspects of the self-control experiments.

An important feature of the self-control experiments which restricts what we can say about self-control in general involves the nature of the self-control tasks at issue. For example, some of the self-control tasks are puzzles which are literally impossible to solve. What the experimenters measured is how long one persists on such a task. The result they found was that those whose resources had not been depleted persisted longer than those whose resources had been depleted. But it is unclear how this conclusion would generalize to all exercises of self-control. Not every exercise of self-control requires sheer persistence in the face of frustration. Take the above example of Jake. He was tempted to skip his jog and relax on the couch, but it's not obvious that just because his resources were depleted, his attempt at exercising self-control would be less likely to succeed because he wouldn't be able to persist in the face of his temptation. Jake could resist his temptation to relax by forming an intention to put on his running shoes and workout clothes now and, by brute force, overcome his countervailing motivation. Exercising self-control for Jake may not require the type of persistence those in the depleted self-control task seemed to lack.

Similarly, some of the self-control tasks that those with depleted resources did worse on involved a series of problems which they took longer to complete than their non-depleted

---

<sup>41</sup> Thanks to Randy Clarke for making this point clear to me.

counterparts. For example, in one of the tasks discussed above, depleted and non-depleted participants were instructed to attempt 80 five-letter anagrams. The results reported by the experimenters showed that those in the depleted group took much longer on the task. Although it's true that some of the participants who were depleted did not complete the task, it's also true that some of the depleted participants did complete the task but, on average, took longer than their non-depleted counterparts. This finding doesn't even suggest that attempts to exercise self-control would be less successful, just that they would take longer.

Another important feature of the experiments which restricts what we can conclude about our capacities for self-control is the fact that the self-control experiments only seem to measure brute resistance of temptation. Brute resistance of temptation is one way in which someone can exercise self-control. But self-control can be exercised in a plethora of other ways. One set of ways called cognitive attentional strategies has been found to be very successful in resisting temptation.<sup>42</sup> Such a technique involves focusing one's attention away from the tempting object and in favor of the object judged to be more valuable in the long run. By focusing your attention in this way you can actually shift your motivational set such that you are no longer more motivated to act against your better judgment. Another technique which may be familiar is simply anticipating when one will become tempted. By doing so, agents may alter their behavior so that they will not be tempted in the future and, therefore, will not have to attempt an exercise of brute resistance. However, the self-control experiments discussed earlier do not speak to such self-control techniques. Why is this important? Consider the following example. Jenny, who has resolved to refrain from snacks in between meals, often finds herself tempted between classes to eat junk food out of the vending machines she frequently passes. On several occasions she has given in and purchased a candy bar. It seems that when she is in the presence of a vending machine her attention is constantly returning to buying some sort of snack. Jenny figures out that she can change her behavior so she will no longer be tempted throughout the course of the day. She decides that she will take a different route to and from her classes so she doesn't pass by a vending machine. Doing so, she reasons, will ensure that her attention is not consumed with purchasing a snack. As it turns out, this trick works for Jenny, and she successfully exercises self-control even though if she were to pass a vending machine she would

---

<sup>42</sup> See Eisenberg et al. (2004) for a review of the literature.



be unable to resist the temptation. She can exercise self-control in a way other than brute resistance.

If a story of this sort is coherent, as it seems to be, then it's not far-fetched to think that even when one's resources for the capacity for self-control in the form of brute resistance are depleted, one may still exercise self-control via some other techniques. In other words, our prospects for exercising self-control, even when our resources are depleted, are not as dim as Baumeister has characterized them to be. However, it might be that depleted agents employ these techniques (or are able to employ these techniques) less often. Perhaps being depleted of self-control resources makes it less likely that an agent will look for and successfully employ techniques to exercise self-control. Whether or not this is actually the case would be an interesting question for future research in this area.<sup>43</sup>

#### **4.4 Concluding Remarks**

So what can we say about the experiments? What implications do they have for free will? What the experiments do show is that our capacity to exercise self-control in the form of brute resistance relies on a limited resource which may become depleted. Once depleted, the likelihood of our making a successful exercise of self-control is diminished. However, it's not at all obvious that the attenuation of this capacity results in any attenuation of our ability to act freely. It's far from clear that the extent to which our resources for this ability are present mirrors the extent to which we have free will.

The experiments on self-control also give us valuable information about why we behave the way we do and how we can exercise more control over our behavior. One example of a phenomenon that ego depletion can potentially explain is the pattern of judicial decisions made throughout the course of a day. Shai Danziger et al. (2010) found that the odds that a prisoner will be successfully paroled at the beginning of the day start off high around 65%. This number starts to drop off over the course of a few hours and then spikes back up to 65% after lunch. Then, once again, the numbers start to dwindle. An explanation for this effect seems to be the glucose levels of the parole board members. When their blood glucose level starts to drop, they

---

<sup>43</sup> I thank Randy Clarke for pointing this out to me.

start spending less time reviewing the individual prisoners and rely on a default judgment of not granting parole. Understanding effects of this sort can help us gain a greater degree of control over our lives. For example, if we understand the phenomenon of ego depletion and its effects on our behavior, we can better arm ourselves to counteract it. Recall, in one of the self-control studies Baumeister et al. reported that the effects of ego depletion could be countered by something as simple as drinking a glass of lemonade.

Now, to be fair, Baumeister's stated aim was to avoid philosophical and metaphysical disputes about the existence of free will. Rather, he claimed, he wanted to attempt to understand free will from the point of view of a psychologist. That said, he took as his starting point that there is a common distinction in folk psychology between free and unfree (or less free) acts, and Baumeister wanted to sort out the difference between these acts. The experiments discussed in this chapter do give us valuable information but I don't think—as I've argued—the self-control experiments warrant Baumeister's skeptical claims about free will.

## CHAPTER FIVE

### FREE WILL AND AUTOMATICITY

In the psychological literature, automaticity generally characterizes mental processes that “operate outside of conscious awareness and even intention,” though, as we will see, the term is used in a number of different ways (Bargh, 2008, p. 128). Psychological research also reveals automaticity to be a ubiquitous feature of human behavior. John Bargh and Tanya Chartrand (1999) speak of not only the “unbearable automaticity of being” but argue that automatic processes perform “the lion’s share of the self-regulatory burden” (p. 462). In fact, many psychologists are convinced that automaticity is such a pervasive feature of human behavior that there is little (if any) room left for conscious control or free will.

The aim of the present chapter is to argue that such convictions are supported on neither empirical nor conceptual grounds. The first step towards this conclusion requires that we more carefully scrutinize the notion of automaticity itself. In section 4.1, I do just this, and although I do not provide a definitive analysis of automaticity, I do distill the hallmark features of automaticity which are most likely the potential threats to our freedom and control. Section 4.2 evaluates the extent to which automatic processes entail a lack of control, and section 4.3 tackles two remaining arguments which make use of the literature on automaticity in order to claim that the freedom and control we have is limited.

#### 5.1 What is Automaticity?

My attempt at answering the title question for this section may seem unsatisfactory to some. I will not be attempting to give an analysis of automaticity and I will not be attempting to give a definition of automaticity which applies to all of its usages. My goal is simply to give a brief understanding of the term ‘automaticity’ by identifying several key features that many have cited as being a threat to freedom and control and then evaluating whether such features are genuine threats.

Characterizing automaticity can be tricky business and has been the source of frustration for many psychologists. This could be because there are many types of automatic things going

on in the human body. Our hearts beat automatically, our eyelids may automatically blink, and our bodies automatically regulate temperature by sweating. In addition, we can perform complex behaviors automatically when we do something without reflection like catching a ball that has been thrown at us when not quite ready. I will be primarily concerned with the automatic things that we do in this latter sense and automatic processes that are related to actions we perform. It seems relatively unproblematic as far as our freedom is concerned that much of our autonomic functioning is automatic.

A popular method for characterizing automatic processes is by juxtaposing them with controlled mental processes—those processes that seem to be careful and deliberate processes that we consciously control like decision making and planning, for example, deciding which college to attend and planning a family vacation. Opting for this strategy, John Bargh (1994) identifies what he calls “the four horsemen of automaticity.” Controlled processes like deliberate decision making have the qualities of awareness, intentionality, being effortful and requiring attentional resources, and being controllable. Automatic processes, on the other hand, have the opposite of these qualities. That is, automatic processes are processes of which we are unaware, are non-intentional, are effortless and don’t require attentional resources (i.e., are efficient), and are not controllable.

It’s also important to note that, contrary to what used to be the consensus view, mental processes are not simply either automatic or controlled.<sup>44</sup> There are some automatic processes that possess some of the features of automatic processes listed above but also possess some of the features of controlled processes listed above. In this sense, mental processes lie on a continuum between automatic and controlled processes. Consider the skills involved in driving a car. The driver must pay attention to the road and other traffic while also monitoring her speed and direction of travel. For the experienced driver, these processes become automatic in the sense that they are effortless for her, require little of her attention, and generally operate outside the scope of her conscious awareness. However, these processes do seem to be controlled processes in the sense that the driver can become aware of them and modify their operation if need be. If she notices that she is driving above the speed limit she may reduce the amount of pressure she is putting on the gas pedal. And though it may be contentious as to whether these individual processes are intentional it seems clear that the driver (and most drivers in normal

---

<sup>44</sup> Bargh (1994) notes that Johnson & Hasher (1987) considered this the consensus view.

circumstances) has the intention to drive and drive in a certain way by making use of those processes. The lesson of this is that rarely do we find mental processes which have all four qualities associated with automaticity. In fact, Bargh and Chartrand (1999) note that there are two major strains of research on automatic processes that are similar only in that the processes identified do not possess all of the four qualities. The first body of research is on skill acquisition which focuses “on intentional, goal-directed processes that [become] more efficient over time and practice until they [can] operate without conscious guidance” (p. 463). The second body of research has to do with “perceptual analysis or encoding of environmental events” (p. 463). One particular experiment which demonstrates this body of research has shown that threatening or emotion-laden words or symbols were shown to be “defended against” through having higher perceptual thresholds than more neutral stimuli.<sup>45</sup>

In addition to the fact that processes can have some of the features associated with controlled processes and some of the features associated with automatic processes, all four features themselves can come in degrees.<sup>46</sup> Some processes may lie so far outside our conscious awareness that it is impossible for us to bring their operation into our conscious thoughts, whereas others are easily brought into our conscious awareness. Similarly for the other three features: efficiency, intentionality, and controllability.

## **5.2 Automaticity and Lack of Control**

As mentioned above, one of the features that generally characterize at least some automatic processes is that they are uncontrollable or they entail a lack of control. Many of our mental processes are said to be automatic in nature. If this is the case and it's true that automatic processes entail a lack of control, then many of our mental processes would appear to be not controlled. This fact might affect our prospects for acting freely at least in the case of mental actions like decisions. But what does it mean to say that automatic processes entail the absence of control, and would this be as big a threat to free will as it seems? In what follows I will attempt to answer these two questions.

---

<sup>45</sup> Bargh and Chartrand (1999) cite Allport (1955) and Erdelyi (1974) in this connection.

<sup>46</sup> Moors and De Houwer (2007) make this point as well.

What does it mean for an automatic process to entail a lack of control? One thing I would like to caution against is thinking of control as an all-or-nothing thing. It should not be surprising that there are many different types of control and many different ways that things can be controlled. Consider a thermostat. There is a sense in which a thermostat controls the temperature of a room. The thermostat operates by a mechanism that detects the air temperature in the room, and if the temperature reaches a certain point the thermostat kicks on the air conditioner or the heater. Once the temperature of the room falls below or rises above a designated point the air conditioner or the heater shut off. The thermostat controls the room's air temperature. This is a type of low-level control that might be similar to the way automatic processes operate. Just as the thermostat monitors the room's temperature and has a plan to regulate that temperature if it "recognizes" a change in temperature, our brains are equipped with regulating mechanisms and processes that can similarly kick in if they "recognize" a change. Maintaining one's balance while riding a bike can partly be regulated by automatic processes in the same low-level way that the temperature regulates the room's temperature.

So what type of control is lacking in automatic processes? With respect to at least some automatic processes the type of control that seems to be lacking is conscious deliberate control. Exercising conscious and deliberate control involves carefully and consciously guiding your thoughts and bodily movements in a way consistent with intentions you may have. But an action's simply lacking conscious control or conscious guidance doesn't seem to be enough to threaten our freedom and agency. Some actions that aren't consciously controlled are still intentional actions. What it means for an action to be intentional has been the subject of much dispute in the philosophy of action. But by assuming a causal theory of agency we can say that what makes an action intentional is that the action was caused, guided, and sustained by an agent's intention in the right sort of way.<sup>47</sup> And as I mentioned above, not all automatic processes seem to completely lack this feature of intentionality. In particular, learned automatic processes or skilled behavior that incorporates automatic processes seem to have aspects of intentional control. Think about an expert pianist. The pianist reads a few bars on a sheet of music and then moves her fingers on the piano in the right sort of way to generate the notes the way they are represented on the page. The pianist (if she has learned to play well) does not need to carefully or deliberately consider each note in order to figure out which key that note

---

<sup>47</sup> Just exactly what the right sort of way is has been the subject of much interest to philosophers.

corresponds to. She can automatically read the note and then press the proper key in a fluid manner that demands almost no attention dedicated to her precise finger movements. Cognitively she uses few resources and the automatic process operates smoothly outside of conscious awareness: initiating upon reading the music and continuing to completion. Though the pianist isn't consciously intending each movement of her finger, I think it's fair to say that she is nonetheless intentionally playing the piano.<sup>48</sup> If the pianist is intending to play music and this intention causally plays the role of guiding and sustaining her behavior, then what she is doing is intentional. What she is doing is exercising her well-developed skill at playing the piano and not merely accidentally or randomly hitting the correct notes.

Now we must consider the second strain of automatic processes discussed above—the automatic processes that seem to lack both conscious and intentional control. In particular, we must ask whether processes that lack both conscious and intentional control are threatening to our freedom and agency. In what follows I will suggest two reasons to think that just lacking these two types of control in and of themselves is not threatening to our freedom and agency. Rather, I will suggest that the types of automatic processes that do seem threatening appear so because they interfere with processes that are more controlled.

The types of processes that Bargh and Chartrand (1999) say fall into the second strain of research differ from skilled and learned behavior in that they are more like cognitive reflexes. Some of these processes we have innately and are born with whereas others are developed from experience at a very young age. Automatic processes of the first sort would include processes that bring particular visual stimuli to our conscious attention, automatic processes that regulate bodily functions, and automatic processes involved in sensory perception. Automatic processes of the second sort include automatic processes that generate our emotions, mood, and temperament, information processing like stereotypes, reading, and evaluation. With respect to all of these automatic processes, I think there is good reason to think it's not problematic that they are neither consciously nor intentionally controlled.

The idea that we in fact have two types of cognitive processes is often referred to as dual-processing theory. Schneider and Chein (2003) argue for such a view. They note that not only

---

<sup>48</sup> Skilled activity has been seen as a threat to not just free will but to the causal theory of action as well. See Clarke (2010) and Adams (2010) for attempts to neutralize the threat. Clarke and Adams each make use of motor schemata in attempting to show that motor intentions involved in behaviors like moving one's hand a certain way can be intentional in virtue of being caused by a higher level intention at some earlier time.

would two processing mechanisms not have evolved if it weren't advantageous to our survival, but we can actually see the advantages by looking at how automatic and controlled processes complement each other. For example, Schneider and Shiffrin (1977) were able to demonstrate how controlled processes are able to contribute to fast learning of a particular search task. Subjects in their experiments were able to learn very quickly to search for a designated set of letters among a set of letters displayed on a screen. Automatic processing complements this by not only operating much more quickly than controlled processing but by being much more robust in the sense that the process is much more resilient under stress. For example, automatic processing has been shown to be quite resilient under the effects of alcohol, fatigue, stress, and vigilance.<sup>49</sup> By having two types processes with complementary features—one type of process that operates well with little input from consciousness and one that operates well with a greater degree of conscious input—we are able to do many more things than we otherwise could. This is not that novel of an idea and is actually quite simple. In the same way that labor is divided up among many members of a business, so too does our brain divide up the labor among its many modules.

The second reason I would like to suggest is that even though automatic processes may not require conscious or intentional control to initiate or sustain the process, this doesn't mean that automatic processes aren't able to be consciously or intentionally controlled. In fact, there is a wealth of evidence that indicates automatic processes can be controlled in a variety of ways. Consider the cognitive process of stereotyping understood simply as categorizing information. The automaticity of stereotyping was demonstrated by Gaertner and McLaughlin (1983) showing that subjects were faster to identify "paired letter strings if they were consistent rather than inconsistent with the stereotype of Black Americans (e.g., Blacks—lazy vs. Blacks—ambitious)." Patricia Devine (1989) was able to go further and show that even subliminally presented stimuli could activate stereotypes, and Irene Blair (2002) reports that over 100 studies have documented that "Whites have automatic negative associations with Blacks (or other non-White groups), young adults have automatic negative associations with the elderly, and both men and women automatically associate males and females...with stereotypic attributes" (p. 242). But even though early understandings of automatic processes like stereotyping had it that such processes were inevitable, impossible to avoid, and uncontrollable, there has been a vast amount

---

<sup>49</sup> See Schneider and Chein (2003) for a review of this literature.



of data that demonstrates a variety of ways that automatic processes like stereotyping can be controlled.

John Bargh (1994) discusses two types of motivation for controlling automatic processes like stereotypes: externally generated motivations and internally generated motivations. Externally generated motivations or situationally induced motivations occur often when experimenters instruct subjects to perform some cognitive task or to think about a task a particular way. Psychologists have found that attention to stereotype-inconsistent information increases when there is outcome dependency—when the subjects have power or control over important outcomes of the people they are making judgments about. An example would be when subjects are told that their input with respect to potential job candidates will be used by a hiring committee. In such cases, people would tend to be more likely to refrain from forming negative judgments that would otherwise result from stereotype automatic processing.

The second type of motivation for controlling automatic processes is internally generated motivations. One of the major motivations that psychologists have found that prompts individuals to control automatic stereotyping is the internal motivation to not appear racist or prejudiced. Devine (1989) showed that even when people are likely to automatically stereotype, under some conditions such stereotyping can be internally controlled. She showed this to be the case by having subjects consciously and deliberately write out ‘thought lists’—lists of associated words that were supposed to correspond to a category. Even though subjects stereotyped a group in previous tasks, during the thought list task they were given the opportunity to override the influence of such automatic stereotyping, and they took advantage of it by inhibiting the activated stereotype. Further research reported by Blair (2002) indicates that automatic stereotypes can be inhibited by several other things: self and social motives, specific strategies, the perceiver’s focus of attention, and the configuration of stimulus cues.<sup>50</sup>

Thus far, the type of control I have considered with respect to stereotyping has focused primarily on controlling automatic processes in a corrective nature. That is, the automatic process of stereotyping is initiated but the subject is able to control the extent to which it shapes the judgment that is reported or given. But there is another form of stereotype control referred to as preconscious control. Moskowitz et al. (1999) showed that there are ways one can control whether the automatic process of stereotyping initiates in the first place. In particular, the

---

<sup>50</sup> Blair, Judd, & Fallman (2004) were also able to show the controllability of racial stereotypes.

experimenters found that participants with egalitarian goals—goals involving fairness to individuals—were able to suppress even the initiation of automatic processes involving stereotypes.

So far what I have said has simply been in the service of demonstrating that there are reasons to think that the mere fact that an automatic process isn't consciously or intentionally controlled doesn't show that our prospects of agential control and acting freely are doomed. In particular, I suggested there were advantages to having a dual process system and that there are actually several ways in which processes that generally operate without conscious and intentional control can become controlled. They can be controlled in a corrective nature, and their initiation can even be controlled. So even if a particular automatic process like stereotyping operates outside of our conscious and intentional control we can exercise control to rein it in, modify it, retrain it, and re-automatize it. The lesson I think we can derive from paying attention to this kind of feature is that it's not the lack of conscious or intentional control that appears to be problematic about automatic processes. Automatic processes appear problematic when they seem to interfere with controlled processes. When the automatic process of stereotyping interferes with controlled processes in the service of judging people equally the control we exercise seems to be threatened. The automatic process doesn't diminish our ability to exercise control, but the control we are exercising seems to be diminished.

An experimental result that demonstrates this pretty well is the now famous Stroop effect named after John Ridley Stroop.<sup>51</sup> In these eponymous experiments participants were presented with a particular color word that is written in a color different than that referenced by the written word and asked to name the color of ink the word is written in. Subjects in the experiment cannot help but process the meaning of the word, and this interferes with the task of naming the color the word is written in as quickly as possible. Subjects are much slower to report the color the word is written in compared to when the word is written in the same color that is referenced. When the meaning of the color word and the color the word is written in are the same, there is no interference. For example, when the word "RED" is written in blue ink, the subjects consistently take longer to report that it is written in blue ink than they take to report the color when the word "RED" is written in red ink. This is the Stroop effect. It is the seemingly uncontrollable nature

---

<sup>51</sup> Kihlstrom (2008) claims that the modern notion of automaticity was inspired by the Stroop effect and related experiments.

of the mental process and the fact that its execution appears to be initiated and then proceeds without the subjects' exerting any effort that stand out as remarkable features. But as I've argued above, I don't think this is the source of the threat to our agency and control. In principle, we might be able to re-train the automatic process involved in reading the word presented to us such that we become faster and better at the Stroop task. Automatic processes seem to be controllable in this sort of way. Rather, the Stroop task seems to demonstrate a lack of control on the subjects' part because processing the meaning of the word presented interferes with other goals and intentions the subjects have, namely, to report the color the presented word appears in.

The Stroop effect and related experimental effects may seem like uninteresting or contrived examples of a lack of control. But consider a real-life example of when automatic processes often interfere with and override other controlled processes operating in the service of a particular goal or intention. I have in mind here the vast literature on cognitive heuristics and biases—a research program begun by Nobel Laureate Daniel Kahneman and research partner Amos Tversky.<sup>52</sup> One bias unearthed by Kahneman and Tversky is the confirmation bias. Nickerson (1998) explains the confirmation bias as the “seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand” (p. 175.). When you are asked “Do you enjoy your job?” you often tend to search for instances or evidence (including memory) that is likely to confirm the belief you already hold. Having a negative attitude toward your workplace may bring to mind those features of your workplace that you dislike, not only prompting a negative answer to the question asked but strengthening your pre-existing attitude.

And the confirmation bias goes well beyond even this. Even when we do not have any pre-existing beliefs we are prone to seek out confirming evidence for a hypothesis rather than disconfirming evidence. What's now called the Wason card task illustrates this nicely.<sup>53</sup> In this type of experiment described by Gilovich (1991), subjects were given four cards each of which has a letter or a number on the side facing up, say, A, B, 2, 3. The participants were then told that each card had a letter on one side and a number on the other. They were asked to determine whether “all cards with a vowel on one side have an even number on the other” by “judiciously

---

<sup>52</sup> Kahneman's *Thinking, Fast and Slow* (2011) does a great job explaining much of this literature. Also, Amos Tversky would have shared to Nobel Prize with Kahneman but he died before he could be awarded.

<sup>53</sup> See Wason (1968).

turning over the proper cards” (p. 33). Subjects often respond by claiming the A and 2 cards ought to be turned over. But a moment’s reflection reveals that the 2 card only offers confirming evidence of the hypothesis: a vowel on the other side confirms the hypothesis whereas a consonant would still be consistent with the hypothesis. Turning over cards A and 3 does a much better job at testing the hypothesis. By turning the 3 card over we are able to potentially disconfirm whether all cards with a vowel on one side have an even number on the other. What this experiment does so well at showing is that even when the hypothesis is one which we aren’t particularly motivated to uphold or in which we don’t have a vested interest we still tend to seek out confirming evidence.

How does this sort of cognitive bias threaten control? If you have the goal or aim to form beliefs which are best supported by the evidence, then cognitive biases like the confirmation bias will tend to thwart that goal. Rather than forming beliefs based on all the evidence, beliefs will be formed based on pre-existing attitudes and the tendency to seek out confirming evidence rather than disconfirming evidence. We are less likely to satisfy our goals carried out by more controlled processes because the automatic process’s operation disturbs them. If this is true, then attention to cognitive biases can help us exercise more control over our attitudes and judgments. By being aware of the confirmation (and related) biases and attempting to counteract and correct them we can increase the amount of control we exercise. Our lack of attention to them may result in exercising less control over our behavior. Corroboration of this comes from a recent article in which Mele and Shepherd (2013) report evidence that knowledge about non-conscious processes that influence our behavior can be mitigated by activating more explicit processes. The authors describe studies in which participants who previously took an implicit association test—a test that measures attitudes that people possess but are not often aware of—were able to control the influence the implicit attitudes had on them. A similar result has been shown about the confirmation bias. Lord et al. (1984) showed that instances in which subjects were likely to be subject to the confirmation bias could be avoided by explicit instruction (to consider alternatives) and through “stimulus materials that made opposite possibilities more salient” (p. 1231). Essentially, these experiments were able to show that agents can exercise more control over processes that might otherwise interfere with goals and plans.

To summarize, I do think that automatic processes that interfere with other controlled processes or even other goals and intentions can indicate a lack of control for an agent even

though it might not indicate a lack of an ability to exercise control. On the other hand, I don't think this is so problematic that it warrants global skepticism about our prospects for acting freely. Much of what I have said about automatic processes that lack conscious and intentional control applies to automatic processes which interfere with goals and intentions. The benefits and advantages of having dual-processing systems allow for us to exercise a greater amount of control than we otherwise could. We can learn to process some information automatically and save valuable resources for processing less familiar information. However, there are occasions in which automatic processes may carry out a task resulting in an undesirable consequence—e.g., a false belief or unsupported judgment. But this isn't the norm for automatic processes—not all of them have this feature. And those that do can often be harnessed. We can learn the situations in which some automatic processes can interfere with standing principles and plans and avoid them or proceed more cautiously. This is good news for our prospects of agency in general, especially if automatic processes are prevalent. Even if this is the case, we need not be subject to them as something that just happens to us.

### **5.3 The Threat of Global Automaticity**

The previous section considered the threat of automaticity to free will based on particular features of automatic processes: lack of conscious control, lack of intentional control, and interference with goals and intentions. In this section, I would like to consider an argument against free will based on the prevalence of automaticity. This argument is most commonly made by John Bargh, though others have sketched similar arguments as well.<sup>54</sup> In broad outline, Bargh (1999; 2008) argues that because automatic processes which lie outside our conscious awareness can explain human behavior, there is no need to posit the existence of free will. There are at least two mistakes I think Bargh makes in his argument against free will. The first is that he is too quick to rule out a causal role for consciousness. The second is that he takes free will to be supernatural and doesn't allow for a more natural understanding of a scientifically friendly account of agential freedom and control. I'll discuss each of these problems in what follows.

---

<sup>54</sup> See Wegner (2002; 2005).

First, I would like to say that Bargh does a great job giving an impressive amount of experimental evidence to demonstrate the pervasiveness of automatic processes and the substantial role they play in our behavior. He does this by providing evidence based on different levels of influence. At the first level there is genetic influence “through which genes drive our present-day behavior through evolved motives” (p. 134). In essence, Bargh says, evolution has shaped us to be “open-ended systems” by which genetic influence can find expression (p. 143). Next, there is the cultural level. The cultural level is described by Bargh as the “software” that is able to be downloaded by the genetic “hardware” and then put to use. Culture being absorbed at a young age allows much of our behavior to be predictable when properly specified: “Culture, including language, norms, values, and so on, is “downloaded” after birth, and it reduces greatly the unpredictability of the child’s world, and his or her uncertainty as to how to act and behave in it” (p. 136). Then there is the psychological level at which what we generally think of as learning takes place. It is on this level that our individual experiences shape the way we behave in the future. For example, traumatic experiences may alter our emotional dispositions and moods affecting our future behavior.

Bargh also goes on to report many of the surprising findings from social and cognitive psychology that appear to reveal how unconscious automatic processes tend to cause our behaviors. There is a long-standing research enterprise on automatic attitudes which demonstrates that people’s attitudes become active in the presence of the attitude object: for example, a negative experience with a goose might cause someone’s negative attitude to activate in the presence of geese. Bargh claims that this research shows that “pretty much everything we encounter is evaluated, unintentionally and unconsciously, as either good or bad immediately after we encounter it (i.e., within 250 milliseconds)” (p. 137). There is also the “similarity-liking effect” in which people develop preferences for people who are similar to themselves in appearance, attitudes, and belief; people tend to like others who resemble significant others and parents; and people tend to like others who share their preferences for places to live and occupations. However, people seem to be unaware that such resemblances and similarities are a factor in their liking.<sup>55</sup>

---

<sup>55</sup> It’s worth noting that Bargh does report some automatic and unconscious priming effects that others have attempted to replicate but failed. In particular, Bargh (2001) reported that when subjects were primed with words like “strive” and “attain” they performed better on demanding cognitive tasks than when primed with more neutral

All of this evidence marshaled by Bargh is impressive, and it demonstrates a myriad of functions able to be performed by the unconscious mind. But I think the objectionable move made by Bargh is his suggestion that because consciousness isn't needed to perform these functions, consciousness is almost entirely absent from human action. Over and over again Bargh is keen on making the point that many of the tasks we take to be the role of consciousness, like goal pursuit and guidance, can actually be performed in the absence of consciousness. Bargh (2005) writes:

I have argued here that conscious acts of will are not necessary determinants of social judgment and behavior; neither are conscious processes necessary for the selection of complex goals to pursue, or for the guidance of those goals to completion. Goals and motivations can be triggered by the environment, without conscious choice or intention, then operate and run to completion entirely nonconsciously, guiding complex behavior in interaction with a changing and unpredictable environment, and producing outcomes identical to those that occur when the person is aware of having that goal. . . . if all these things can be accomplished without conscious choice or guidance, then the purpose of consciousness (i.e., why it evolved) probably lies elsewhere (p. 52).

But the move from this position to the view that consciousness actually doesn't play much of a role in action production or guidance is quite a leap. Bargh (1997a) makes this position clear when he writes that it is "hard to escape the forecast that as knowledge progresses regarding psychological phenomenon, there will be less of a role played by free will or conscious choice.... That trend has already begun..., and it can do nothing but continue" (p. 1). In the same article, Bargh not only says that he is attempting to remove "consciousness from its privileged place at the meditational center of everything" but that he is claiming consciousness is actually in a sense unnecessary (p. 50, 52).

John Kihlstrom (2008) discusses what he calls the "Automaticity Juggernaut"—the idea that behavior is governed primarily by automatic processes. Kihlstrom does an

---

words. However, Harris et al. (2013) report that they have failed to replicate such results on two direct replication attempts.

excellent job explaining the origins of the notion of automaticity and in particular the evolution of John Bargh's position with respect to the prevalence of automatic processes. For example, Kihlstrom notes that Bargh (1984) is actually critical of the view that social interaction operates automatically. However, in his later work, Bargh expands the role of automatic processes, claiming that conscious involvement is almost entirely absent from social interaction.<sup>56</sup> It should be noted, though, that Bargh (1997b) did admit his commentators pushed him back from a position of 100% automaticity to "an Ivory soap bar degree of purity in my beliefs about the degree of automaticity in our psychological reactions from moment to moment" (p. 246). Apparently, this means that Bargh thinks only 99.44% of our behavior is automatic, leaving the rest of the work to be done by consciousness and conscious processes.

Unfortunately for Bargh, his position doesn't seem to be supported by the empirical evidence he presents. It is true that experimental evidence demonstrates that automatic processes play a role in many aspects of human behavior, but this is quite distinct from Bargh's suggestion that automaticity pervades all aspects of human life leaving little room for consciousness and control. Nowhere does Bargh provide evidence for the claim that consciousness plays only a miniscule role in human behavior. The mistake, as I've said before, seems to be in thinking that because consciousness isn't necessary for some processes, that consciousness doesn't play any role with respect to those processes. But as I have already shown, many processes that are automatic still have a role for consciousness in several ways. First, skilled behavior that is learned develops into an automatic behavior in virtue of consciously and deliberately practicing the skill. It is only after conscious effort is exerted that the process becomes automatic. Second, many automatic processes which operate outside the realm of conscious awareness aren't forbidden from this space. We can become consciously aware of many automatic processes, modify them and then retrain them. Simply because they ordinarily operate outside the realm of conscious awareness doesn't mean that they must always operate outside the realm of conscious awareness. Neglecting this role of consciousness to become aware of and modify automatic processes ignores an important type of control we can exercise over our behavior. Just because unconscious automatic processes can

---

<sup>56</sup> For example, (John A. Bargh and Chartrand 1999; John A. Bargh 1997a).



perform certain tasks doesn't mean that consciousness doesn't have a role to play with respect to those same tasks.

The second feature of Bargh's argument that I think is mistaken is the fact that he characterizes free will much too narrowly. In particular, even though Bargh associates consciousness with free will he seems to think of free will in supernatural terms. This leads him to attempt an all-encompassing explanation for human behavior that doesn't require us to posit consciousness or free will. But as I will show, if Bargh were not so quick to understand free will in this supernatural sense, there would be room for free will, consciousness, and automatic unconscious processes in a picture of agency.

One need not look very far for evidence that Bargh understands free will in supernatural terms. A 2008 essay title says it all: "Free Will is Un-Natural." In this essay Bargh says he takes the issue of whether free will exists to boil down to "whether *undetermined choices* of action exist and occur" (p. 130). But then Bargh writes, "Yet any scientific—as opposed to philosophic—approach to the question of free will cannot rely upon extraphysical explanatory concepts," specifically concepts that refer to things which "originate in the mind and are not themselves the causal product of any physical or mechanical forces" (p. 131). Similarly, in a separate article, Bargh and Ferguson (2000) write:

...we reject the thesis of free will as an account of the processes that require conscious control. Instead we embrace the thesis that behavior and other responses are caused, including a person's choices regarding those responses; every deliberation, thought, feeling, motivation, and impulse, conscious or nonconscious; is (often multiply) caused (p. 926).

From the above quotations it's clear that Bargh is making a number of mistakes. One mistake that Bargh is clearly making is conflating something's being caused with something's being determined. Peter van Inwagen's 1983 book titled *An Essay on Free Will* sets out in the first few pages the distinction between determinism and the principle of universal causation. Determinism is defined as the thesis "that there is at any instant exactly one physically possible future" (1983, 3). The Principle of Universal Causation

is the thesis “that every event (or fact, change, or state of affairs) has a cause,” and it is false that this principle entails determinism (p. 3). Even though these characterizations are about the nature of the world, it’s easy to see that there will be a difference when applied to individual events—an event’s being caused doesn’t mean that the event is determined. It’s entirely possible that an event can be probabilistically or indeterministically caused. Robert Kane (1996; 1999) and Alfred Mele (1995; 2006) each develop libertarian accounts of free will that make use of indeterministic causation. So there seem to be more theoretic options open than Bargh is actually acknowledging; there is room for choices to be caused that are not determined.

Similarly, Bargh seems to be assuming that if mental processes like decisions are determined that they can’t be free. Historically, compatibilists like Hume (2000) and A. J. Ayer (1954) have attempted to show that determinism need not rule out alternative possibilities and free will by highlighting the differences between actions that are simply determined and those that are compelled or coerced. And more recently Harry Frankfurt (1969) has argued that alternative possibilities may not be required in order to act freely, so it doesn’t matter if determinism rules them out. And even though compatibilism seems to be the more popular view among philosophers writing on the topic of free will, it is by no means the consensus view. There are those who think the arguments in favor of compatibilism fail. But Bargh doesn’t seem to think that these arguments are failures, he seems to be unaware of the conceptual possibility and ignores it completely, providing no argument whatsoever.

Bargh makes use of some very interesting findings as far as automaticity is concerned, but the implications he draws about the experiments with respect to our agency don’t seem to be warranted. Part of the problem seems to be a supernatural understanding of free will. This aside, Bargh also makes several assumptions about causation and determinism. These issues together result in Bargh making insufficiently supported claims about the implications that the experiments on automaticity have for our prospects of acting freely.

## 5.4 Conclusion

This chapter has focused on attempting to refute arguments against free will made on the basis of the psychological literature on automaticity. After attempting to get clear on just what automaticity is and how it might threaten free will, I argued that automaticity appears to be threatening because it seems to entail a lack of conscious and intentional control. However, I suggested that the lack of conscious and intentional control alone doesn't warrant skepticism about our prospects for free agency. A more potent threat appears to be automatic processes which interfere with other more controlled processes. But here again I think there isn't a need to worry that our freedom is threatened. The last section of this chapter dealt with Bargh's argument that automaticity is so prevalent that consciousness and free will need not be posited as an explanation for human behavior. In response, I claimed that Bargh does not adequately support the irrelevance of consciousness and he holds consciousness and free will to be supernatural, ignoring a host of other conceptually possible explanations for the exercise of human agency.

## CHAPTER SIX

# FREE WILL AND THE ALLEGED ILLUSION OF CONSCIOUS WILL

The relationship between consciousness and free will is complicated and not yet well understood, but even so, many people tend to think that some level of consciousness is required in order to act freely. This natural idea has it that our ability to act freely is at least partly in virtue of our ability to consciously reflect on what we have reason to do, weigh our options in light of our goals and desires, and consciously decide on a course of action, and then act accordingly. This natural picture of our agency, however, is the subject of critical scrutiny most notably by psychologist Daniel Wegner. In Wegner's book, *The Illusion of Conscious Will* (2002), he advances the idea that our conscious will is, in fact, an illusion.<sup>57</sup> Rather than our conscious mental states causing our actions, he argues, such states do not cause anything and simply give rise to the illusion that we consciously control our behavior. If this is indeed the case, then it would be devastating to our prospects for acting freely. The aim of this chapter is to take a closer look at Wegner's claim that conscious will is an illusion and evaluate whether the evidence he amasses provides reason enough to accept such a claim.

### 6.1 Conceptual Investment in Consciousness

The relationship between consciousness and free will is a complex issue, and there is no consensus on just how these two things are intertwined. Part of the difficulty is that there are so many different ways to understand the concept of consciousness itself. David Chalmers (1995) describes consciousness as an ambiguous term “referring to many different phenomena” (p. 200). And Ned Block (1995) even goes as far as to say that consciousness is a “hybrid” or “mongrel” concept.<sup>58</sup> Some philosophical projects involving consciousness require making distinctions among the various kinds or types in order to make progress. However, for now, I can make do

---

<sup>57</sup> Wegner advances this and related ideas elsewhere, also. See Wegner (2005; 2008) and Wegner and Wheatley (1999)

<sup>58</sup> Block (1995) aims to distinguish between what he calls phenomenal consciousness—the what-it's-likeness of a state—and access consciousness—availability for use in reasoning.

with speaking of consciousness in more general terms. It is the goal of the following section to sort out exactly what type of consciousness is alleged to be illusory. Only then will I begin to evaluate the extent to which free will is threatened. That being said, it appears that most people (theorists and non-philosophers alike) believe that some level or type of consciousness is required in order to act freely. Let me begin with the non-philosophers.

Through survey-style experimental philosophy, Joshua Shepherd (2012) has found that consciousness plays a central role in the folk conception of free will. In particular, Shepherd found that when conscious states are said to cause behavior people tend to judge the resulting action as free, and that when it was unconscious states that are said to cause behavior, people tend to judge that the action was not free. Such experiments provide substantial evidence that consciousness plays a central role in the folk conception of free will. Additional evidence that non-specialists associate consciousness with free will comes from a different type of study conducted by Stillman, Baumeister, and Mele (2011). In this study, participants were asked to write a short essay describing a particular event in their life. One group of participants was asked to describe an event they considered to have been of their “own free will”; the other group of participants was asked to relate an event that was “not the result of free will” (p. 387). The essays were then coded by independent research assistants on several dimensions, one of which was consciousness understood as careful consideration of one’s actions. Not surprisingly, Stillman et al. found that “the level of conscious thought was higher among those reporting free actions than among those reporting unfree actions, as predicted” (p. 389). These two studies alone provide strong evidence that non-philosophers strongly associate the concept of free will with consciousness.

Non-philosophers aside, many philosophers have noted the importance of consciousness for free will. P. F. Strawson asserted that genuine freedom depends on one’s behavior being “intelligible in terms of conscious purposes rather than in terms only of unconscious purposes.”<sup>59</sup> Alfred Mele (2008a) even reports that “If all behavior were produced only by unconscious processes, and if conscious choices and their neural correlates were to play no role at all in producing any behavior at all, free will would be in dire straits” (p. 332).

In addition, nearly all philosophical accounts of free will require some level of consciousness. Some accounts require an agent to at least be able to consciously reflect on his

---

<sup>59</sup> Cited by Dennett (2008, 249).

reasons for acting and to consciously decide on the basis of those reasons. Randolph Clarke (2003) notes that an agent must have the capacity for rational self-determination which involves reflecting on and modifying one's reasons and the ability to exercise this capacity consciously (p. 16). Imagine a creature with no consciousness at all. Such a creature would likely perform no actions, let alone free actions.

One speculation as to why consciousness is so closely aligned with free will is that many free will theorists focus their attention not only on overt bodily actions but also on mental actions such as decisions. If we follow Mele in understanding decisions as momentary actions of intention formation which resolve uncertainty about what to do and often times follow careful reflection and deliberation, then it seems clear that most of the decisions we are familiar with are conscious or at the very least seem to be in the forefront of our awareness.<sup>60</sup> This might help explain the following two claims. Mele (2003b) writes that “the power to make practical decisions lies at the heart of much of the literature on freedom of the will and freedom of action (p. 212). And Derk Pereboom (2001) claims that moral responsibility and the control most associated with it applies primarily to decisions. If most decisions we make are conscious decisions, then it's no surprise that consciousness is so closely connected to free will.

## **6.2 What is the Illusion of Conscious Will?**

The previous section showed that consciousness appears to be a necessary component of many accounts of free action and that the folk conception of free will is associated with consciousness. And as I reported earlier, Daniel Wegner has claimed that our conscious will is actually an illusion. Wegner's claim, if true, would seem to be problematic in the face of the role consciousness appears to play in the folk concept of free will and our free will theories. I say this would “seem to be problematic” because it actually being problematic hinges on what the illusion of conscious will actually is. Wegner gives several characterizations of this claim, and they're not all equivalent to one another. In this section I will try to pin down exactly what Wegner's claim is supposed to mean. So my guiding question is, of course, what is the illusion of conscious will. One explanation is given by Wegner in the following passage:

---

<sup>60</sup> Mele (2003b) ch. 9.

This means, though, that conscious will is an illusion. It is an illusion in the sense that *the experience of consciously willing an action is not a direct indication that the conscious thought has caused the action* (2002, 2).

In a footnote, Wegner says that calling this phenomenon an illusion might be “a bit strong, and it might be more appropriate to think of this as a construction or fabrication” (2002, 2, N2). He uses the term ‘illusion’ because it conveys the possibility that we falsely place a large emphasis on how our will appears to us and assume that the appearance is a deep insight. In any case, this explanation of the illusion describes the relationship between two parts of an action: the experience of consciously willing an action and the action being caused by our conscious thought. This characterization of the illusion of conscious will still lacks precision. Even so, I will argue that there is a *prima facie* understanding of it that is intuitive. Unfortunately, I don’t think this understanding produces any claim that is threatening to free will. Upon digging deeper into what Wegner means by this characterization of the illusion of conscious will, I will show that it resists attempts to distil a precise claim due to its lack of conceptual clarity.

So how can we intuitively understand this initial characterization of the illusion of conscious will? Wegner writes about the conscious will as understood in two ways: as an experience and as a force (2002, 3). On the one hand, the conscious will is an experience or a feeling we have when we perform an action—it is a feeling of voluntariness or doing a thing on purpose (2002, 3). On the other hand, Wegner writes about the conscious will as a force which causes action and says it’s appropriate to speak of the conscious will in this sense as “a name for the causal link between our minds and our actions” (2002, 3). Take, for instance, the action of driving a nail into the wall to hang a picture. In performing this action we do many things, such as hold the hammer in one hand and hold the nail in the other. We rely on our perception of the nail held up to the wall to guide our hand to properly place the nail. And we must then begin to tap the nail into the wall by applying enough force with the head of the hammer so that the nail goes in straight. We exert effort with our arm and continue to pound the nail into the wall until it has sunk far enough. Part of Wegner’s point is that performing such an action is accompanied by an experience or feeling that we performed the task on purpose or voluntarily—there is something it is like to have performed said action.<sup>61</sup> However, Wegner claims, it’s also

---

<sup>61</sup> See also Ginet (1990), though Ginet aims to provide a necessary condition for actions.

appropriate to talk about the conscious will in the sense that it was our mental state that produced the action—it was our intention to hammer the nail into the wall (of which we were consciously aware) that caused our body to move in the right sort of way that would carry out the action. The source of the illusion of conscious will, according to Wegner, is confusing the experience of conscious will—the experience of driving the nail into the wall—with the conscious will as a force—our conscious mental state that initiated and sustained the action of driving the nail into the wall. So the rough idea behind Wegner’s characterization of the illusion of conscious will is that just because we have the feeling of doing (some action) it doesn’t mean it was our conscious mental state or thought that caused the action.

At a first glance, this sort of thesis doesn’t strike me as terribly detrimental to my sense of agency or free will. It’s hard to see how one could confuse these two senses of conscious will. Even once we make the distinction between the experience of conscious will and conscious will as a force, it doesn’t make too much sense to say that what is going on is that we often mistake the two. Why would one think that it is the experience or the what-it’s-likeness that is doing the causing in performing an action? Though the experience may accompany performing the action, it’s unclear why one would think the experience is causally efficacious.

Similarly, it’s not surprising that our sense of doing something could come apart from our actually doing something. I could have a sense of doing some action though it turns out that I didn’t do it or that something else caused me to do it in a different way. I have flipped through the channels on the car radio then stopped only to have the channels continue to change because the radio was set to scan the whole time. I experienced changing the radio station each time I pressed the seek button, but it was my accidentally pressing the scan button initially that was causing the radio station change. Even though I was still doing something, namely, pressing the button on the radio, my sense of what I was doing didn’t jibe with what I was actually doing.

When we think about it, it seems that these kinds of thing happen to us more frequently than we realize, according to Wegner, so we shouldn’t be surprised by the claim that the conscious will is an illusion—that the experience of consciously willing something doesn’t mean that our conscious thoughts actually caused anything. But, in fact, these and similar cases are unusual. Wegner himself says that most of the things we do in everyday life, the usual cases, are



ones in which action and the experience of will correspond—we experience doing something and we actually have done it (2002, 11).<sup>62</sup> So what’s all the fuss about the illusion of conscious will?

Wegner seems to want to say that the unusual cases demonstrate to us that the experience of consciously willing can come apart from the production of action; they let us see these two sides of the will. This causes us to wonder, says Wegner, whether the feeling and the actual production of action may be produced by different systems. So maybe there is more to the illusion of conscious will than just the unusual cases where the feeling comes apart from actually doing. Maybe there is more to Wegner’s characterization of the illusion of conscious will. However, I will go on to argue that once we look a little closer at how the terms in Wegner’s characterization of the illusion of conscious will are explained and employed, we will see that they offer us no further understanding as to how the illusion of conscious will threatens free will.

I’ve discussed Wegner’s distinction between the experience of conscious will as a feeling of doing and the will as a causal force which produces action. Wegner uses this distinction in order to argue for the illusion of conscious will, so it’s important for Wegner’s purposes that this distinction is properly laid out. But even though this distinction between conscious will as an experience and as a force has some prima facie intuitive pull, other things Wegner says with respect to the conscious will seem to confuse rather than clarify the distinction. Timothy Bayne (2006) notes that, in general, Wegner’s distinction between the experience of conscious will and conscious will is easily understandable: “The experience is one thing, the property or object experienced is another” (p. 3). But Bayne goes on to argue that in several places Wegner associates the conscious will with a feeling like happiness or sadness. This, Bayne says, threatens to collapse the distinction between the will itself and the will as an experience. Collapsing this distinction not only harms Wegner’s argument for the illusion of conscious will but leaves us wondering how to even understand what the illusion of conscious will is.

Wegner’s distinction seems to raise more questions than answers by requiring the experience of willing an action in order for some behavior to count as being consciously willed or “truly *willed*” (2002, 3–4). Consider the following passage:

---

<sup>62</sup> Bayne (2006) makes the point that there is a slight tension here: Wegner’s entire account of the illusion of conscious will relies on the fact that the experience of conscious will corresponds with the action occurring.

If a person plans to take a shower and says that she intends to do it as she climbs into the water, spends fifteen minutes in there scrubbing up nicely, and then comes out reporting that she indeed seems to have had a shower but does not feel she had consciously willed it—who are we to say that she did will it? Consciously willing an action requires a feeling of doing, a kind of internal “oomph” that somehow certifies authentically that one has done the action (2002, 4).

Requiring experiencing willing an action for something to qualify as truly willing an action makes it unclear what happens when the experience of willing an action comes apart from the will itself as a force. What is an action that is not consciously willed? Is it not an action at all but something closer to mere behavior akin to a sneeze? Wegner doesn't explain, and his characterizations of the conscious will and the experience thereof don't help elucidate what the illusion of conscious will might be.

Wegner's characterization of the illusion of conscious will also involves the notion of conscious thought causing action. Rather than help us understand what the illusion of conscious will is, though, this notion confuses things even more. Wegner expands upon what he means by 'conscious thought' when he writes, “The new idea introduced here is the possibility that the experience of acting develops when the person infers that his or her own *thought* (read: intention, but belief and desire are also important) was the cause of the action” (p. 66). Wegner understands intention as “an idea of what one is going to do that appears in consciousness just before one does it” (p. 18). This account is clearly not sufficient for an intention. If the idea of falling into a large hole in the ground appears in my consciousness just before I actually do fall into the hole, I would not be inclined to say that I intended to fall in the hole.<sup>63</sup> A better understanding of intentions has it that intentions are executive attitudes towards plans of action about what to do.<sup>64</sup> On this view, intentions having an executive element means, roughly, that intending to go to the grocery store entails being settled (but not irrevocably) on going to the grocery store (Mele 2009, 6–7). This element of intention helps distinguish it from other attitudes we may have toward our plans e.g., believing one has such a plan and desiring to execute such a plan.

---

<sup>63</sup> Mele (2009) makes this point in chapter 1.

<sup>64</sup> See, for example, Mele (1992; 2003b; 2008a; 2009)

Furthermore, intentions can be occurrent or standing. Mele proposes that there are two ways an intention can be occurrent. The first way is that the intention is “suitably at work at the time in producing relevant intentional actions or in producing items appropriate for the production of relevant intentional actions” (2009, 4). The second way is “roughly, for it to be a conscious intention at that time, provided that the intention “is not wholly constituted by a disposition to have occurrent intentions to A” (2007; 2009, 4).<sup>65</sup> My intention to take the garbage out right now may count as an occurrent intention if it is a conscious intention which initiates and guides my intentional action of taking out the garbage and all the steps involved in this process. Standing intentions, on the other hand, Mele analyzes as “a disposition of a certain kind to have a corresponding occurrent intention” (2009, 4). Having decided to go to the symphony, I purchased my tickets and made a plan to attend the symphony on the date of the concert. Even though I am aware of my intention now, I was not aware of it earlier today or last night when I was asleep and not even conscious. Nonetheless, the intention to go to the symphony can still be attributed to me because I have a disposition to have the occurrent intention to go to the symphony. The type of intention which can be attributed to me even when I’m not conscious is best characterized as a standing intention.

Additionally, intentions can be proximal or distal. Distal intentions are intentions about the non-immediate future e.g., my intention to send my brother a birthday card this week; proximal intentions are intentions about the immediate future (Mele 2009, 10). My intention to call my mother right now would count as a proximal intention.

And finally, we need to understand what Wegner means by “direct indication” in his characterization of the illusion of conscious will. Mele (2009) writes that he prefers not to speculate about exactly what Wegner means by “direct indication” and instead considers more precise versions of Wegner’s claim about the illusion of conscious will (p. 95). Bayne (2006) is similarly puzzled by what Wegner means by “direct indication” in the passage cited above. Bayne notes that one possibility is that a direct indication view of the phenomenology of agency involves thinking that one’s experiences and beliefs about one’s own agency involve no sub-personal processing or inferential mechanisms of any kind. But it seems that such direct

---

<sup>65</sup> Of course not all dispositions to have a corresponding occurrent intention are standing intentions. See Mele (2007, 746–54) for the details of his analysis.

indication or transparency views of mind and agency have been widely rejected.<sup>66</sup> In addition, Bayne notes that there doesn't seem to be any reason for thinking that experiences of agency could be reliable only if the direct indication theory were true and that the rejection of direct indication views of visual perception has not led theorists to reject visual experiences as illusory (2006, 181). Therefore, the rejection of direct indication accounts of the phenomenology of agency should not lead us to reject the phenomenology of agency as an illusion.

Another possibility is that Wegner means that the experience of consciously willing an action doesn't guarantee that the intention caused the action. But if this is what Wegner means then I don't think we should be worried about the illusion of conscious will. Part of the reason we shouldn't be worried is that, as Wegner has shown, the experience of doing can come apart from actually doing. In fact, we shouldn't be surprised that the experience of willing an action doesn't guarantee that our intention actually caused the action; the intention needs to accompany the experience of doing. One can experience performing an action that they do not actually perform. Perhaps Wegner has in mind the claim that the experience of consciously willing an action doesn't count as evidence that the intention actually caused the action. But this claim seems implausible for a number of reasons. Our conscious experiences seem to be a primary source of evidence for cognitive goings on. Also, as we shall see in a later section, Wegner's model of apparent mental causation relies on the conscious experience serving as evidence for the things we actually do. In any case, from what has been said so far, the phenomenon of the experience of doing coming apart from actually doing doesn't seem to threaten the possibility of acting freely. So Wegner must mean something else.

We can get an idea of what that something else is by examining an analogy Wegner draws between the illusion of conscious will and an illusion performed by a magician. When a magician performs a magic trick there is a set of "perceived causal sequences" and a set of "real causal sequences." The magician's task in performing the trick is to make "magic" the easiest explanation for the perceived causal sequence while the real causal sequence orchestrated by the magician remains out of sight. Wegner says the illusion of conscious will has much of the same structure. There is a perceived causal structure to the nature of our actions and there is a real causal structure. The illusion of conscious will is thinking that the perceived causal sequence is what causes our actions rather than the real causal sequence. It seems that, given this analogy,

---

<sup>66</sup> See, for example Damasio (1994).

the real causal sequence is “a massively complicated set of mechanisms” that operates below the level of consciousness. This would make the perceived causal sequence our conscious mental states like intentions and our conscious mental actions like decisions. The illusion of conscious will is thinking that our mental states like intentions and mental actions like decisions actually cause our actions. According to Wegner, they do not. This claim, if true, would be devastating for our prospects of free will.

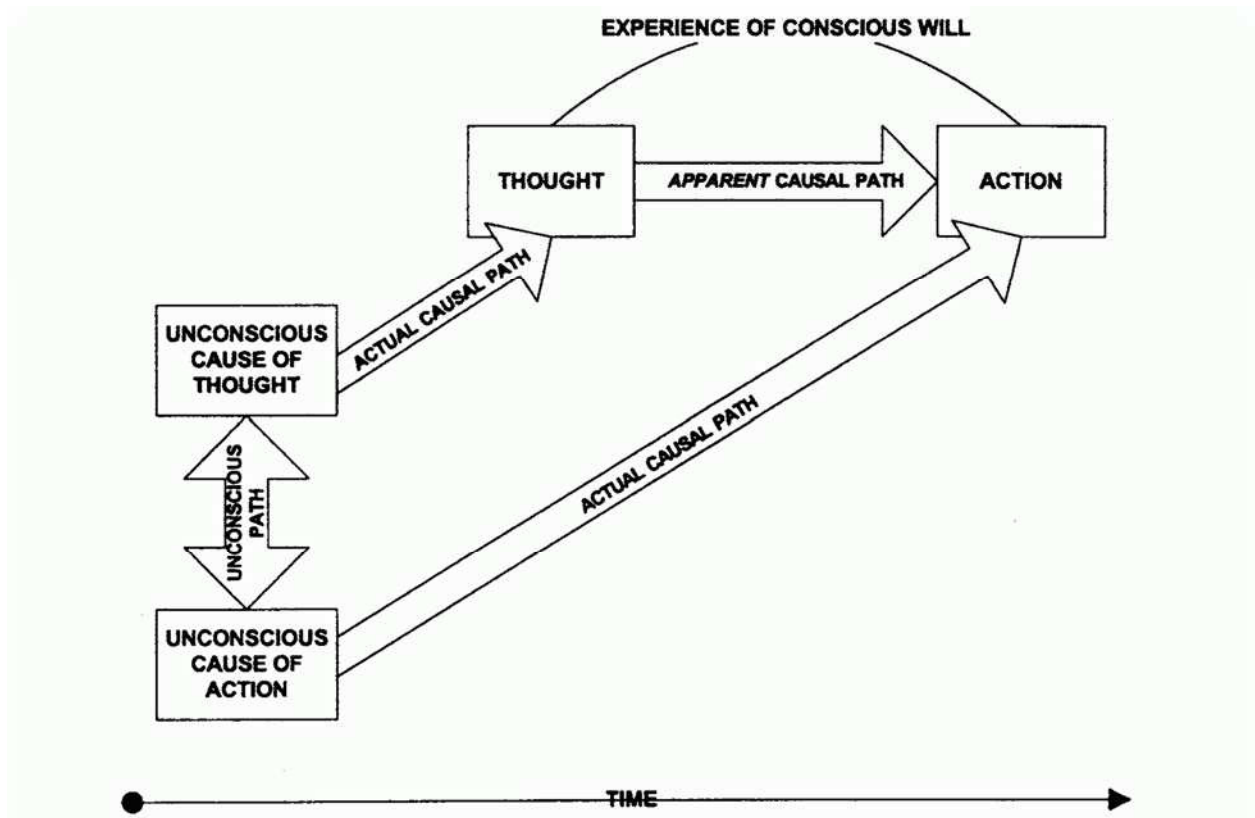
This sort of view is usually labeled ‘epiphenomenalism’, the thesis that “although all mental events are caused by physical events, no mental events are among the causes of any physical events” (Mele 2008b, 119). There are at least two reasons which suggest that Wegner is not arguing for this type of view. The first consideration is Wegner’s model of apparent mental causation, which is illustrated by Figure 6.1.<sup>67</sup> According to this model, the actual causal sequence resulting in an action begins with an unconscious event or state. At the same time there is another unconscious event or state which causes the conscious mental state that precedes the action. These two unconscious events are distinct on Wegner’s view. He also writes the following:

There may or may not be links between these underlying unconscious systems (as designated by the bidirectional unconscious potential path), but this is irrelevant to the perception of the apparent path from conscious thought to action. Any actual path here cannot be directly perceived, so there may be no actual path (2002, 67).

Epiphenomenalism as characterized above allows for the possibility that the neural correlates of mental events are among the causes of physical events such as actions. What epiphenomenalism disallows is for the mental events (or states) to be among the causes of physical events like actions. Wegner’s model of the experience of conscious will, however, doesn’t disallow this possibility. It is consistent with his model that the unconscious event which is the cause of a physical event such as an action is a mental event (or state).

---

<sup>67</sup> This is Wegner’s own illustration and can be found on p. 68 of Wegner (2002).



**Figure 6.1:** Wegner’s Model of Apparent Mental Causation

The second consideration in favor of thinking Wegner is not advocating epiphenomenalism as characterized above is because Wegner does think that consciousness and conscious mental states can play a kind of causal role. Wegner claims that the experience of conscious will does a couple of things. First, it is an interpretive system that recognizes the correlation between thoughts and action (and thus generates the illusion of conscious will). This system alerts us and brings to our attention actions that are ours or result from our own agency. In this way, Wegner claims the experience of conscious will is an “authorship emotion” (2002, 265). The experience of conscious will helps us to know what events we have authored and what events we have not.

Secondly, the conscious will as an interpretive system “organizes and informs our understanding of our own agency” (2002, 318). The conscious will serves as a sort of collector of which actions we have authored and helps us maintain a sense of our own identity. This means, though, that the experience of conscious will can serve a kind of causal role insofar as being aware of and being able to remember what one has done may shape one’s future actions.

Wegner notes that this feature of conscious will has been studied by researchers investigating the phenomenon of perceived control. A number of psychological studies indicate that perceiving one's own behavior as more controlled or skilled can have dramatic effects on subsequent behavior. Bandura (1993) has shown that one's perceived self-efficacy influences how one performs in an academic setting. Self-efficacy with respect to one's mathematical abilities has been shown to be a greater predictor of achievement than previous scores on similar tasks.<sup>68</sup> Similarly, negative views of abilities or control can have a negative impact. What this shows is that the experience of conscious will, the way in which the conscious will organizes and informs our understanding of ourselves plays a causal role on our future behavior. Such a causal role is disallowed by epiphenomenalism.

### 6.3 Wegner on Consciousness, the Will, and the Self

Before looking at and evaluating the evidence that Wegner appeals to in order to support the claim that neither consciousness nor conscious mental states cause actions, it is worth explaining in more detail Wegner's understanding of consciousness and conscious mental states. After all, it is Wegner's understanding of consciousness and conscious mental states that leads Eddy Nahmias (forthcoming) to claim that Wegner's data should not be interpreted as threatening free will.

The first important feature I want to bring attention to is that Wegner writes about the conscious will and psychological mechanisms as two competing and mutually exclusive explanations for human behavior. At the very beginning of his book Wegner writes, "The idea of conscious will and the idea of psychological mechanisms have an oil and water relationship, having never been properly reconciled" (2002, 2). And later on Wegner urges us to think of it "in terms of lenses":

If each person has two general lenses through which to view causality—a mechanical causality lens for objects and a mental causality lens for agents—it is possible that the mental one *blurs* what the person might otherwise see with the mechanical one. The

---

<sup>68</sup> See Pajares & Miller (1994).

illusion of conscious will may be a misapprehension of the mechanistic causal relations underlying our own behavior that comes from looking at ourselves by means of a mental explanatory system (2002, 26).

And Wegner claims that it is

The tendency to view the world in *both* ways...[that] has created in us two largely incompatible ways of thinking. When we apply mental explanations to our own behavior-causation mechanisms, we fall prey to the impression that our conscious will causes our actions (2002, 26).

As for why these two ways of viewing our behavior are incompatible, we are given no argument or explanation. Rather, it seems to be a preconceived notion of Wegner's that he begins with. The reason why Wegner has such preconceived notions of incompatibility, it seems, is due to his conception of the ontological nature of consciousness, which is the second distinguishing feature I would like to bring attention to.

Wegner uses many terms to explain his view of conscious will, but among the synonyms for the conscious will are consciousness and the self. The way Wegner explains these terms is quite telling; in particular, Wegner explains the conscious will or the self as a non-natural entity, going so far as to call the self "magic." In Wegner's (2008) article "Self is Magic," he claims "the magic self stands squarely in the way of the scientific understanding of the psychological, neural, and social origins of our behavior and thought," suggesting that science can in no way capture the self (2008, 226–7). But the view of ourselves as magical selves, Wegner says, is natural and part of being human: "Seeing one's own causal influence as supernatural is part of being human, though" (2008, 228). Perhaps I am out of touch with folk conceptions of agency, but I do not consider this part of being human.

Elsewhere, Wegner writes that thinking that "consciousness" or "the will" are legitimate possible causes of behavior "undermines the possibility of a scientific theory of psychology" (2005, 20). In this article, Wegner goes to great lengths to argue that scientific (including psychological) theories need to abandon the notion of a homunculus—an inner executive agent who is responsible for performing actions. Behavioral theories that appeal to a homunculus are



charged with being vacuous when the behavior the theory was supposed to explain is simply attributed to the homunculus without further explanation.<sup>69</sup>

Where does this leave us? Some may reject the claim that consciousness and conscious mental states don't cause any of our actions simply because Wegner understands consciousness and the self in a supernatural way.<sup>70</sup> On the other hand, Wegner does consider a massive amount of evidence for his illusion thesis. Even if Wegner does have a supernatural understanding of consciousness, it could turn out that the evidence he marshals for the illusion thesis nonetheless supports its central tenets. This means that there is still a possibility that Wegner's evidence could have important implications for whether consciousness or conscious mental states cause behavior, and thus have important implications for acting freely.

## 6.4 The Apparent Mental Causation Argument

One way Wegner argues for the thesis that consciousness and conscious mental states do not cause our actions is by developing a theory of apparent mental causation. The theory of apparent mental causation is supposed to account for why it is that we experience consciously willing our actions when in fact our conscious mental states do not cause our actions. The experience of consciously willing our actions arises when we believe our thoughts have caused our actions. And we believe our thoughts have caused our actions “when we have thoughts that occur just before the actions, when these thoughts are consistent with the actions, and when other potential causes of the actions are not present” (2005, 23). Wegner suggests that these are the three principles—Priority, Consistency, and Exclusivity—that result in our making the faulty inference that our mental states have caused our behavior. The aim of this section is to show that Wegner's account of apparent mental causation is a far cry from providing evidence for the illusion of conscious will. In what follows I will consider each of the principles that constitute Wegner's account of apparent mental causation.

---

<sup>69</sup> Wegner notes that Freud's theory of the id, ego, and superego has been criticized as being a homunculus-based explanatory system (2005, 20). He also notes that in more contemporary theories, the homunculus is often implicit rather than explicit (2005, 21).

<sup>70</sup> Eddy Nahmias (2002) seems to take this route.

### 6.4.1

The first principle of Wegner's account of apparent mental causation is the priority principle. According to Wegner, "The priority principle suggests that the thought must appear in a timely way just before the action for the action to be experienced as voluntary" (2005, 24). When this principle is violated people experience involuntariness or less voluntariness of their actions. The main idea driving the priority principle is appealing—we certainly expect there to be certain thoughts present before we actually perform actions. Wegner's idea is that the thought must occur within a certain timeframe before action. Occurring too early, the thought doesn't result in a sense of voluntariness, and occurring too late (say, after the action has been initiated) also doesn't result in a sense of voluntariness. According to Wegner, "If will is an experience fabricated from perceiving a causal link between thought and action, it should be possible to lead people to feel that they have performed a willful action when in fact they have done nothing" (2002, 74). So Wegner set out to test whether the timing of the thought's occurrence would affect attributions of voluntariness.

The study Wegner appeals to is Wegner and Wheatley (1999). The goal was to learn whether people will feel they willfully performed an action that was actually performed by someone else "when conditions suggest their own thought may have caused the action" (2002, 74). This experiment is often referred to as the I Spy Study. In this experiment, a subject was seated at a table facing a confederate and each of them were instructed to place their hands on a square board that sat atop a computer mouse. The mouse was connected to a computer the screen of which was visible to both persons and showed about fifty small objects.<sup>71</sup> The pair were asked to move the mouse around together in sweeping circles and

...to stop moving the mouse every 30 seconds or so, and that they would rate each stop they made for personal intentionality. That is, they each would rate how much they had intended to make each stop, independent of their partner's intentions. The participant and confederate made these ratings on scales that they kept on clipboards in the laps. Each scale consisted of a 14-centimeter line with end points "I allowed the stop to happen" and "I intended to make the stop," and marks on the line were converted to percent intended (0-100) (2002, 74–5).

---

<sup>71</sup> The photo was a "Tiny Toys" photo from the book *I Spy* by Marzollo and Wick (1992).

The participant and confederate were told that they would hear music and words through headphones they were instructed to wear. Each trial would involve a 30-second interval of movement followed by a 10-second clip of music indicating that they should make a stop. They were told that they would be hearing different audio tracks but that they would hear music at about the same times “and should wait a few seconds into their music before making the stops to make sure they were both ready” (2002, 75). Also, they were both told that they would hear words over the headphones to provide a mild distraction and that the words each person heard would be different.

The words, however, served to prime thoughts about objects on the screen for the participant, and the confederate, in fact, heard neither words nor music but rather instructions to make particular movements at particular times. Four of the trials were “forced stops” in which the confederate received instructions to move to an object on the screen followed by a countdown to stop on the mentioned object. These forced stop instructions were timed to occur midway through the participant’s music and the participant heard the word consistent with the stop either 30 seconds before, 5 seconds before, 1 second before, or 1 second after the confederate stopped on the object, which was different on each trial. The four forced stops were embedded in a series of other trials—“unforced stops”—in which the confederate received no instruction to stop the movement of the mouse. On these unforced stop trials, the participant heard a word 2 seconds into the music which corresponded to an object on the screen half of the time and didn’t correspond to an object on the screen the other half of the time.

Wegner and Wheatley report that on the forced stop trials a pattern of “perceived intention emerged as predicted by the priority principle” (2002, 77). What they noticed was actually a general tendency overall for participants to perceive the forced stops as somewhat intended (the mean intentionality was 52 percent), but the degree to which intentionality was perceived fluctuated depending on the timing of the word. Perceived intentionality was lower when the word appeared 30 seconds before the forced stop than when it appeared either 5 seconds or 1 second before the forced stop. Intentionality was also lower when the word appeared 1 second after the forced stop. Wegner, therefore, claims to have shown that “the experience of will can be created by the manipulation of thought and action in accord with the principle of priority, and this experience can occur even when the person’s thought cannot have

created the action” (2002, 78). But as I will argue, there are several reasons that this claim is dubious.

One issue of paramount importance is what Wegner means by ‘thoughts’. But since this was discussed in an earlier section I will not dwell on it anymore here except to point out that it is unlikely that priming participants with certain thoughts sheds any light on the role intentions play in generating an experience of conscious will. However, even if the priming done in this experiment did lead the participants to form the corresponding intention, Wegner would have a necessary component for the case that conscious mental states don’t cause our actions but, as I will show, there would still be plenty of reason to think that his case falls short of convincing.

Another glaring concern with this study is with the particular instructions that were given to the participants. They were specifically instructed to “stop moving the mouse every 30 seconds or so” (2002, 74). Then after stopping the mouse, they were asked to rate the intentionality of the stopping. These instructions are so vague that it’s unclear how the participants understood them. It seems as though they would need to have a kind of standing intention to stop moving the mouse at some point while the music is playing. If this is the case, then given these instructions, it should not be a surprise to discover that even when the participants are moving a mouse cursor on a computer screen with another person they will rate the action of stopping the mouse as somewhat intentional—it’s what they were told to do! Even during the so-called forced stop trials, the participants were supposed to follow the instructions of stopping the mouse within the same window of time that the confederates forced them to. Therefore, it seems rather hasty to conclude that the participants’ “thought cannot have created the action,” as Wegner puts it.

Wegner maintains that the participants could not have played a role in stopping the mouse at all for the forced stops by appealing to an analysis of unforced stops. The point of this, he claims, is to see whether participants might naturally stop on the objects that were mentioned during the unforced stops. Wegner claims that if the cursor did stop on the items that were mentioned over the participants’ headphones for the unforced stops, it would suggest that the participant may have played some role in the forced stops. So what Wegner and Wheatley did was the following:

Distances between stops and objects on the screen were computed for all unforced stops (all trials in which the confederate heard no instruction and simply let the participant make the stop). The mean distance between the stop and an object on the screen (e.g., dinosaur) was measured separately for stops when that object was the mentioned word and for stops when the mentioned word was something not shown on the screen (e.g., *monkey*). The mean distance between stop and object when the word referred to the object was 7.60 centimeters, and this was not significantly closer than the distance of 7.83 centimeters when the word did not refer to the object. Thus, simply hearing words did not cause participants to stop on the items. The forced stops created by the confederate were thus not likely to have been abetted by movement originated by the participant (p. 76-7).<sup>72</sup>

But this method of analysis doesn't make sense to assess whether participants played a role in stopping during the forced trials. The mean distance between where the participants stopped and the object when the word they heard referred to the object compared to the mean distance between where the participant stopped and the object when the word they heard didn't refer to the object isn't evidence of any sort for the causal role played by the participant during the forced stops. Given this in conjunction with how the participants were instructed to perform the task, it seems likely that the participants still played some role in stopping the mouse.

Adding further complications to this is the scale the participants were given to rate the intentionality of stopping the mouse. Their rating had to fall on a continuum between "I allowed the stop to happen" and "I intended to make the stop," which don't seem to be mutually exclusive or represent all the varying degrees of intentionality of an action. Allowing the stop to happen can still be fully intentional. But according to Wegner and Wheatley's study, this is considered to have 0 percent intentionality. Similarly, one can intend to make the mouse stop but not execute the intention in the way they planned. According to Wegner and Wheatley, this would be 100 percent intentional.

Nonetheless, the effect that was purported to be found in the experiment was that the degree of intentionality increased when the word appeared 5 seconds and 1 second before the

---

<sup>72</sup> Note that it's unclear how the mean distance between the stop and object was measured when the word mentioned didn't refer to an object on the screen.

forced stop. This is actually good evidence for the priority principle, but it's not the priority principle that I think is problematic. What I do think can be said is that given the lack of precision about the term 'thoughts', the particular instructions given to the participants, and the scale of intentionality used, it seems unlikely that the I Spy experiment actually demonstrates a feeling of doing absent any actual doing on the part of the subjects.

## 6.4.2

Wegner's second principle—The Consistency Principle—requires the mental attitude (e.g., an intention) to be 'consistent' with the action that occurs for there to be the appearance of mental causation. This requirement makes sense for understanding mental causation. If I intend to change the television channel by pressing a button on the remote but my hand throws the remote across the room, I would be less likely to think my conscious thoughts have caused my behavior. I might think, instead, that I have a neurological disorder or that something else is severely wrong with me.

One problem with this principle is raised by Bayne (2006). Bayne notes that it's unclear what the relationship is "between the awareness of a match between one's intentions and actions (on the one hand) and the experience of doing" on the other (2006, 19). Bayne illustrates the problem by describing a case of causal deviance.<sup>73</sup> In Davidson's (1963) original case, a climber belaying another has the desire to drop his partner and the belief that loosening his grip would accomplish this. But his very recognition that he has this belief/desire pair causes him to lose his grip on the rope resulting in his partner falling. In this case, the reason matches the behavior it has caused but not quite in the right sort of way. Bayne says that cases of this sort appear to show that one can experience one's movements and behavior as caused by their intention without experiencing a sense of agency that one is likely to experience in ordinary cases of action. Something more is needed, Bayne claims, and Wegner does not develop this part of his theory.<sup>74</sup>

Despite this theoretical shortcoming, Wegner does seem to garner empirical support for this principle. In particular, Wegner reports a study in which subjects experienced a sense of performing some action in virtue of having thoughts consistent with the action being performed.

---

<sup>73</sup> Causal deviance is often employed as a problem for the causal theory of action. Donald Davidson (1963), a proponent of the causal theory of action, brought attention to this problem.

<sup>74</sup> Bayne (2006) goes on to explore one response but finds it wanting.

Wegner, Sparrow, and Winerman (2004) had participants watch themselves in a mirror while another person behind them, hidden from their view, extended their hands forward on each side and performed a series of movements. When the subjects were played a tape recording with instructions to perform particular movements that matched the movements of the hands, they experienced a greater sense of control over the hands that were not their own. Similarly, Wegner reports another study demonstrating this effect involving magical thinking. Pronin et al. (2006) had participants read an article suggesting the possibility of voodoo curses before the experiment. Then the experimenters had the participant stick pins in a voodoo doll said to represent another individual, the victim, who was in fact a confederate that later faked a headache. The experimenters were interested in finding out whether the subjects would feel causally responsible for the victim's headache as a result. It turned out that many of the participants did accept the idea that their behavior caused a headache in the other person.<sup>75</sup> In a second experiment, the experimenters tested magical thinking with respect to a positive outcome. They had participants envision the success of a basketball player shooting free throws. When the player shot the ball and made six out of eight baskets the subjects reported having a greater sense of causal responsibility compared to subjects who instead envisioned the player lifting barbells.

There seems to be a good deal of empirical evidence to support Wegner's consistency principle. The experiments do seem to show that an agent's thoughts need to be consistent with the perceived behavior in order to acquire a sense of agency. The problem is that the consistency principle doesn't do any work in distinguishing cases of apparent mental causation from actual mental causation. The consistency principle that Wegner says is required for a theory of apparent mental causation is also a principle that is required for actual mental causation! If one were to give an account of what is required for an intention to appropriately cause an intentional action, one would want to include an appropriate match between the representational content of the intention and the action that is performed. The experiments appealed to by Wegner are all examples of instances we know are impossible for there to be any actual causation. We know, for example, that pricking a voodoo doll won't cause pain in another individual no matter how much of a likeness the doll is of that person. And we know that our positive thoughts—no matter how strong and vivid—didn't cause the batter to hit that game winning homerun. But

---

<sup>75</sup> Wegner (2008) notes that the subjects didn't have to believe in voodoo but only needed to believe that the potential victim might be "stressed into a headache by the shock of receiving a curse" (p. 229).

when we start looking at our everyday actions like deciding to wash the dishes and then going about the chore, we may become more skeptical that the causation is merely apparent. Unlike the experiments, we have a working explanation for how such causation might work. Unless Wegner's account can give a satisfactory explanation for why our conviction that there is a distinction between actual and apparent mental causation persists, it's unlikely to be very convincing.<sup>76</sup>

### 6.4.3

The last principle Wegner discusses with respect to the theory of apparent mental causation is the Exclusivity Principle: "people see their thoughts as causing events to the degree that there are no other plausible candidate causes" (2008, 233). Wegner goes on to say that "when the thought and only the thought precedes an event...the person will experience the event as flowing from that thought" (p. 233). Part of the support that Wegner marshals for this principle is from the famous Milgram experiments. Stanley Milgram (1974) conducted experiments in which participants were led to believe they were teaching another person by applying electrical shocks whenever the other person (a confederate) performed incorrectly on a task. Milgram found that many participants were willing to apply electrical shocks to the confederate to the point of apparently placing him in physical danger. However, because the subjects were instructed to apply the shocks by a lab technician whom they perceived as an authority figure, the subjects were only willing to take partial responsibility for their actions. According to Wegner and Milgram, the subjects no longer viewed themselves as acting for their own purposes and experienced an *agentic shift* in which the "person entering an authority system no longer views himself as acting out of his own purposes but rather comes to see himself as an agent for executing the wishes of another person" (Stanley Milgram 1974, 133).

This is an interesting case and though it may seem to offer support for the exclusivity principle, I think the support is mere appearance. In the first place, the Milgram experiments show only that people felt less responsible for electrocuting the confederate which is an issue that, while related, can come apart from the question of freedom and control. In particular, it seems to be that the Milgram experiments show that people are more willing to decide to shock someone when instructed to, not that the control they exercise over deciding is diminished. They

---

<sup>76</sup> Thanks to Randy Clarke for making this point clear to me.



may be pressured to act on reasons they think aren't justified or that they don't entirely accept, but the control they exercise over their own bodily movements and decisions doesn't seem to be threatened in the way that Wegner suggests. It's not as if the subjects claimed to not have the experience of acting in a certain way (i.e., pressing the shock button). Rather, they just felt less responsible because it wasn't their idea to shock the other person.

Additionally, it might be fair to say that the exclusivity principle is just a bad principle. Al Mele (2008a), in commenting on Wegner's theory of the illusion of conscious will, writes that "Obviously, even people who believe that some of their conscious intentions (or the neural correlates thereof) play a role in causing some of their behavior should not believe that "the conscious mind is the sole player" (2008a, 341). Even among those people who do not think that the conscious mind is just an illusion it should be obvious that there are a myriad of other factors that play a role in influencing actions one performs.

## **6.5 Conclusion**

The three previous subsections were dedicated to evaluating the three principles involved in Wegner's theory of apparent mental causation: priority, consistency, and exclusivity. The principle of exclusivity was found to lack not only plausibility but adequate empirical support as well. And while the first two principles aren't too controversial, I think that they do not do the work that Wegner thinks they do. Wegner's theory of apparent mental causation doesn't actually explain why mental causation is illusory as opposed to actual. The three principles might be some of the necessary ingredients that would be required if mental causation were illusory, but we lack sufficient empirical support to claim that mental causation is illusory. In fact, the theory of apparent mental causation seems to presuppose that mental causation is illusory and then Wegner goes about attempting to explain why we fall prey to this illusion. At the end of the day, we still lack any convincing evidence that mental causation is actually illusory.

## CHAPTER SEVEN

### CONCLUSION

This dissertation has looked at several research programs in psychology in an attempt to determine if the negative claims they have been used to make about our freedom are warranted. The main conclusion I have arrived at is that such negative claims about free will appear to be ill supported by the evidence.

Situationism, on the face of it, appears to be threatening to our freedom. When we hear that there is evidence in psychology that our behavior is causally influenced by features of our situation or environment that we are unaware of, we can tend to feel like we are manipulated by our environment or that we are not really in control of our behavior. But after looking more closely at the arguments for exactly how situationism might show we are less free, it appears that we are not forced to concede that situationism rules out our free will. In particular, situationism doesn't seem to show that we lack the required awareness to act freely, and situationism doesn't appear to show that we lack the capacity or ability to act freely.

Baumeister's work on self-control is interesting and provides an illuminating model for understanding how our capacity for self-control works. But his use of this research to support negative claims about free will does not seem to be justified. Part of the reason is that Baumeister seems to be thinking of free will as being constituted by the capacity for self-control. If he was right about this, then his experiments might show that we act freely less often than we think, but this is an idiosyncratic conception of freedom. Free will involves more than just our capacity for self-control and it's not obvious that a diminished capacity for self-control diminishes our free will.

The literature on automaticity which has also been used to support the idea that we are less free than we think I also found to be wanting. The presence of automatic processes, though they may be pervasive, do not show that we lack free will. There is room for a picture of agency to include controlled and intentional behavior along with automatic processes. Though there may be types of automatic processes that could potentially interfere or disrupt controlled processes in a way that we undermine our freedom, such processes do not do this in virtue of being automatic processes.

Lastly, I looked at Wegner's research on the illusion of conscious will. Wegner's theory, if it could be empirically supported, might show that we lack free will. If it could be demonstrated that our consciousness plays no role at all in producing our actions, then free will would be severely threatened. However, Wegner's view is far from being empirically supported. Wegner's theory of apparent mental causation assumes rather than proves that conscious will plays no role in action production, and the empirical evidence he uses doesn't do much to convince one that the conscious will is illusory. Wegner's claims seem to have been arrived at too hastily.

So, does all of this evidence show that we are, in fact, free? By challenging some of the strongest empirical evidence used to support the idea that we are not free, I hope to have provided some room for optimism about the prospects for our freedom. This dissertation hasn't shown that we do, in fact, possess all the required capacities and awareness to act freely. However, it has shown that some of the arguments against free will are severely flawed.

## APPENDIX

### COPYRIGHT PERMISSION FOR FIGURE 6.1

**From:** Zachary Martin  
**Date:** Monday, July 7, 2014 6:06 PM  
**To:** Pamela Quick <[quik@mit.edu](mailto:quik@mit.edu)>  
**Subject:** New message from Zachary Martin

There is a new contact message form Zachary Martin. You may login and administer messages [here](#).

Name: Zachary Martin

Message:

My name is Zachary Martin. I am a Ph.D. candidate at Florida State University's Department of Philosophy. I am writing to ask for permission to include a copyrighted figure in my doctoral dissertation. I would like to use a figure from Daniel Wegner's "The Illusion of Conscious Will" (2002). It is figure 3.1 on page 68. All elements of the figure would remain the same. I am including it simply to illustrate Wegner's model of apparent mental causation. If you need more information from me, please do not hesitate to contact me.  
Best, Zachary Martin

**From:** Pamela Quick <[quik@mit.edu](mailto:quik@mit.edu)>  
**Date:** Monday, July 7, 2014 7:41 PM  
**To:** Zachary Martin  
**Subject:** New message from Zachary Martin

Dear Zachary,

Thank you for your message. I am happy to grant to you non-exclusive permission to reprint figure 3.1 from *The Illusion of Conscious Will* in your PhD thesis for Florida State University. Please credit the reprinted figure to Daniel Wegner, *The Illusion of Conscious Will*, and The MIT Press.

Please let me know if you have any questions.

Very best,

Pamela Quick  
Permissions Manager  
The MIT Press  
One Rogers Street  
Cambridge, MA 02142  
tel: 1-617-253-0080

## REFERENCES

- Adams, Frederick. 2010. "Action Theory Meets Embodied Cognition." In *Causing Human Actions: New Perspectives on the Causal Theory of Action*, edited by Jesús H. Aguilar and Andrei A. Buckareff, 229–52. Cambridge, MA: The MIT Press.
- Allport, Floyd H. 1955. *Theories of Perception and the Concept of Structure*. New York: Wiley.
- Anseel, F., and W. Duyck. 2008. "Unconscious Applicants: A Systematic Test of the Name-Letter Effect." *Psychological Science* 19 (10): 1059.
- . 2009. "Implicit Letter Preferences in Job Choice: An Experimental Test of the Role of Cognitive Load." *The Journal of Psychology: Interdisciplinary and Applied* 143 (2): 207–24.
- Aristotle. 1985. *Nicomachean Ethics*. Translated by Terence Irwin. Indianapolis: Hackett.
- Asch, S. E. 1951. "Effects of Group Pressure on the Modification and Distortion of Judgments." *Groups, Leadership and Men*, 177–90.
- Ayer, A. J. 1954. "Freedom and Necessity." *Philosophical Essays*, 271–84.
- Bandura, Albert. 1993. "Perceived Self-Efficacy in Cognitive Development and Functioning." *Educational Psychologist* 28 (2): 117–48.
- Bargh, J. A. 2005. "Bypassing the Will: Toward Demystifying the Nonconscious Control of Social Behavior." In *The New Unconscious*, edited by Ran R. Hassin, James S. Uleman, and John A. Bargh, 37–58. New York: Oxford University Press.
- . 2008. "Free Will Is Un-Natural." In *Are We Free? Psychology and Free Will*, edited by John Baer, James C. Kaufman, and R. F. Baumeister, 128–54. New York: Oxford University Press.
- Bargh, J. A., and T. L. Chartrand. 1999. "The Unbearable Automaticity of Being." *American Psychologist* 54 (7): 462.
- Bargh, John A. 1984. "Automatic and Conscious Processing of Social Information." *Handbook of Social Cognition* 3: 1–43.
- . 1997a. "The Automaticity of Everyday Life." In *Advances in Social Cognition*, 10:1–61.

- . 1997b. “Reply to the Commentaries.” In *Advances in Social Cognition*, 10:1–61.
- Bargh, John A., and Tanya L. Chartrand. 1999. “The Unbearable Automaticity of Being.” *American Psychologist* 54 (7): 462.
- Bargh, John A., and Melissa J. Ferguson. 2000. “Beyond Behaviorism: On the Automaticity of Higher Mental Processes.” *Psychological Bulletin* 126 (6): 925.
- Bargh, John A., Peter M. Gollwitzer, Annette Lee-Chai, Kimberly Barndollar, and Roman Trötschel. 2001. “The Automated Will: Nonconscious Activation and Pursuit of Behavioral Goals.” *Journal of Personality and Social Psychology* 81 (6): 1014.
- Baumeister, R. F., K. D. Vohs, and D. M. Tice. 2007. “The Strength Model of Self-Control.” *Current Directions in Psychological Science* 16 (6): 351–55.
- Baumeister, Roy F. 2004. *Handbook of Self-Regulation: Research, Theory, and Applications*. New York: The Guilford Press.
- . 2008a. “Free Will in Scientific Psychology.” *Perspectives on Psychological Science* 3 (1): 14–19.
- . 2008b. “Free Will, Consciousness, and Cultural Animals.” In *Are We Free? Psychology and Free Will*, by Roy F. Baumeister, John Baer, and James C. Kaufman, 65–85. New York: Oxford University Press.
- Baumeister, Roy F., C. N DeWall, N. J Ciarocco, and J. M Twenge. 2005. “Social Exclusion Impairs Self-Regulation.” *Journal of Personality and Social Psychology* 88 (4): 589–604.
- Baumeister, Roy F., Erin A. Sparks, Tyler F. Stillman, and Kathleen D. Vohs. 2008. “Free Will in Consumer Behavior: Self-Control, Ego Depletion, and Choice.” *Journal of Consumer Psychology* 18 (1): 4–13.
- Bayne, Tim. 2006. “Phenomenology and the Feeling of Doing: Wegner on the Conscious Will.” In *Does Consciousness Cause Behavior? An Investigation of the Nature of Volition*, edited by S. Pockett, W. P. Banks, and S. Gallagher, 169–86. Cambridge, MA: The MIT Press.
- Beaman, A.L., P.J. Barnes, B. Klentz, and B. McQuirk. 1978. “Increasing Helping Rates through Information Dissemination: Teaching Pays.” *Personality and Social Psychology Bulletin* 4 (3): 406.

- Blair, Irene V. 2002. "The Malleability of Automatic Stereotypes and Prejudice." *Personality and Social Psychology Review* 6 (3): 242–61.
- Blair, Irene V., Charles M. Judd, and Jennifer L. Fallman. 2004. "The Automaticity of Race and Afrocentric Facial Features in Social Judgments." *Journal of Personality and Social Psychology* 87 (6): 763.
- Block, Ned. 1995. "On a Confusion about a Function of Consciousness." *Behavioral and Brain Sciences* 18: 227–227.
- Bratman, Michael E. 1996. "Identification, Decision, and Treating as a Reason." *Philosophical Topics* 24 (2): 1–18.
- Chalmers, David J. 1995. "Facing up to the Problem of Consciousness." *Journal of Consciousness Studies* 2 (3): 200–219.
- Chisholm, Roderick. 2001. "Human Freedom and the Self." In *Agency and Responsibility: Essays on the Metaphysics of Freedom*, edited by Ekstrom, Laura Waddell, 126–37. Boulder, CO: Westview Press.
- Clarke, Randolph. 1992. "Free Will and the Conditions of Moral Responsibility." *Philosophical Studies* 66 (1): 53–72.
- . 2003. *Libertarian Accounts of Free Will*. Oxford: Oxford University Press.
- . 2010. "Skilled Activity and the Causal Theory of Action." *Philosophy and Phenomenological Research* 80 (3): 523–50.
- Damasio, Antonio R. 1994. *Descartes' Error: Emotion, Rationality and the Human Brain*. New York: Grosset/Putnam.
- Darley, J. M, and C. D Batson. 1973. "'From Jerusalem to Jericho': A Study of Situational and Dispositional Variables in Helping Behavior." *Journal of Personality and Social Psychology* 27 (1): 100.
- Davidson, Donald. 1963. "Actions, Reasons, and Causes." *The Journal of Philosophy* 60 (23): 685–700.
- Dennett, Daniel C. 1984. *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, MA: The MIT Press.

- . 2008. “Some Observations on the Psychology of Thinking About Free Will.” In *Are We Free? Psychology and Free Will*, edited by John Baer, James C. Kaufman, and Roy F. Baumeister, 248–59. New York: Oxford University Press.
- Devine, Patricia G. 1989. “Stereotypes and Prejudice: Their Automatic and Controlled Components.” *Journal of Personality and Social Psychology* 56 (1): 5.
- Doris, J.M. 2002. *Lack of Character: Personality and Moral Behavior*. Cambridge, MA: Cambridge University Press.
- Doris, J.M., and D. Murphy. 2007. “From My Lai to Abu Ghraib: The Moral Psychology of Atrocity.” *Midwest Studies in Philosophy* 31 (1): 25–55.
- Double, R. 1991. *The Non-Reality of Free Will*. New York: Oxford University Press.
- Eisenberg, N., C. L. Smith, A. Sadovsky, and T. L. Spinrad. 2004. “Effortful Control: Relations with Emotion Regulation, Adjustment, and Socialization in Childhood.”
- Ekstrom, L. 2002. “Libertarianism and Frankfurt-Style Cases.” In *The Oxford Handbook of Free Will*, edited by Robert Kane, 309–22. New York: Oxford University Press.
- Erdelyi, Matthew H. 1974. “A New Look at the New Look: Perceptual Defense and Vigilance.” *Psychological Review* 81 (1): 1.
- Fischer, J. M., and M. Ravizza. 2000. *Responsibility and Control: A Theory of Moral Responsibility*. New York: Cambridge University Press.
- Frankfurt, Harry G. 1969. “Alternate Possibilities and Moral Responsibility.” *The Journal of Philosophy* 66 (23): 829–39.
- . 1971. “Freedom of the Will and the Concept of a Person.” *The Journal of Philosophy* 68 (1): 5–20.
- . 1988. “Identification and Wholeheartedness.” In *The Importance of What We Care About*, 159–76. New York: Cambridge University Press.
- Gaertner, Samuel L., and John P. McLaughlin. 1983. “Racial Stereotypes: Associations and Ascriptions of Positive and Negative Characteristics.” *Social Psychology Quarterly*, 23–30.



- Gailliot, M. T., R. F. Baumeister, C. N. DeWall, J. K. Maner, E. A. Plant, D. M. Tice, L. E. Brewer, and B. J. Schmeichel. 2007. "Self-Control Relies on Glucose as a Limited Energy Source: Willpower Is More than a Metaphor." *Journal of Personality and Social Psychology; Journal of Personality and Social Psychology* 92 (2): 325.
- Gilovich, Thomas. 1991. *How We Know What Isn't So: The Fallibility of Reason in Everyday Life*. New York: Free Press.
- Ginet, Carl. 1990. *On Action*. New York: Cambridge University Press.
- Gollwitzer, Peter M. 1999. "Implementation Intentions: Strong Effects of Simple Plans." *American Psychologist* 54 (7): 493.
- Haney, C., C. Banks, and P. Zimbardo. 1973. "Interpersonal Dynamics in a Simulated Prison." *International Journal of Criminology & Penology*.
- Harris, Christine R., Noriko Coburn, Doug Rohrer, and Harold Pashler. 2013. "Two Failures to Replicate High-Performance-Goal Priming Effects." *PLoS ONE* 8 (8): e72467.
- Harris, Sam. 2012. *Free Will*. New York: Free Press.
- Hume, David. 2000. *An Enquiry Concerning Human Understanding: A Critical Edition*. Vol. 3. Oxford University Press.
- Jones, J. T, B. W Pelham, M. Carvallo, and M. C Mirenberg. 2004. "How Do I Love Thee? Let Me Count the Js: Implicit Egotism and Interpersonal Attraction." *Journal of Personality and Social Psychology* 87 (5): 665.
- Jones, J. T, B. W Pelham, M. C Mirenberg, and J. J Hetts. 2002. "Name Letter Preferences Are Not Merely Mere Exposure: Implicit Egotism as Self-Regulation." *Journal of Experimental Social Psychology*.
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Macmillan.
- Kane, R. 1996. *The Significance of Free Will*. New York: Oxford University Press.
- . 1999. "Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism." *The Journal of Philosophy* 96 (5): 217–40.

- Kihlstrom, John F. 2008. "The Automaticity Juggernaut—or, Are We Automatons after All." In *Are We Free? Psychology and Free Will*, edited by John Baer, James C. Kaufman, and Roy F. Baumeister, 155–80. New York: Oxford University Press.
- Latane, B., and J. Rodin. 1969. "A Lady in Distress: Inhibiting Effects of Friends and Strangers on Bystander Intervention\* 1." *Journal of Experimental Social Psychology* 5 (2): 189–202.
- Levy, Neil. 2011. "Expressing Who We Are: Moral Responsibility and Awareness of Our Reasons for Action." *Analytic Philosophy* 52 (4): 243–61.
- Li, W., I. Moallem, K. A Paller, and J. A Gottfried. 2007. "Subliminal Smells Can Guide Social Preferences." *Psychological Science* 18 (12): 1044.
- Libet, B. 1999. "Do We Have Free Will?" *Journal of Consciousness Studies*, 6 8 (9): 47–57.
- Lord, Charles G., Mark R. Lepper, and Elizabeth Preston. 1984. "Considering the Opposite: A Corrective Strategy for Social Judgment." *Journal of Personality and Social Psychology* 47 (6): 1231.
- Malle, B. F. 2006. "Of Windmills and Strawmen: Folk Assumptions of Mind and Action." In *Does Consciousness Cause Behavior? An Investigation of the Nature of Volition*, edited by S. Pockett, W. P. Banks, and Shaun Gallagher, 207–31. Cambridge, MA: The MIT Press.
- Marzollo, Jean. 1992. *I Spy: A Book of Picture Riddles*. Scholastic Publications.
- Mele, Alfred R. 1987. *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*. New York: Oxford University Press.
- . 1992. *Springs of Action: Understanding Intentional Behavior*. New York: Oxford University Press.
- . 1995. *Autonomous Agents: From Self-Control to Autonomy*. New York: Oxford University Press.
- . 2003a. "Agents' Abilities." *Nous* 37 (3): 447–70.
- . 2003b. *Motivation and Agency*. New York: Oxford University Press.

- . 2006. *Free Will and Luck*. New York: Oxford University Press.
- . 2007. “Persisting Intentions.” *Noûs* 41 (4): 735–57.
- . 2008a. “Psychology and Free Will: A Commentary.” In *Are We Free? Psychology and Free Will*, edited by John Baer, James C. Kaufman, and Roy F. Baumeister, 325–46. New York: Oxford University Press.
- . 2008b. “Recent Work on Free Will and Science.” *American Philosophical Quarterly* 45 (2): 107–30.
- . 2009. *Effective Intentions: The Power of Conscious Will*. New York: Oxford University Press.
- Mele, Alfred R., and Joshua Shepherd. 2013. “Situationism and Agency.” *Journal of Practical Ethics* 1: 62–83.
- Milgram, S. 2009. *Obedience to Authority: An Experimental View*. New York: Harper Perennial Modern Classics.
- Milgram, Stanley. 1974. *Obedience to Authority*. New York: Harper & Row.
- Moors, Agnes, and Jan De Houwer. 2007. “What Is Automaticity? An Analysis of Its Component Features and Their Interrelations.” In *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes*, edited by Bargh, John A., 11–50. New York: Psychology Press.
- Moran, Richard. 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press.
- . 2002. “Frankfurt on Identification: Ambiguities of Activity in Mental Life.” In *Contours of Agency: Essays on Themes from Harry Frankfurt*, 189–217. Cambridge, MA: MIT Press.
- Moskowitz, Gordon B., Peter M. Gollwitzer, Wolfgang Wasel, and Bernd Schaal. 1999. “Preconscious Control of Stereotype Activation through Chronic Egalitarian Goals.” *Journal of Personality and Social Psychology* 77 (1): 167.
- Nahmias, Eddy. forthcoming. “The Psychology of Free Will.” *The Oxford Handbook on the Philosophy of Psychology*. Oxford University Press, Oxford

- . 2002. “When Consciousness Matters: A Critical Review of Daniel Wegner’s The Illusion of Conscious Will.” *Philosophical Psychology* 15 (4): 527–41.
- . 2007. “Autonomous Agency and Social Psychology.” In *Cartographies of the Mind: Philosophy and Psychology in Intersection*, edited by M. Marraffa, M. Caro, and F. Ferretti, 169–85. Springer.
- Neisser, Ulric. 1988. “Five Kinds of Self-Knowledge.” *Philosophical Psychology* 1 (1): 35–59.
- Nelkin, D. K. 2005. “Freedom, Responsibility and the Challenge of Situationism.” *Midwest Studies in Philosophy* 29 (1): 181–206.
- Nickerson, Raymond S. 1998. “Confirmation Bias: A Ubiquitous Phenomenon in Many Guises.” *Review of General Psychology* 2 (2): 175.
- Nisbett, R. E, and T. D Wilson. 1977. “Telling More than We Can Know: Verbal Reports on Mental Processes.” *Psychological Review* 84 (3): 231.
- Nuttin Jr, J. M. 1985. “Narcissism beyond Gestalt and Awareness: The Name Letter Effect.” *European Journal of Social Psychology* 15 (3): 353–61.
- O’Connor, Timothy. 2002. *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press.
- . 2009. “Degrees of Freedom.” *Philosophical Explorations* 12 (2): 119–25.
- Pacherie, Elisabeth. 2006. “Toward a Dynamic Theory of Intentions.” In *Does Consciousness Cause Behavior? An Investigation of the Nature of Volition*, edited by S. Pockett, W. P. Banks, and Shaun Gallagher, 145–67. Cambridge, MA: The MIT Press.
- Pajares, Frank, and M. David. 1994. “Role of Self-Efficacy and Self-Concept Beliefs in Mathematical Problem Solving: A Path Analysis.” *Journal of Educational Psychology* 86 (2): 193–203.
- Pelham, B. W, M. C Mirenberg, and J. T Jones. 2002. “Why Susie Sells Seashells by the Seashore: Implicit Egotism and Major Life Decisions.” *Journal of Personality and Social Psychology* 82 (4): 469.
- Pereboom, Derk. 2001. *Living Without Free Will*. Cambridge: Cambridge University Press.

- Pronin, Emily, Daniel M. Wegner, Kimberly McCarthy, and Sylvia Rodriguez. 2006. "Everyday Magical Powers: The Role of Apparent Mental Causation in the Overestimation of Personal Influence." *Journal of Personality and Social Psychology* 91 (2): 218.
- Schneider, Walter, and Jason M. Chein. 2003. "Controlled & Automatic Processing: Behavior, Theory, and Biological Mechanisms." *Cognitive Science* 27 (3): 525–59.
- Schneider, Walter, and Richard M. Shiffrin. 1977. "Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention." *Psychological Review* 84 (1): 1.
- Shepherd, Joshua. 2012. "Free Will and Consciousness: Experimental Studies." *Consciousness and Cognition* 21 (2): 915–27.
- Slattery, Trick. 2014. *Breaking the Free Will Illusion for the Betterment of Humankind*. 1 edition. Working Matter Publishing.
- Sternberg, Eliezer J. 2010. *My Brain Made Me Do It: The Rise of Neuroscience and the Threat to Moral Responsibility*. Amherst, N.Y: Prometheus Books.
- Stillman, T. F., R. F Baumeister, and A. R Mele. 2011. "Free Will in Everyday Life: Autobiographical Accounts of Free and Unfree Actions." *Philosophical Psychology* 24 (3): 381–94.
- Strawson, G. 1986. *Freedom and Belief*. Oxford: Oxford University Press.
- Ulrich, Roger. 1984. "View through a Window May Influence Recovery." *Science* 224 (4647): 224–25.
- Van Inwagen, Peter. 1983. *An Essay on Free Will*. Oxford: Clarendon Press.
- Vargas, Manuel. forthcoming. "Situationism and Moral Responsibility: Free Will in Fragments." In *Decomposing the Will*. New York: Oxford University Press.
- Velleman, J. David. 2002. "Identification and Identity." In *Contours of Agency: Essays on Themes from Harry Frankfurt*, 91–123. Cambridge, MA: MIT Press.

- Vohs, K. D., R. F. Baumeister, B. J. Schmeichel, J. M. Twenge, N. M. Nelson, and D. M. Tice. 2008. "Making Choices Impairs Subsequent Self-Control: A Limited-Resource Account of Decision Making, Self-Regulation, and Active Initiative." *Journal of Personality and Social Psychology* 94 (5): 883.
- Wallace, R. J. 1994. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.
- Wason, Peter C. 1968. "Reasoning about a Rule." *The Quarterly Journal of Experimental Psychology* 20 (3): 273–81.
- Watson, Gary. 1975. "Free Agency." *The Journal of Philosophy* 72 (8): 205–20.
- Wegner, Daniel M. 2002. *The Illusion of Conscious Will*. Cambridge, MA: The MIT Press.
- . 2005. "Who Is the Controller of Controlled Processes?" In *The New Unconscious*, edited by Ran R. Hassin, James S. Uleman, and John A. Bargh, 19–36. New York: Oxford University Press.
- . 2008. "Self Is Magic." In *Are We Free? Psychology and Free Will*, edited by John Baer, James C. Kaufman, and Roy F. Baumeister, 226–47. New York: Oxford University Press.
- Wegner, Daniel M., Betsy Sparrow, and Lea Winerman. 2004. "Vicarious Agency: Experiencing Control over the Movements of Others." *Journal of Personality and Social Psychology* 86 (6): 838.
- Wegner, Daniel M., and Thalia Wheatley. 1999. "Apparent Mental Causation: Sources of the Experience of Will." *American Psychologist* 54 (7): 480.
- Whittle, A. 2010. "Dispositional Abilities." *Info: December 2010* 10 (12).
- Wolf, S. 1993. *Freedom within Reason*. New York: Oxford University Press.

## **BIOGRAPHICAL SKETCH**

Zachary Thomas Martin was born and raised in central Ohio. He attended Bowling Green State University, graduating in 2006 with a B.A. in Philosophy and a minor in Sociology. Zachary entered graduate school at Florida State University in 2006, earning his Master's degree in Philosophy in 2009. His philosophical interests include free will, philosophy of mind and action, philosophy of psychology, and Nietzsche. Outside of philosophy, Zachary enjoys backpacking, running, competing in triathlons, rock climbing, chess, and playing with his dog, Rupert.