AN ABSTRACT OF THE THESIS OF

Erik G. Helzer for the degree of Honors Baccalaureate of Arts in Philosophy and Psychology presented on November 29, 2005. Title: Beyond Free Will: An Empirical Case for a Capacity Model of Human Agency.

Abstract approved:

_____
John Edwards

Although the question of free will has persisted throughout the history of philosophy as an important metaphysical consideration, rarely does the inquiry include empirical evidence from other disciplines concerned with human behavior. In this interdisciplinary paper, I combine the wisdom of psychologists, metaphysicians, and philosophers of science to comment on the problem. Research on cognitive organization, priming, spreading activation, and attribution is reviewed, and I challenge the notion of autonomous controlled processing. I argue from empirical psychological evidence that the concept of free will is an incoherent model for understanding the complex interrelatedness of human behavior. It amounts to little more than an historical placeholder that should be replaced by a descriptive framework rooted in science. I suggest as a replacement a capacity model of human agency, which allows for active, but not free, agents. Important metaphysical concepts such as personal identity, control, and responsibility can be maintained under this new framework and are, in fact, superior insofar as they are rooted in an accurate description of human functioning.

Beyond Free Will:

An Empirical Case for a Capacity Model of Human Agency

by

Erik G. Helzer

A PROJECT

submitted to

Oregon State University

University Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Arts in Philosophy and Psychology (Honors Scholar)

Presented November 29, 2005
Commencement June 2006

<u>Honors Baccalaureate of Arts in Philosophy and Psychology</u> project of <u>Erik G. Helzer</u> presented on <u>November 29, 2005.</u>

APPROVED:

_____

Mentor, representing Psychology

_____

Committee Member, representing Philosophy

_____

Committee Member, representing Philosophy

_____

Chair, Department of Psychology

_____

Chair, Department of Philosophy

_____

Dean, University Honors College

I understand that my project will become part of the permanent collection of Oregon State University, University Honors College. My signature below authorizes release of my project to any reader upon request.

_____

Erik G. Helzer, Author

ACKNOWLEDGEMENT

I am grateful to my academic mentors who have been vital to my growth as a student. Thank you to those professors who excited me to learn and challenged me to give it my all. I will remember with fondness Professor Bill Uzgalis, Professor Eric Hill, Professor Jennifer Connor-Smith, and Professor Jim Blumenthal.

Thank you to my helpful, supportive, and enthusiastic thesis committee—the ones who kept me writing and thinking: Professor John Edwards, Professor Flo Leibowitz, and Professor Sharyn Clough.

And to my those who have questioned me, encouraged me, and above all, cared for me, making it all worth while, I extend an appreciative hug and a big sigh of relief: my family, my friends, and the symposium cohort.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

DEDICATION

After two years of thinking about what makes a person a person and writing about the

(often un- or under-recognized) importance of each interaction, I have come to appreciate

to a great degree all those who have brought me to this point. This thesis is dedicated to

my parents.

Beyond Free Will:

An Empirical Case for a Capacity Model of Human Agency

INTRODUCTION

The human organism resides in a sea of information. Research suggests that in any one second, the eye transmits around one million bits of information to the human brain, the ear and nose contribute one hundred thousand pieces of data each, and so on for each sensory apparatus (Nørretranders 1991). From this vast body of stimuli, the brain extracts and processes only a very small portion, and we, as conscious beings, are aware of only about five percent of that sampling (Bargh & Chartrand 1997). The conscious experience is a vastly incomplete and potentially deceptive one, akin to watching the action of Wall Street through a small keyhole.

What is incredible, actually, is how clear of a picture we seem to get from our conscious experience despite the fact that almost all processing and activity happens beyond awareness, and how confident we are about the accuracy of our conscious experience. We attribute a number of advantages to this evolutionary phenomenon: self-awareness, planning, and high-order cognitive functioning. In addition, we believe because of our conscious experience that we are distinct from the rest of the natural world in our ability to reason and make decisions. Most people believe in free will as a result of the conscious experience, too. We believe that our self-awareness gives us the ability to freely choose the course of our lives, to separate ourselves, as it were, from the environment around us, and to exert our will wherever we deem it appropriate.

In this paper, I challenge these beliefs about our conscious experience of free will. The conscious experience, I argue, that keyhole view into our psychology that gives us the *feeling* of free will, is a self-reinforcing phenomenon. Because we only have access to processing that happens on a conscious level, we are unaware of what lies beyond awareness and, more importantly, ignorant of those factors that shape the conscious experience. We therefore assume that all behaviors, thoughts, and emotions are generated consciously. I begin this investigation by reviewing evidence that the conscious experience is indeed shaped by factors that lie outside of the grasp of our awareness. These internal and external factors are constantly at play and very rarely do we realize the extent of their effects.

In addition to internal processing that goes on beyond awareness, the environment within which we operate is continuously providing cues that interact constantly with our conscious and non-conscious experience so that we may function effectively out in the world. We are not usually aware of our perceptions of these cues; we are even less aware of how they are accommodated into our ideas about the world and alter our behavior. I present research about the effects of environmental and social cues on perception and judgment to show that we humans are beings of interaction in the world. Our internal states interact causally with external phenomena, and these external factors, in turn, affect our cognitive-emotional experience, our perceptions, and our judgments.

If this is the case, I ask, how could we have the possibility of free will? We might conclude from all of this that the idea of free will is incoherent, and that at the very best, a "free" conscious decision is something akin to reaching our arm into a deep sack full of

prizes and pulling out the one that seems unique, to the exclusion of other prizes whose use and identity remain unknown.

Alternatively, we might focus not on the hand that grasps the prize, but on the relationship between our hand, the prizes, the bag, and the act of grabbing. This is the view I support: that the prize we eventually win becomes less important than the interactive process that results in grasping the prize. I suggest that the idea of free will, like the hand grasping the prize, is an inarticulate conceptualization of the actual process of human functioning. In what follows, I will address the philosophical problem of free will from the standpoint of empirical psychology.

An Overview of the Paper

In section one, I present the distinction between automatic and controlled processing as a tool for understanding our various mental activities. Automatic processing happens below the level of conscious awareness, and, due to its efficiency, makes up the bulk of our mental activity. Controlled processing, by comparison, happens much less frequently, but because it occurs consciously, we have some degree of access into its workings. I discuss the interaction between these two forms of processing and review research revealing that automatic, non-conscious mental activity underpins our controlled processing and outward behaviors in ways that we do not consciously realize. Indeed, all information that comes to conscious awareness is first processed and sorted automatically, such that it arrives to the stage of consciousness packaged in predefined meaning.

I then present a host of studies that illustrate the powerful effect of environmental cues on shaping thoughts and behaviors. I will explore the automatic

pathways of thinking and reflect upon the consequences of our cognitive organization, which is a network of schemas, or ideas, that are activated based upon internal (emotional and cognitive) and external (environmental) cues. Since these cues are sensed and processed regardless of whether or not we are conscious of them, many of our conscious thoughts, behaviors, and actions, even those labeled "controlled" phenomena, are triggered automatically and in interdependence with our internal and external environment.

This discussion leads naturally to the conclusion that the distinction between automatic and controlled processing is flexible and overlapping. Indeed, the term "controlled" processing may be deceptive given the amount of automaticity that occurs behind the scenes. It may actually be the case that level of awareness is a better taxonomic tool.

I then question the conscious mind's ability to offer an accurate picture of the world, which will set us up for understanding how it is that we can arrive at mistaken conclusions about how our own minds work and the degree to which our conscious thoughts are causally efficacious. I present data showing that human beings are relatively poor judges of causes and that our conscious attributions about the world can be manipulated through environmental and other factors without our knowing it. We very regularly offer causal stories about the world that turn out to be wrong, and we overestimate the degree to which our thoughts correspond to reality. These errors have drastic consequences for the problem of free will because they suggest that despite the feeling that we are acting autonomously and freely, our feelings and conscious attributions about our own ability to be "self-caused" probably have little to do with our

actual causal efficacy in the world. This distinction between the inert *feeling* of free will

and the *efficacy* of free will becomes important to the view I offer here.

Based upon the above empirical evidence, it may come as little surprise that I

reject the notion of free will. That is, I deny that there is any construct of "will" that is

self-caused, independent, and the source of human agency. By extension, then, I

disallow that human beings are self-caused, independent agents. Empirical data aside, I

hold that free will is an incoherent concept; a mysterious placeholder that may have

been historically useful in a pre-scientific world, but that is outdated given our current

knowledge of human functioning. Philosophers who argue for its existence have done

little to explain what, in fact, free will is and how it operates. Instead, the construct is

almost always –implicitly or explicitly- backed circularly by experience: "I did this

action in accordance with my own free will, and that is proof that I have it." The other

non-empirical route for defense of human liberty is to use some form of straw-person

argument: "free will must exist in order for us to be responsible for our actions." I

discuss these problems later on in section two.

Before concluding section one, though, I offer two different, but complementary,

theories for the illusory experience of free will. Daniel Wegner and Thalia Wheatley's

(1999) theory of apparent mental causation offers a model from which to understand the

power of the illusion. They hypothesize that our experience of conscious will is

reinforced by the fact that we are limited observers of our own behavior. We

consciously experience thoughts and we also consciously experience behaviors and we

make causal connections suggesting that the former activates the latter, even though

such connections are not necessarily there. I offer additional philosophical support of

Wegner and Wheatley by way of David Hume, whose account of causation nicely describes the ease with which human beings can misjudge causes.

Second, I review Efran, Lukens, and Lukens' (1990) structural determinism thesis. According to these researchers, human beings are complex, dynamic, and self-organized entities who are in constant interaction with, and are, in fact, an indivisible part of, their environment. Causal stories about the world are arbitrarily drawn insofar as they isolate interactions apart from their ecological existence. Our idea of free will represents a causal story we tell in which we separate ourselves from the interactions with which we engage. The structural determinism of Efran and colleagues posits that we construct causal stories by freeze-framing parts of an indivisible stream of events. Free will is a story told from the first-person perspective.

Section two responds to the question of "If not free will, then what?" Traditionally, the philosophical answer defaults to determinism. If we are not free, then we are controlled by external causes. This is usually where defenders of free will, known as libertarian philosophers, object to a view of humanity in which people are reduced to mere cogs in a machine with no purpose in life, no morality, and no ultimate responsibility. Not only do I challenge these philosophers' claims that free will is the only way to establish these important metaphysical concepts, I wholly doubt that free will is the best way to go about it, if for no other reason than that libertarian foundationalism is illusory. Throughout the process of this investigation I have also come to question whether the polarized debate of "either free will or determinism" (with the middle-ground stance of compatibilists, who believe that free will and responsibility are compatible with determinism) is helpful or accurate, given our

understanding of cause and effect relationships. And so, in section two I begin with a review of philosophical theories of causation.

The third theory presented begins with John Dupré's (1993) treatment of causation, which offers an historical perspective of the philosophical issues. Laplacean determinism, the linear, one-way, mechanistic brand of determinism, the consequences of which libertarian philosophers regard as little better than Sisyphean, has largely been abandoned by philosophers interested in causation because of its metaphysical and methodological short-comings. As a result, many philosophers have taken solace in probabilistic theories of causation, which posit causes as those things that, in all contexts, increase the probability of the occurrence of a certain effect. Dupré offers reasons to reject this form of causation, as well, primarily because of the problem of universal context, which he views as an unreachable goal. He suggests as a replacement a capacity model of causation, in which factors become causes when and if their capacity to cause is actualized by the context of the interaction.

The capacity model denies the existence of *a priori* causes and instead requires that causal relationships are interdependent. That is, in order for a factor to act causally, it must be supported by a causal context, which includes the potential-effect factor and all other factors involved. Each factor in an interaction has a capacity to affect causal relationships but only does so when other factors support the relationship both by increasing its chance of occurrence and by not interfering with its chance of occurrence. Dupré conceptualizes people, then, as bundles of capacities. The thing that makes human beings unique is our extraordinary level of capaciousness: our ability to interact with the environment in varied and malleable ways.

I extend Dupré's model to construct a capacity theory of agency that fits with the empirical conclusions about the human condition reached in section one. I reason that if human beings are bundles of capacities and capacities act only in contextual interdependence with other factors, then human beings must also be bundles of receptors. Our agency, our capacity to affect change in the world, is powered not by "will" but rather by our constant interrelation with the contextual world. That is, we are efficacious only insofar as we are receptive products of causes and conditions ourselves. I show that this viewpoint is alluded to by a social psychological view of persons, and that all that is needed is to make explicit the similar conclusions of both positions. The capacity model may be a useful tool for explaining certain empirical phenomena, as well as a helpful theoretical framework for understanding the complexity and dynamical interactions that are the objects of interest in the study of the human mind.

In the remainder of section two, I address the constructs that libertarians claim are lost when free will is rejected. Three constructs are considered: personal identity, control, and responsibility. For each of these, I attempt to show that our ideas about each, including their metaphysical foundations and consequences, are situated within cultural and rhetorical contexts. Moreover, I claim that, in practice, our attitudes about each of these are flexible and tend to change, to a greater or lesser degree, to suit our needs. I argue that the presumption of free will is unnecessary for establishing these important metaphysical constructs. I investigate alternative conceptions of personal identity, control, and responsibility, using theoretical frameworks that do not share the same individualistic preference that is embodied in our cultural ideals. With regard to responsibility, I show that freedom of the will is neither a necessary nor sufficient condition for accountability, but rather that active, rather than passive, agency is the

requirement. Under a capacity model of human agency, responsibility can be achieved in a way that fosters greater social accountability.

In my concluding remarks I recognize the limitations that are present when the study of humans is conducted by humans. For this reason, I advocate an interdisciplinary approach to questions about the human condition as a means of checks and balances for teasing out some of the disciplinary biases that color our treatment of any phenomenon. Seen from the standpoint of empirical psychology, I reiterate that the traditional philosophical concept of free will is both illusory and not useful for understanding human behavior and the rich and dynamic stream of interactions that comprise human experience. I urge that the question of free will verses determinism can serve to blind researchers to asking the more important question: how is it that people actually interact in the world? Such open inquiry leads to non-binary, interesting, and plausible accounts of the human condition.

SECTION 1: EMPIRICAL FINDINGS OF PSYCHOLOGY

Reformulating the Problem and Clarifying Terms

In order to proceed, a language common to both the philosophical and psychological inquiries must be set. We must, for the sake of argument, understand the empirical conditions that would need to be met in order to support philosophical libertarianism. To this end, I would like to substitute the old conceptual debate of philosophers with a nuanced theoretical alternative better suited for empirical investigation. This redefining of terms allows for a more specific reformulation of the philosophical problem; moreover, it sets the appropriate burden of proof for establishing free will. In other words, I will provide the psychological language that allows us to maneuver within the philosophical structure.

We know much more about the brain and its functions now than at any time in the past. Indeed, our rapidly expanding knowledge of neurophysiology and neurobiology continues to inform our understanding of what it is to be a thinking being. The practical effect of this expansion is that many folk theories of brains and behavior have run their course, replaced by empirical, physical explanations. For example, we no longer suppose, as Descartes did in the 17th century, that "animal spirits," in connection with the pineal gland, bring about mental behavior.

In fact, given the scope and progress of neuroscience, the question of "determinism" may seem largely settled. Most reasonable people with only slight knowledge of neuroscience, for example, would assent to the fact that we can identify and posit physical and chemical causes for each behavior we observe in others and ourselves. The act of raising my left arm can be explained by physical causation—we

can see corresponding brain activity in the areas associated with control of the left side of the body; we can discuss action potentials and the shifting levels of sodium and calcium; and we can identify neurological pathways linking the brain to the rest of the body. That there are physical causes of behavior that originate in the brain is a matter of little contention. That many of these processes happen without our knowledge of them seems to indicate a physical deterministic system underlying all behavior.

In this sense, the question of determinism is settled. The question of free will, however, remains. When we attempt to polarize the debate as free will versus determinism, we must understand both of these terms on the basis of their metaphysical significance. When we speak of the former, what we really mean is a kind of meta-determinism, a mind-determinism. That is, in order to support the notion of free will we are not interested in whether all behavior has physical causes, but whether there is some entity, a "will," that operates independently of these physical explanations, such that it is *free* to make choices, while at the same time retains its ability to affect, that is make changes in, the physical processes. We are interested in whether mental activity (which I am distinguishing from brain activity) has the efficacy to cause brain activity. We are also interested in whether our predominantly mental behaviors, things like imagination, rational thought, and planning, can operate in some sense independently of the brain states—that is, whether the activities of the mind cannot be explained completely by physical and chemical processes which happen prior to corresponding changes in mental behavior.

In this paper, I am presupposing the construct of mind, a conscious realm of activity whose origin cannot be reduced simply to one physical area of the brain. This is by no means an obvious premise: as my professor, Bill Uzgalis has said, there was a

time when philosophers were sure they had minds but weren't at all sure that they had bodies; now philosophers are willing to admit bodies, but not at all willing to assume the existence of minds. Nonetheless, I am making this presupposition for a number of reasons, not least of which is the principle of charity: I consider it a necessary premise if we are to grant any validity to the doctrine of free will. Without the construct of mind as it is defined here, the human experience is one of physical processes and two-way brain-body determinism. The only possibility for free will within this system resides in the hope for indeterminism; however, even the prospect of randomness provides little comfort, for reasons I explain below.

Second, that there is conscious mental activity is a matter of common human experience. We may theorize about the underpinnings of mental events, but the overwhelming and exhilarating realm of consciousness is one that I would be premature to disregard as inconsequential, especially given the relative lack of progress psychologists and philosophers have made in attempting to shed light on this phenomenon. It very well may be that one day we will be able to explain consciousness in terms of physical processes, and I imagine the future model will reveal consciousness as several different brain pathways operating in concert with one another to create a metaphenomenon that is somehow more interesting than the sum of its parts. When this paradigm arrives, we will be faced with newer and more detailed questions about the present debate. However, for now, the existence of mind, if we define it as a conscious realm of activity whose origin is still unclear, is a premise I am willing to grant. That said, I would be just as willing to revoke its validity once our understanding of consciousness improves.

The philosophical problem of free will, then, must be formulated as a question about the causal efficacy of the conscious mind. There are three separate considerations:

1. Are the processes that happen at the conscious level self-controlled, that is, is mental activity controlled by mental activity, independent or semi-independent of physical processes?

2. If mental processes are controlled by other mental processes or environmental cues, can we then say that these activities are *freely chosen*?

3. Can these mental processes interact with physical processes; is the mind causally efficacious?

When we consider question 2 above, we see that the problem of the self becomes essential. I cannot help but visualize the analogy of the peach and the onion. Depending on the analogy we use to conceptualize the mind, we are led to certain conclusions about the problem of free will. If, for example, mental processes are controlled by other mental processes, and we conceptualize the mind as a peach, then free choice happens within the pit—the core source of mental activity (the "self") that creates the workings of the mind and, perhaps, acts upon the physical processes of the brain. If, though, the mind is conceptualized as an onion, free will seems impossible. The conscious mind is nothing other than the ebb and flow of various mental activities, layers upon layers of thoughts and emotions that are activated either deterministically or indeterministically. A final conception of mind might be a flashlight model, in which consciousness is the roaming flashlight whose attention is directed by the physical and environmental circumstances of the individual, thus shedding light upon certain brain activities at various times (Trincker 1965 in Nørretranders 1991).

In any event, we are certainly interested in conscious processing—how information reaches consciousness, how consciousness works as a processing tool, and how consciousness relates to the physical brain and body. Specifically, we are interested in two cognitive routes: automatic processing and controlled processing (Posner & Snyder 1975). Mental phenomena that occur reflexively, below the level of conscious awareness are known as automatic processes. Psychologists characterize these kinds of mental activity as *if-then* relations: if the appropriate preconditions exist, then the automatic process will take place (Bargh 1997). Automatic processes dominate mental activity because they require no conscious attention; they are unintentional, uncontrollable, and efficient (Shiffrin & Schneider 1977). Controlled processing, by comparison, requires effortful conscious attention, and because of its inefficiency, occurs far less frequently. Controlled processing is intentional and deliberate, and can only occur when sufficient cognitive resources are available.

In addition to the standard-cited differences in efficiency and effort that separate automatic processes from controlled ones, Timothy Wilson (2002) lists a number of other attributes. Automatic processing, which takes place in Wilson's model of the "adaptive unconscious," involves multiple systems that span a number of different mental and behavioral abilities; such processes are adept at identifying and fitting incoming stimuli to established patterns; they are concerned with the present moment; and Wilson suggests that automatized processing is especially sensitive to negative information. As a result of its evolutionary primacy and its efficiency, automatized processing tends to be rigid and slower to accommodate data that conflict with established mental patterns.

Wilson identifies the following characteristics of controlled (conscious) processing. Conscious processes can only handle one task at a time, and as such,

consciousness represents a single system whose primary concern is checking and balancing information interpreted by rapid and sometimes imprecise automatized pattern-detectors. In other words, conscious processes function as "override mechanisms" (Baumeister & Sommer 1997) that interfere with and override undesired predictable automatized processing and behavior (see also Hefferline, et al. 1959). Finally, Wilson's model identifies conscious processing as that which incorporates past, present, and future into decision-making, and suggests that these processes may show greater sensitivity to positive information. As an evolutionary add-on, consciousness develops more slowly within the individual (a theory of mind, for example, has been shown to develop in children around age three or four, whereas automatic, reflexive processing is active *in utero*), which makes it at once less intuitive and slower (relative to automatic processing), but also more flexible and accommodating of new information.

These two cognitive processes interact with one another. Most of our so-called "higher-order" mental events and behaviors rely upon both automatic and controlled processing. That is, most behaviors alternate between conscious, effortful behavior and non-conscious, habituated behavior. Walking, for example, is a behavior that is learned by controlled processes. A toddler attempting to ambulate from one end of the room to the other must exercise effortful and deliberate attention to her movements and to the environment in which she travels. As the toddler becomes practiced at walking, though, the act becomes more and more familiar and automatic. By the time the toddler reaches college, say, she can walk around campus with little or no conscious effort and can allocate such controlled processing to more pertinent matters. However, if situational factors dictate the need for conscious attention (there are icy patches on the sidewalk, her shoe laces are tied together, or her legs have fallen asleep), walking can once again

become an effortful and careful behavior. Almost all behaviors in which we engage on a regular basis are automatized. Only rarely do we allocate conscious attention to their doing.

In terms of a common language, then, philosophical determinism is the condition of automatized processing. For obvious reasons, those behaviors of which we have no awareness cannot be said to be consciously willed. Free will, if it exists at all, must reside within the realm of the controlled processing of the conscious mind and must in some sense be independent of automatic processes. I will proceed, then, first by situating these two psychological processes in a useful, historical analogy, and then by examining these two cognitive processes, focusing specifically on the overlap between them. As we will see, controlled processes, though they are conscious, are by no means free of automaticity. In fact, the very small amount of conscious processing that does occur is, in large part, regulated by, and thus subjected to the preconditions of automatic processes.

Freud's Iceberg and Controlled Processing as an Exception

When Freud first introduced his theory of the unconscious mind, he used the analogy of an iceberg. Consciousness was for Freud, as it is for many researchers in automaticity today, only a small window into the workings of the mind; it was only the tip of the iceberg (according to Wilson (2002), consciousness might better be understood in this analogy as "the… snowball on top of that iceberg" [p. 6]). Freud believed that the vast majority of mental processes occur under the surface of the water, below the level of conscious awareness (Corey 2001). While many psychologists have since moved on from Freud and his theories of personality to a greater or lesser degree, his basic

assumptions about the structure of the mind predicted the empirical hypotheses and findings of current research psychologists.

First, Freud's conjecture about the proportionality of unconscious and conscious mental life is accepted by many of today's cognitive researchers. Automaticity researchers commonly agree that about 95 percent of all mental activity occurs automatically and below the level of consciousness (Bargh & Chartrand 1999). Second, it is useful for our thinking that Freud's iceberg model connotes a deceptive element: at sea level, we see only the tip of the iceberg and may confuse the visible chunk for the iceberg itself. Because we only have the experience of conscious, controlled processing, we assume that consciousness is the rule and not the exception. Finally, Freud theorized that consciousness was the window into unconscious processes, or put another way, that unconscious processes "poke through" into conscious awareness. This premise, too, is supported by our physical model of the stimuli pathway to conscious awareness (outlined below). Indeed, as research continues, automatic (un- or non-conscious) underpinnings of consciously-experienced phenomena are being identified: psychologists recognize influential automatic processes underlying perception, trait categorization, stereotyping, and attitude formation (Bargh 1997), aggression (Berkowitz 1997), emotion (Clore & Ketelaar 1997), and virtually every other behavior (mental or otherwise) in which human beings engage.

Automatic Processing and the Path to Awareness

Roughly 95 percent of all mental behavior is automatized (Bargh & Chartrand 1999). Put another way, only about 5 percent of our mental world reaches the level of conscious attention. Those events that do become conscious reach that level only after

"clearing" many levels of pre- and non-conscious processing. In short, only the most salient and important information receives effortful attention. That information comes into conscious awareness through the lenses of non-conscious processing, and to this extent, even conscious information is enveloped in automatic biases. The journey of a stimulus to the world of conscious awareness is marked by discrimination, refinement, and bias.

When looked at physically, information received from the environment must be received by many different brain structures. The reticular formation, a structure located in the brain stem, the oldest area of the brain, determines arousal and attention. When the recticular formation becomes aware of incoming information, it sends an alert to the rest of the brain. The thalamus, also a part of the reptilian brain, is responsible for the first filtering process. All information received by the brain is first "checked" at this point. The vast majority of stimuli are prevented from proceeding past the thalamus; only that information that is remarkably salient or relevant is sent to further brain structures. From the thalamus, information is processed according to initial impulses. Information that is interpreted as threatening, for example, may be sent to the amygdala and then to the hypothalamus. Only after the approval of four or five brain structures does information become available to the prefrontal cortex, a structure located in the neocortex, the evolutionary advance of the human brain. In other words, consciousness is only given access to information that has been processed hurriedly through "ancestral" brain structures (Jacobs 2003).

The assembly line that brings information into conscious awareness is fully automatized. Effectively, this means that the conscious processing deck is somewhat stacked from the very start. Only information that is relevant to current functioning or

salient is allocated further attention. Later, we will discuss the criteria by which

information is deemed relevant or salient. For our purposes now, however, it is

important to recognize the chief implication of this process: all information that reaches

the level of controlled processing has been fully subjected to automatized processes,

which discard the majority of information and frame the remaining information in

categorical terms. Because the automatized processes that govern information

processing are meant to be efficient and quick, their accuracy very often lacks. A dog

barking, for example, might be processed rapidly to the amygdala, which sends off a

fear reaction, so that a friendly hello can translate consciously into a fearful warning. In

this sense, all information that reaches conscious awareness has pre-determined

meaning. In fact, information that reaches conscious awareness only does so because of

its initial assigned meaning.

The Need for Speed: Automaticity as both crucial and unavoidable

    From the standpoint of cognitive efficiency, none of these results should be

surprising. The complexity of human functioning, the intricacies of our environment,

indeed, the incredible array of stimuli to which at any moment we might give attention,

and the inefficiency of conscious processing mandates that the vast majority of mental

activity happen without our attention. I picture myself on a sunny afternoon outside of

my apartment playing Frisbee with a friend. While I am physically tossing the disc, I am

also talking with my companion, noticing the warmth of the sun, recognizing the

sharpness of a rock by my left foot, feeling happy, monitoring the area for walkers

whom I do not want to hit with the Frisbee, laughing, and changing positions in order to

catch my friend's toss. There are countless other processes taking place as well to keep

me oriented to my environment. Without efficient automatic processes underpinning all of these behaviors, action would be impossible. To put this in perspective, consider that research in cognitive science suggests that people can consciously remember chunks of only 5-7 pieces of information at one time with any sort of accuracy (Miller 1956). Consciousness, it seems, is only equipped for a very specific and limited kind of processing.

As I have mentioned above, the result of this inefficiency is that behaviors and thought processing tend toward automaticity with repeated occurrence. It is important to note that "inefficient" does not mean "inept," "unnecessary," or even "wasteful." Rather, in terms of an organism of complex functioning, "inefficient" mental processes refer to processes that are "not primary" or "the exception." Baumeister and Sommer (1997) categorize consciousness or controlled processing as an override mechanism that "intervene[s] occasionally to take behavior out of …rutted pathways" (p. 77). Bargh (1997) comments that continued findings in automaticity are "inevitable," and argues against the *prima facie* presupposition that both psychologists and philosophers tend to accept: that consciousness is, in some sense, where the real action occurs. Indeed, it is true that the histories of both psychology and philosophy, as well as folk assumptions about human behavior, point to controlled processing as the root and trigger of human behavior. That is, from the perspective of both of these disciplines, the burden of proof has historically been placed upon determinists (or automaticity researchers) to show exceptions to the rule of controlled processing (or conscious will). However, from the standpoint of empirical research and, to some degree, informed common sense and reflection (see the Frisbee example above), controlled processing is, and clearly must be, the exception to the rule of automatic mental activity. Despite the fact that our *awareness*

of mental activity circularly (i.e., fallaciously) leads us to the conclusion that conscious

processes dominate behavior, we must abandon this notion, and shift our paradigm if

we are to gain a more thorough understanding of human behavior and consciousness

itself.

As my thesis advisor has reminded me on a number of occasions, psychologists

have been not very good at defining and understanding consciousness. Perhaps one of

the reasons for this is that consciousness, and its role in human processing, has been

overestimated or misunderstood. For purposes of this investigation, my assumption,

based on empirical support to be presented, is that automatic processes govern behavior

under almost all circumstances. The text that follows will examine so-called

"controlled" mental activities and the interplay between automatic and controlled

mental processes.

"Controlled" Processing and Automatic Encroachment

We have already seen that the information made available to consciousness is

preselected by non-conscious, automatic processes. It comes pre-sorted, pre-

categorized, and pre-judged, having filtered through the non-conscious structures and

constructs that prevent our conscious mind from being inundated. The classical view of

information processing (that held by many philosophers, for example) is that once

conscious light is shed upon these data, they can be addressed objectively or subjectively

by the willful agent:

> This, at least, I think evident, -- That we find in ourselves a power to begin or
> forbear, continue or end several actions of our minds, and motions of our bodies,
> barely by a thought or preference of the mind ordering, or as it were
> commanding, the doing or not doing such or such a particular action. This
> power which the mind has thus to order the consideration of any idea, or the

forbearing to consider it; or to prefer the motion of any part of the body to its rest, and *vice versa*, in any particular instance, is that which we call the *Will*. (Locke 1690/1974, p. 42-43)

But, investigation of higher-order mental functioning reveals that it, too, is driven by automatic processes. Thought and ideas are generated and extrapolated on by priming processes and spreading activation. Recent research in thought control (Wegner 1994), for example, reveals that the attempt to suppress thoughts in many cases actually leads to cognitive hypersensitivity (i.e., increased occurrence) to the suppressed thought. Judgment is impinged upon by a number of cognitive biases that skew anything from our estimations of probability to our stereotyping and of other people; even our perceptions of cause are regularly misattributed because of automatized heuristics. Recently, evidence has emerged supporting the premise that goal setting, too, is a function of automatic processes (Bargh 1997; Bargh & Ferguson 2000; Wilson 2002). In short, the extent to which even conscious processes show evidence of automation suggests that the term "controlled" processing may be misleading. As we will see, our tendency to misunderstand and misattribute even our most seemingly controlled processes casts a great deal of doubt on the libertarian hypothesis of the freely-willing agent.

Schematic Organization, Priming, and Spreading Activation

Cognitive psychologists consider the mind's organization to be a web of concepts, or schemata (schemas). Although there is disagreement about the actual schematic mental representation of concepts (see Smith & Medin 1981; Medin 1989; Smith 1990; Gelman & Markmam 1986) and the nature of the connections between schemas (Anderson 1983; Collins & Loftus 1975), many psychologists agree that (1) our

ideas, concepts, and mental representations are stored cognitively as schemas, and (2) schemas are interconnected in a complex and vast web, but connected at a basic level nonetheless (Rosch, et al. 1976).

Due to their interconnectedness, schemas do not act independently. When one schema is activated, other, related schemas become accessible, as well. This triggering process is known as spreading activation (Kunda 1999). Some strong connections between schemas are formed empirically—if I say "salt" and you think "pepper," for example, this association has probably been made by your experience that salt and pepper tend to co-occur. Other connections may be psychological in nature. If I say "tree" and you say "childhood," this may be because you remember climbing trees when you were a child and trees are part of your childhood schema. In either case, spreading activation is an automatic process. Typically spreading activation occurs and directs our thoughts without our awareness of its sway.

As I stated above, interconnections between schemas are complex and interwoven such that activation of one schema may lead to activation of any number of other schemas (for example, "tree" might activate "childhood," but it might also activate "bird," "leaf," "grass," etc.). Spreading activation is influenced by interplay between internal and external factors (Kunda 1999). The connections we make between ideas are driven by our personality, our philosophies and worldview, our current mood, our likes and dislikes, and our recent thoughts and feelings. Likewise, spreading activation is motivated by environmental stimuli and our subjective reaction to them. The process of "priming" occurs when a particular schema or concept is activated by environmental cues (Kunda 1999). Our perceptions and judgments about the world shape, and are shaped by, the processes of spreading activation and priming.

Priming can occur with or without conscious awareness on the part of the individual being primed. The studies that follow outline both levels (non-conscious and conscious) of priming. What is most interesting for our purposes is that behavioral effects result from priming irrespective of a person's awareness of being primed: a target's behavior will be influenced by the triggering stimulus in significant ways that she does not realize.

Non-conscious priming is automatic and influences judgment and subsequent behavior without conscious awareness (Word, Zanna, & Cooper 1974). Because of this, psychologists can gauge people's attitudes indirectly by priming words related to a particular subject and measuring behavioral effects and spreading activation. Schemas for African Americans, for example, have been shown to be linked to words like lazy, aggressive, welfare, and ghetto (Bargh & Pietromonaco 1982; Devine 1989) based upon the social stereotype of this target group. Several studies have assessed racial attitudes based upon the observed effects of priming.

Bargh and colleagues (1996) used non-conscious priming to measure the effects of stereotypes on behavior. Researchers asked participants to complete the tedious task of counting circles on a computer screen. Before each new puzzle, a face was flashed on the screen for 26 milliseconds and then covered by a masking picture. Because of the very short duration of time, these pictures were perceived below the level of conscious awareness (participants had no knowledge of their having seen them). Half of the participants were presented subliminally with pictures of young African American men; the others were presented with pictures of young Caucasians. After several trials, participants received an error message indicating that it would be necessary for them to begin the procedure again. Reactions to this message varied systematically between the

two groups: those who had been primed with the African American faces responded with greater hostility than did those who had received the other subliminal prime. In other words, priming effects for the negative stereotype of African Americans actually increased hostile behavior, regardless of participants' lack of knowledge about the priming process.

Another study demonstrated the lasting power of conscious priming. Participants were asked to complete a word-scrambling test. For half of the participants, the list included words related to the elderly schema (Florida, gray, wrinkles), for the others, the test contained neutral words. After completing the task, participants were thanked and told that they may leave. Bargh and colleagues (1996) then covertly timed participants walking down the hall as they left. Those participants whose elderly schema had been activated during the experiment showed lasting effects after the experiment—they walked more slowly down the hall than did participants in the control group.

The results of this latter study are important to our understanding of conscious, controlled processing. Although participants may have been aware of the theme underlying their set of words, they most certainly were not making a conscious effort to walk more slowly (i.e., to accord with the stereotype) after the experiment was over. Neither were participants in the first study attributing their hostile reaction to stereotypes involving aggressiveness. The executive status of controlled processing appears more dubious when we consider that without any knowledge whatsoever of the causal mechanism of both of these behaviors, participants exhibited very real, observable changes in their interactions with the environment.

<u>Individual Differences in Spreading Activation and Priming: "Higher order" processing</u>

Some schemas are chronically accessible, such that they have a constant influence on spreading activation (Markus 1977). On a cognitive level, individual differences in ideology and worldview become chronically accessible schemas through which information is processed by default (Lambert & Raichle 2000). Narvaez and colleagues (1999) have shown that about 80 percent of an individual's moral thinking can be explained and predicted given knowledge of the person's religious identity, political beliefs, and moral judgment development. In other words, processing associated with moral issues is less a matter of conscious deliberation and effortful challenging of hypotheses, and more a matter of matching the situation in need of moral consideration with our schema of morally-correct behavior. This schema turns out to be a composite of the three dimensions above read through current cognitive priming, including current affect and arousal and pre-conscious judgment. We will return to the subject of automatic higher order processing in the discussion of attribution below.

<u>I Think, Therefore I Automatize?</u>

It is becoming more and more apparent that cognitive processing, from the most fundamental networks formed between related ideas and representations to the "higher order" processes of social and moral judgment, is dependent upon automatized, non-conscious processing. The fact that higher order processing, which is typically considered the domain of the consciously-willing and controlling agent, is so entrenched in automization suggests that the division between these two forms of processing is not entirely distinct, and may, in fact, one day be proven arbitrarily. Bargh and Ferguson (2000) comment that "the real difference between automatic and controlled processes is

not that one form is caused and the other not caused but that psychologists have at present, through research, discovered the mechanisms for the one form and not yet for the other" (p. 938).

In addition to cognitive automaticity, psychologists also recognize that emotions and moods, both of which are caused by individual responses to internal and external circumstances, deeply affect our cognitive judgments and processing in ways that we may or may not realize (Reeve 2001). A brief review of some of the research in this area follows.

The Breakdown of Emotional and Rational Barriers

Current mood and affect play an important role in a person's perception and judgment. Damasio's (1994) now famous book *Descartes' Error* addresses the blunder in assuming that emotions and reasoned judgment are separate processes, and lays the groundwork for understanding how it is that emotions underlie many judgments that we might otherwise label "controlled." Damasio's work with brain-damaged patients has shown that individuals with damage to the prefrontal cortex, the emotion center of the brain, show egregious errors in daily judgment. Even though these patients retain their abilities to perform on tests of "cold" rationality (puzzles, intelligence tests, etc.) and retain a theoretical understanding of social behavior, their judgment in daily life lacks greatly. Damasio theorizes that emotions provide "somatic markers" — physiological body responses to our situation — that allow us to reason with regard for salient situational factors. While these factors may not come into play in tests of pure logic (a puzzle, say), they are all too necessary for social reasoning. Slight changes in physiological arousal that indicate anxiety, excitement, or relaxation set the stage, as it

were, for relevant cognitive events to take place. Patients without the use of their prefrontal cortex and the emotional reactions it provides, lack this emotional priming for thought. The result is that these individuals show maladaptive, "irrational" decision making in everyday life.

Research has shown that depressed and non-depressed individuals differ in their abilities to process self-relevant information. In one study, depressed and non-depressed people were given either success or failure feedback and then were asked to perform a recall test (Ingram, et al. 1983). Researchers found that while non-depressed individuals could recall favorable self-references when presented with favorable feedback, depressed individuals receiving favorable feedback could not recall corresponding self-references. In other words, depressed mood "shuts off" a person's ability to think about non-depressive thoughts. Clinicians accept this cognitive reinforcement model: depressed mood triggers depressed/depressing thoughts, which in turn reinforce depressed mood (Scher, et al. 2004; Beck 1967, 1976).

Expanding these results more broadly, then, any mood or array of moods that dominate the emotional life of an individual will serve as a filter through which the world is viewed and information is processed. Current mood has been shown to affect judgments as they pertain to stereotypes and members of target groups, blame for and severity of conflict, intrapersonal interpretation of somatic symptoms, competence, life satisfaction, likelihood of threat, evaluation of public figures, persuasiveness of messages, and a number of other areas (Kunda 1999; Forgas 1995). Additionally, numerous studies have demonstrated the effect of mood and affect on memory. Emotions influence our storage, interpretation, and recall of past events (Blaney 1986; Ornstein 1991).

What is "Controlled" Processing?

In sum, the controlled processing of which we are aware is tied in significant, influential, and inextricable ways to cues and triggers over which we have no control and about which we are typically ignorant: priming, spreading activation, and emotional arousal. That is, these processes automatically create associations and preconditions for mental events, which are then played out on the stage of consciousness. The behaviors and thoughts that make up who we are have been shown by the foregoing research and related studies to be deeply automatized. Not only are we only aware of about 5 percent of our mental processing, but even *that* small slice of mental activity is caused and determined by other processes for which we can provide no conscious account.

Attribution Theory: We see what we want to see

It is commonly accepted among social psychologists that people are relatively inaccurate judges of causes, even though we think otherwise. The attributions we assign to events, including those events that very immediately concern ourselves, are wrapped up in a number of judgment biases and assumptions that lead to incorrect (though plausible) explanations for behavior and events. In order to understand how human judgment about causes can and does go wrong, I will present first a general understanding of the criteria that must be met so as to make an attribution.

David Hume, in both his *Treatise of Human Nature* and *Enquiry Concerning Human Understanding*, presents a controversial account of causality, one that continues to fuel debates about this subject even today. Hume's analysis of causation (or attribution) can

be viewed in two stages: the first, a skeptical inquiry into the limits of our experiential understanding of causation, challenges many of the common sense assumptions held by academics and laypeople alike. The second stage addresses the skeptical conclusions of its predecessor and offers a positive account of causation, one based in probability and constant conjunction between causes and effects.

Hume's is a psychological theory of causation. That is, his skepticism about causality is not based in ontological causation—how cause and effect relationships actually work—rather, it is based phenomenologically—how it is that human beings perceive cause. Given this, Hume's theory may be the first comprehensive theory of attribution we have. We see cause and effect, Hume asserts, when two events are related to each other by priority (event A happens before event B), contiguity (event A happens in a similar time and place to event B), and constant conjunction (event A and event B co-occur repeatedly) (Hume 1738/2001). Hume's classic example is billiard balls. When ball 1 rolls across the table and strikes ball 2, we say that ball 1 caused ball 2 to move because (a) ball 1 was in motion before ball 2, (b) ball 1 shared the same physical space with ball 2 upon impact, and (c) given many repeated trials, ball 2 has consistently moved upon impact with ball 1. In other words, through the senses we perceive the motion of the balls, and through reason we assign a causal relationship that tells us something about how these otherwise disconnected events relate to one another.

However, "assigning a causal relationship" is a mental event, and does not necessarily reflect any inherent causation. This is where the distinction between ontology and psychology becomes important. What Hume provides in the above account is a psychological theory: he explains the conditions that must be met in order for an observer to recognize cause-effect relationships. This is very different from an

ontological theory, which would concern itself with the actual relationship between ball 1 and ball 2, and the power or causal efficacy between one event and the other.

Another way of thinking about this is to consider Hume's distinction between necessary connections and constant conjunctions. For Hume, a necessary connection between events is not something that can be accounted for by actual empirical experience. We see ball 1 move towards ball 2, we see ball 1 and ball 2 related contiguously, and we see ball 2 move. We do not see, however, the power of the actual causal relationship that causes a change in ball 2. That is, we do not see a relationship of necessity between cause and effect. Rather, we see that each time ball 1 strikes ball 2, ball 2 moves, and we reason from this experience to causal understanding. A necessary connection, then, is an ontological phenomenon—it is a connection inherent in the relationship between ball 1 and 2. A constant conjunction, though, is a psychological phenomenon involving both memory and reason. In order to recognize constant conjunction, an observer must remember past examples of the interaction between ball 1 and ball 2 and infer from these past examples some general relational pattern; in this case, causality.

The Strength of Hume's Account for Understanding Attribution

If Hume's account is correct, we should expect that the psychological processes that recognize cause-effect relationships might easily render false conclusions some of the time. That is, we should expect that people see cause-effect relationships where there are none or conflate causes and effects. In fact, psychologists have long held that people are generally hasty judges of causes even though we tend to think of ourselves as

accurate perceivers. In reality, Hume's psychological account of causal associations does a fine job of explaining situations where causal relationships are misattributed.

One of the most infamous examples of misattribution of causes is the case of facilitated communication. In the 1970s, Australian scientists developed a revolutionary method for allowing communication in patients with disabilities who formerly had no means of communicating. With the aide of a facilitator, patients with severe developmental disabilities were communicating with others through a modified keyboard. Facilitators provided support for a patient's arm while the patient typed messages to the world.   As a result of this magnificent achievement, parents of children with disabilities were able to speak to their children for the first time. Patients wrote "I love you" to dear ones with whom they had never been able to converse. For many, this was a miraculous achievement (Shane 1993).

In addition to this, many patients with access to this new method of communication reported instances of abuse. Hundreds of people were implicated by patients—social workers, parents, friends, and others. Some went to jail. It was not until the early 1990s that psychologists Donnellan and Silliman called for a closer examination of this technique. The results were devastating. Under empirical testing, it was found that communication attributed to patients was actually being done by facilitators, non-consciously. As many of the facilitators were parents of patients, their deep desire to make contact with children actually drove their behavior, such that patients' arms were moved toward appropriate keys on the keyboard. A series of tests confirmed that despite the assurance from facilitators and scientists alike that the communication observed was coming from patients, it was actually the result of non-conscious intentions of facilitators. In this case, the belief in the constant conjunction

between a communicator's thoughts and subsequent expression led facilitators to assume that the cause of the communicated message was the thoughts of the patient.

It is also the case that our perception of causes changes from experience to experience, based on a number of factors, both internal and external. Research in this area suggests that there are a number of cognitive biases that skew our perception of cause-effect relationships. The "fundamental attribution error," for example, suggests that people tend to over-emphasize dispositional (internal) causes for observed behaviors, ignoring important situational causes (Jones & Harris 1967). Relating this to the larger question, then, this appeal to independent agency means that observers will often fail to take into account important, and perhaps more plausible, causes of behavior when assessing a sequence of events or behaviors. We bias ourselves toward causal explanations that favor free agency, and insofar as the question of free will is concerned, circularly conclude free agency to be true.

Research on patients with brain damage also suggests a gulf between the causal attributions of behavior (apparent explanation) and actual causes of behavior (actual explanation). Gazzaniga and LeDoux (1978) reported the case of P.S., a young man with damage to the corpus collosum, the "bridge" between the left and right hemispheres of the brain, which prevented his brain from sending messages between hemispheres. In the study, researchers displayed pictures in such a way that only one hemisphere received sensory information, and then asked P.S. to choose with either the left or right hand a card that related directly to the picture he had just been shown (in the study, only one of the choice cards was obviously appropriate). When the picture (say, a snow scene) was sensed by the right hemisphere and P.S. was told to select a card (say, a shovel) with his left hand (or when the picture was shown to the left hemisphere and

P.S. could use his right hand), he performed well at the combined tasks, due to the fact that the sensory task and the choice task both resided in the same hemisphere. However, when P.S. was asked to explain why he had chosen the pictures, he offered logical explanations for the processing in the left hemisphere, the language center of the brain, but offered a fabricated causal explanation for processing in the right hemisphere. That is, processing that happened in the left hemisphere allowed P.S. to create a cogent causal explanation for his choice because the verbal center of the brain also had knowledge of the triggering stimulus. When asked to explain the choices in the right hemisphere, though, P.S. could not offer a sensible explanation even though he had chosen the acceptable card, since the "explaining part" of the brain (the left side) had no access to the sensory information received by the right side. As such, researchers found that P.S. fabricated a likely causal explanation for his left brain choices, even though the explanation given had nothing to do with the actual triggering mechanism (the input to the right side of the brain). In addition, P.S. felt as confident with his accurate (left brain) attributions as he did offering his inaccurate (right brain) attributions. In short, even though only one set of causal explanations actually resembled the real causal process, P.S. had no way to discern between those attributions that were true and those that were false.

In another classic study of misattribution, Zillmann and Bryant (1975) had college-age men watch pornography. One-third of the participants were required to exercise for 10 minutes, rest for 15 minutes, and then watch pornography. One-third were required to exercise for 10 minutes and then watch pornography directly thereafter. One-third were simply told to watch pornography. Participants in the first condition rated the pornography as more erotic than did participants in the other two

conditions. In their conclusion, the researchers argued that participants in the first condition had given a higher rating to the pornography because the physiological arousal from the exercise had waned from conscious attention, though its physiological effects were still manifest. When participants in this condition viewed the pornography, they misattributed their exercise-based arousal to the effect of the pornography. Participants in the second condition, by comparison, could recognize and distinguish between physiological and sexual arousal, and participants in the third condition served as a control. Again we see the use of Hume's constant conjunction:

1. I know from experience that there is a constant conjunction between pornography and arousal;

2. I am now watching pornography;

3. So my arousal must be caused by the pornography.

While we cannot discard this syllogism altogether (certainly pornography did have something to do with arousal), we recognize that participants in this condition were only aware of one cause of their arousal. The fallacy of the single cause, which bases itself in oversimplification of ontological relations among things, applies here as it does more broadly to Hume's psychological account of causation.

In another study of emotional arousal, Schachter and Singer (1962) observed participants' attributions when experiencing drug-induced physiological arousal. Experimenters varied the level of information given to participants in order to vary their expectations of the experience, and achieved telling results. What the researchers found is that when participants felt physiological arousal for which they saw no clear explanation, they would attribute the arousal to whatever cognitions were available to them. By manipulating environmental conditions, the researchers went a step further

and demonstrated that participants could be primed with a certain cognitive schemas to create distinct attributions. If a participant were in the room with a nosy and irritating confederate, for example, physiological arousal would most likely be attributed to irritation. That same arousal though, in a different environment would elicit other causal attributions. Schachter and Singer (1962) end with the observation that "precisely the same state of physiological arousal could be labeled 'joy' or 'fury' or 'jealousy' or any of a great diversity of emotional labels depending on the cognitive aspects of the situation" (p. 398).

There is also evidence that people create causal explanations for non-consciously activated behavior. That is, individuals acting involuntarily (due to reflexes, neuronal stimulation, or suggestion) may create causal stories for their behavior even though the behavior lacks conscious "causes." In studies of individuals who receive posthypnotic suggestions, Wegner (2002) reports that very early research (Moll 1889) found a tendency for these individuals to invent causal explanations for hypnotically-induced, involuntary behavior. Participants were asked to perform certain uncommon behaviors while hypnotized (individuals were asked to wrap a flower pot in a cloth, set the new package on a sofa, and bow three times, for example), and then were asked to explain their reasons for acting as they did. Some individuals invented highly-complex causal stories—"I saw the flower pot there [and] I thought it was rather cold…and I bowed because I was pleased with myself for having such a bright idea" (Moll 1889 in Wegner 2002, p. 150)—others denied that the behavior had occurred, while still others offered simpler, internal attributions for behavior ("I felt as if I must").

Nisbett and Wilson (1977), in an article that reviewed numerous psychological studies involving processes of introspection and higher order processing, concluded that

people tend to "tell more than they can know." Due to a host of research suggesting the individual's very limited access to her or his higher order functioning, Nisbett and Wilson argue that when asked to provide reasons for behavior, individuals will rely upon *a priori* causal explanations; that is, will call to mind likely causal stories and offer these as justification even if these causal schemas are, in fact, inaccurate. Causal theories employed to explain behavior may be set by cultural rules and policies, implicit theories of causation, or empirical observation. Although some situations may elicit more or less accurate attribution, the vast majority of our so-called "privileged access" is actually illusory and arbitrary. Theories by Efran, Lukens, and Lukens (1990), which will be reviewed later, support this constructionist model, as well.

What these studies and others tell us is that there is a noticeable gap between actual causality and psychological perception of causality. In other words, the ways we understand causality may or may not match up to real processes. We must bear in mind that the psychological literature investigating attribution has identified a great number of heuristics and biases involved in judgment of cause, such that our subjective perceptions of cause seem to be at least as dependent upon external factors as internal ones. Still other research suggests that the judgments we make are not only biased, but also actually fabricated to a greater or lesser degree. In short, the study of attribution teaches us that we have very limited knowledge about the actual processes of causation. Moreover, because we organize the world in order to suit our needs, we tend to see people as final causes for behavior. We filter the world through schemas that allow us to see what we are looking for, and as later theorists (Wegner & Wheatley 1999; Efran, Lukens, and Lukens 1990) contend, this may actually serve to make us "strangers to ourselves" (Wilson 2002).

Summation of Our Condition

The foregoing has been offered to suggest that romantic ideals surrounding controlled processing—including reasoning, judgment, and objectivity—ought to be reconsidered by a humble, empirically-backed reader. Indeed, it becomes clearer with each addition to the automaticity literature that (1) there is no support for the claim that controlled processing is autonomous processing, but rather that it is dependent upon automatic processes; (2) our "controlled" judgments are typically a response to automatic processes and perform a corrective purpose; and (3) our conscious experience is only somewhat aware of the entire functioning of the human mind. I have mentioned that the idea of autonomous controlled processing is backed circularly by experience: we experience conscious processing consciously and therefore conclude that it is in control. We neglect the importance of non-conscious, automatic processing because we do not have access to its functioning. While these conditions used to be sufficient for positing an executive function of consciousness, current research suggests a paradigm shift that emphasizes the importance and predominance of automaticity.

On balance, conscious controlled processing affords us certain advantages that automatic processing simply cannot provide. Because we have cognitive access to mental representations of past events (memories) and mental speculation about future events, indeed, even that we have concepts like "past" and "future," we are allowed a long-range view of functioning that allocates energy to planning and anticipation. We are also afforded the opportunity for self-awareness and self-reflection, which, though often taken for granted, is certainly a fantastic and unprecedented achievement. Of course all of these "advantages" may be shifted easily to maladaptive functioning: I may be able to plan and anticipate, but I am also able to worry and feel disappointment.

I can be self-aware and reflective, but I can also be depressed that I am not the person I would like to be and get caught in the existential angst of facing an absurd cosmos. The evolutionary blade of consciousness cuts both ways.

In sum, I have offered a number of reasons to doubt the causal efficacy of consciousness. Moreover, research in attribution has given us reason to doubt people's ability to judge cause altogether. Given relevant research, we cannot at all be sure that our conscious minds are in control. Most of our behavior happens below conscious awareness and the behavior about which we *are* consciously aware may or may not be explained by our conscious attributions. However, the feeling of conscious efficacy, or conscious will, is undeniable. While we may not have an actual understanding of causal relationships, we certainly feel that we do; and while we may misattribute non-consciously-caused behavior to conscious processes, we definitely feel as though we have consciously decided upon our current course of action. This distinction between the "force" of conscious will and the "feeling" of conscious will, and further, the support for the latter but not the former, is accounted for by Daniel Wegner and Thalia Wheatley's (1999) theory of apparent mental causation. I turn to this now.

Wegner and Wheatley's (1999) Theory of Apparent Mental Causation

In light of burgeoning research suggesting that the notion of free willing agents is more a useful heuristic than an empirical reality, Wegner and Wheatley offer a very cogent explanation of why we seem to be deceived by our own conscious processes. "People experience conscious will when they interpret their own thought as the cause of their action," they argue, not necessarily when their own thought *is* the cause of their

action (Wegner & Wheatley 1999, p. 64). David Hume's account of causality, which has been discussed above, is the model from which this explanation springs.

Recall that Hume's account leaves ample room for skepticism about whether our perceptions of causation actually correspond to real causation. In fact, Hume's explanation only described the conditions under which people see causes, not the actual ontological relation between cause and effect. Moreover, according to Hume, such insight into actual causal relationships is impossible to verify empirically: there is no evidence for a necessary connection between causes and their effects; rather, cause-effect relationships are constantly conjoined to a probabilistic degree only.   All that is needed to perceive a cause and effect relationship between events A and B is for event A to have happened before B, for events A and B to have occurred in a contiguous reference in space and time, and for events A and B to co-occur repeatedly.

Wegner and Wheatley agree with Hume and the skepticism his account promotes. The experience of will, they argue, "is not a direct readout of some psychological force that causes action from inside the head… rather, will is experienced as a result of an interpretation of the *apparent* link between the conscious thoughts that appear in association with the action and the nature of the observed action" (Wegner 2002, p. 65). The supposed causal link between a conscious thought and an action is perceived when a mental event and a behavior are related by priority, consistency, and exclusivity. As I sit at my computer with a cup of coffee at my right arm, I experience a thought, T, that tells me, "Have a sip of this delicious coffee;" I also experience a behavior, B: my right hand grabs the cup and puts it to my lips. Typically, we conclude from this relationship that T caused B. We do this because we observe a relationship between T and B that includes these three criteria. We know that T and B are temporally

consistent, such that T seemed to occur prior to B. Events T and B are also consistent

with one another in content, insofar as T, which indicates a desire for B, and B, which is

the fulfillment of the desire, seem to be related to each other in commonsensical ways.

Finally, there seems to be a relationship of exclusivity between T and B: T seems to be

the most exclusive possible cause of B.

However, as Wegner, Wheatley, and Hume point out, there is no empirical

evidence to justify this, and moreover, given the research that has been presented thus

far, it is unlikely that such a simple, consciously-controlled causal relationship exists.

Conscious thought does not exist in a vacuum and is not autonomous. As we have

already seen, conscious thoughts are susceptible to a wide range of influences: from

emotions to environment, from cognitive priming to non-conscious data selection, from

judgment biases to subtle physiological and non-conscious cues. In short,

consciousness, the snowball on the tip of the iceberg, knows very little about the origins

of its subjects, and is frequently overconfident in its assessments.

According to the theory of apparent mental causation, our mistaken notions

about conscious will are due to two factors: (1) our inability to recognize non-conscious

third factors acting as causes and (2) classical conditioning between perceived conscious

causes and "resultant" behaviors. In the case of (1), psychologists and other empirical

scientists are constantly concerned about the problem of the third cause. If we take the

example given above, where T is the thought that I would like a sip of coffee and B is the

behavior that makes me drink, we see these two events related by priority, consistency,

and exclusivity and we conclude that T caused B. However, the problem of the third

cause says that some other, unnoticed factor, A, may have caused both events, T and B.

So, say event A was the physiological signal, "I need some caffeine," "I want something

warm," or say it was the non-conscious cue, "Reset your thought processes for a moment by redirecting attention." In such case, A is the overarching third factor that brought about both events, T and B.
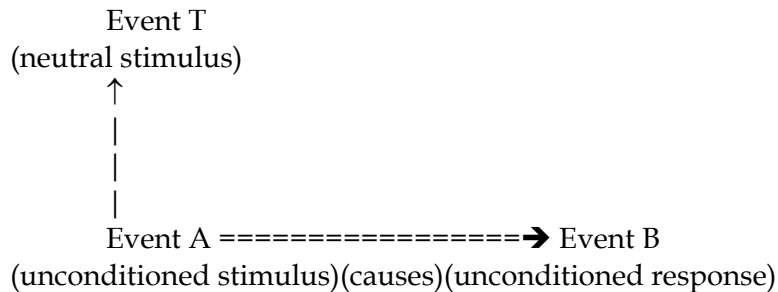
This relationship alone would not lead to the mistaken conclusion that T causes B. Indeed, as Hume tells us, two events can be observed to be simultaneous or nearly simultaneous without drawing causal inferences. The psychological perception of cause happens when the conjunction between the two events is constant; we become habituated to seeing two events repeatedly co-occur and conclude causality. Hume's habituation resembles the psychological concept of classical conditioning. In classical conditioning, a neutral stimulus is paired with an unconditioned stimulus repeatedly until the neutral stimulus becomes so closely tied to the unconditioned response that it, neutral though it may be, can elicit the unconditioned response. Once such a relationship has been formed, the neutral stimulus becomes known as the conditioned stimulus. Let us return to the coffee example in order to concretize this.

Event A is the physiological need ("I need caffeine) that brings about drinking behavior, B. The relationship between A and B is both sufficient and necessary for event B to occur. Furthermore, event A causes B even without conscious mediation. In this respect, the relationship between A and B is unconditioned: it will happen despite conscious thought. So:
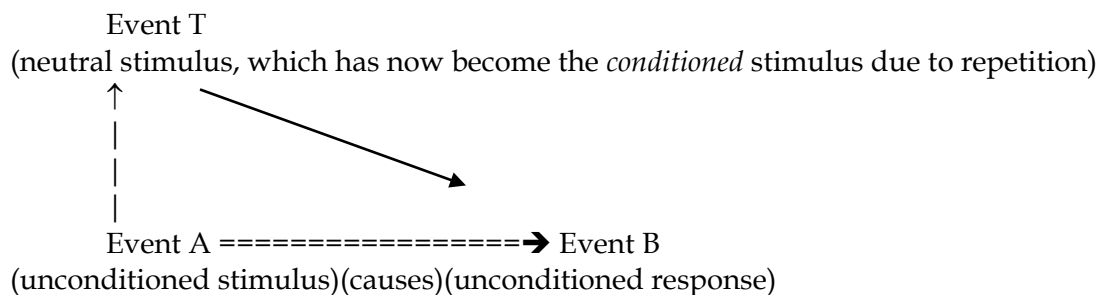
Event A ================➔ Event B
(unconditioned stimulus)(causes)(unconditioned response)

But now let us say that event T, the conscious thought, "Have a sip of this delicious coffee," is added to the relationship. This event would also be caused by A, the

physiological need, but, when taken alone, would be neither necessary nor sufficient for

event B to occur. At first, it is as if event T were floating above events A and B,

completely neutral. Therefore:

        Event T
(neutral stimulus)
        ↑
        |
        |
        |
        Event A ================➔ Event B
(unconditioned stimulus)(causes)(unconditioned response)

However, event T has a bias working in its favor: it is a *conscious* thought, whereas

event A is a non-conscious (i.e., unnoticeable) mental event. Furthermore, event B is a

consciously-observed behavior. And so, with repeated trials, and employing the use of

memory, we become habituated to fact that when event B occurs, so too does event T.

We are aware of the fact that these two are constantly conjoined. We are not aware,

though, of event A because we have no conscious access to this automatized process.

Therefore, with time, the neutral stimulus, T, becomes related by priority, consistency,

and exclusivity, as well as by constant conjunction with the unconditioned response, B.

Consciously, we form a connection between event T and B that leads us to believe that

the thought, "Have a sip of coffee" leads to the behavior of drinking. Thus:

        Event T
(neutral stimulus, which has now become the *conditioned* stimulus due to repetition)
        ↑   ＼
        | ＼
        | ＼
        | ➔
        Event A ================➔ Event B
(unconditioned stimulus)(causes)(unconditioned response)

In this model, event T becomes associated with event B, but is neither sufficient nor necessary for event B to occur. However, consciously, we have learned an association between events T and B, such that when event T is observed, we expect event B to follow.

A Complementary Perspective: Constructing a case for agency

Efran, Lukens, and Lukens (1990) offer a perspective on self-awareness that supports the skeptical hypothesis of consciousness lain out by Wegner and Wheatley. Their model uses a structural determinism conceptualization for understanding human behavior. According to Maturana (1970), who offered the first structural determinism view of humans, we are best understood organismically, as beings who are inseparably related to their environment in complex and important ways. The theory of structural determinism comes from this holistic understanding, stressing several main principles. Efran and colleagues see human beings as self-created (or self-organizing) beings who are constantly and subjectively wrapped up in their environment. The human condition is one in which we assign purpose to an otherwise meaningless set of life conditions, and our habituation towards one kind of framing determines worldview and information processing. Structural determinism furthermore asserts that rational thought is founded upon arrational premises and that language, as a communal entity, defines the ways in which we understand the world.

Efran, Lukens, and Lukens note that human beings are essentially meaning-makers and storytellers. We observe an event and then come up with a likely narrative to explain it. And, more often than not, we take that explanation as true: "A description is turned into a purpose that is then asked to account for the description" (p. 99).

Echoing the voice of attribution theory, these researchers suggest that people assign purposeful and intentional attributions to even the most ambiguous events. People who are shown a video of moving shapes, for example, explain the movement of one shape toward the other as intentional, and assign teleological explanations (e.g. the triangle *wants* to bump the square) to make sense of the event (Bruner 1986).

Philosophically, Efran and colleagues differ from Wegner and Wheatley in terms of their ontology. Whereas the latter theorists posit automatized non-conscious causes for behavior (in a more or less linear fashion), Efran, Lukens, and Lukens argue, in accord with their structural determinism model, that "recursively, everything feeds into everything else" (p. 29) such that it is impossible to determine actual causes of behavior. They see human beings as complex, dynamic, and self-organized entities who, at any given moment, are responding to any number of environmental and internal cues. Both of these theories, though, agree to two conclusions: (1) we have little or no actual knowledge about what causes our mental and physical behavior, even though (2) we really and truly believe that the conscious explanations we offer are accurate.

Both sets of theorists, moreover, recognize that our conscious attributions are, in many cases, arbitrarily drawn. In the case of Wegner and Wheatley, neutral stimuli are converted into conditioned stimuli and become conscious representations of non-conscious mental events. Efran, Lukens, and Lukens point out that assigning causal attributions to things in the world is the way in which we divide our environment into chunks of meaning. As they put it, "It is difficult to keep in mind that the purposes we infer belong to us, not to the [world]" (p. 97). This radical thesis suggests not only that we are unaware of the causes of our behavior, but that, moreover, there are no independent, identifiable causes of behavior, rather that behavior (mental and physical)

is the expression of a dynamic set of circumstances played out in the self-organizing system.

The theory of apparent mental causation in concert with the structural determinism thesis concludes that conscious thoughts are incidental to "resultant" behaviors. What purpose do conscious thoughts serve, then? There are a number of perspectives on this. One thesis is that consciousness serves little or no purpose—that it is an epiphenomenonon, an evolutionary leftover. The epiphenominalist believes that consciousness is a stage upon which life is more or less played out, in which the "self" cannot interfere (Wilson 2002). Another interpretation is that consciousness serves as a mental preview of what is about to occur, thus it has some adaptive preparatory value (Wegner 2002; Jaynes 1976). An argument might also be made that consciousness is important for attentional reasons. That is, conscious attention (mental energy) is allocated to tasks that are most pressing in order for them to be done well.

Despite years of research and thousands of articles on the mind and conscious processes, the perceived role of consciousness is still quite unclear. Whether this is because consciousness acts as the illusive ghost in the machine or not is a topic for debate. Perhaps the difficulty of researching consciousness is in part due to our theoretical frameworks and the questions psychologists have been asking. The research I have presented in this paper has made the case that automatic processing is not only the dominant activity of the mind as a whole, but also the driving force behind controlled processing itself. In fact, upon close analysis, the distinction between the two types of processing, while conceptually helpful, is not at all clear in reality.

Relating this to the larger scope of my thesis, empirical investigation does not lend support to the idea of a conscious agent who is free to act and think as she desires.

Rather, the view of a human being espoused by this research is one of a complex organism whose behaviors, thoughts, and environments are engaged in constant interaction. Free will, the illusory notion that "we" are somehow independent of the causal stream, simply does not and cannot reside next to the extensive evidence that illustrates the probabilistic and deterministic behaviors of a person.

And yet, as Wegner (2002) and others have pointed out, we still *feel* as though we are able to act freely and many people, philosophers especially, deeply want conscious, executive control for individuals. In the next section, I will turn to the philosophical side. Why do some philosophers want free will? And, what, if anything, do we lose if the incoherency of free will leads us to reject this concept altogether?

SECTION 2: PHILOSOPHICAL ANALYSIS

Putting it All Together: Natural philosophy and capacity causation

Contemporary philosophers who concern themselves with the question of causation and determinism have, in the last fifty years or so, tried to accommodate the empirical claims of other disciplines into their treatment of the "cement of the universe" (Mackie 1974). The well-documented and oft-repeated finding of quantum mechanics has incorporated into our understanding of the universe an indeterminacy that violates the Newtonian vision of a linear cause-effect ontology (for a compelling, accessible review, see Johnson 1995). Much like Democritus of 2500 years ago, philosophers in the 20th century have cited recent empirical justification for navigating the course between the random universe (the "dice of the gods" realm of the ancient Greeks), and the universe of necessity and natural law (*logos*). The emergence of chaos theory in the physics of the last half-century, and its iterations in physical, life, and social sciences, attempts to bridge the gap between the microcosmic disorder of quantum randomness and the apparent orderliness of macrocosmic determinism (Gleick 1987).

Of course, skepticism about simple, linear causation has been well documented in the philosophical literature since the days of David Hume. Recall from section one that Hume raised the phenomenological question regarding the human ability to perceive causal relationships. Our perceptions of cause are habitual and probabilistic ideas, and are ultimately, according to him, unverifiable by the empirical microscope. Hume's skepticism about the necessary connection between cause and effect fuels theories of causation to this day. It opens up a Pandora's Box that cuts straight to our

basic understanding of the world around us, and warns us against simple and common-sense notions of causality.

The overarching effect of Humean skepticism and scientific paradigm shifts has set philosophers on a course to examine new conceptions of causality. Moreover, a larger question has emerged which prompts investigation into the limits of causal explanations. John Dupré (1993) offers a chronology (both historical and ideological) of the debate from the philosophy of science, which juxtaposes the metaphysical question of causation with empirical scientific methodology. In the following, I will review his treatment of this progression. In addition, Dupré offers his own framing of the problem of causation in a way that, when extended into a macroscopic picture, offers a strong philosophical account of human agency that complements the psychological model of human functioning presented in section one.

The doctrine of determinism is most famously offered as the Laplace hypothesis. According to this theory, all things in the universe follow rigid deterministic paths that might be represented mathematically. If scientists knew the exact conditions of the universe at any given time, they could plug these coordinates into a cosmological formula and predict all future states of the universe. Under this Laplacean hypothesis, then, the world is reducible to omnipresent and definitive laws of nature. To the extent that all things in the universe have the essential properties of these laws (and they must, in order to be a part of the universe), they are governed by the universal pattern of cause and effect.

At the heart of the Laplace hypothesis is the optimistic scientist with the desire to predict future states. Dupré points out that there are two implicit pre-conditions for such a claim. First, we must be able to pinpoint to a degree of exacting accuracy the

present state of our phenomenon of investigation; second, there must exist a magic equation that represents our universe. In addition, Laplacean determinism requires a mechanistic understanding of the universe, such that all its constituent parts may be reduced down to linear relationships of cause and effect. According to Dupré and other philosophers, there are many reasons to doubt this model of the universe. Quantum indeterminacy aside, the assumptions of a Laplacean mechanistic universe are not empirically justifiable. First, the reductionism required in order to meet criteria for Laplacean pre-conditions, that is, the isolation of the phenomenon of investigation, may not (and in the case of many sciences does not) accurately capture the behavior of the phenomenon in its actual environmental relationships. Emergent properties, so-called because of their "emergence" from the interaction between phenomena, cannot be explained by reductionistic study of isolated phenomena, for example. Second, Dupré points out that even machines do not fit the rigorous standard of mechanization that must be assumed in a Laplacian universe. At some point, parts in a machine break down, and when this happens the deterministic order is disturbed in unpredictable ways.

Because of the limitations of Laplacean determinism, many philosophers have shifted (rightly, in Dupré's estimation) to "probabilistic uniformitarianism." Probabilistic accounts of causality suggest that a factor can be causal when and only if it raises the probability of an effect occurring in every possible context. The advantage of a probabilistic account of causality is that it provides for and explains exceptions to a supposed rule or pattern. Anyone with basic understanding of statistics understands that probability claims refer to groups and not to individuals. If smoking is correlated with cancer such that there is a 75 percent co-occurrence of the two, we can only assume

that in a randomly selected group of 100 individuals who smoke, about 75 of them will develop cancer in most cases. Our probability estimates will tell us nothing about the eventual fate of one particular smoker. Probabilistic causality, then, never gives us certainties about causal statements, only estimates about likelihood (in statistical language, our estimates almost never reach 1, which is the requirement of the Laplacean model).

The probabilistic account allows room for interactions between variables, such that contextual factors may affect the probability of a given causal relationship. Moreover, and more problematically, the probabilistic model gives reason to doubt the possibility of causal completeness, insofar as an identified cause actually may be the repeated expression of another, third variable read through a repeated context. This possibility, labeled "Simpson's paradox", is explained in a way that is not dissimilar from Wegner and Wheatley's (1999) model for apparent causation presented above. Dupré uses the example of smoking, and asks us to imagine the following scenario.

Suppose that a gene that predisposes people to smoking is also linked to both an increased prevalence of heart attacks and an increased prevalence of exercise. Under empirical investigation, scientists find that the probability of heart attack given that the person smokes is greater than the potentially-offsetting effect of increased exercise. Scientists thereby conclude that smoking causes increased prevalence of heart attacks. However, imagine that the same scenario were occurring in a location where tobacco is inaccessible. Given a change in environmental context, the same gene that causes heart attacks in smokers, now, in the absence of the possibility of such expression, expresses itself as an increased tendency toward exercise, and thus, decreases a person's chance of heart attack. The same factor, therefore, can cause opposite effects depending on the

environment that shapes its expression. This presents a two pronged problem: first, probabilistic causal relationships may only be provisional insofar as apparent causes may be the repeated expression of some overarching third variable, and second, contextual universality nexuses are untenable. In an interactive causal theory like this one, the presumption of universal presence of contextual factors is unlikely. Dupré points out that the scientific project requires an unceasing search for additional causal and contextual variables.

So, context, it turns out, is quite important to a useful account of causation. The example given above illustrates the dual-but-opposed-capacity for the expression of factors acting as causes. In every causal relationship we identify, there is an interdependent relationship between the labeled cause and the surrounding environment. Accordingly, *a priori* causes (factors that are causal simply by way of their causal nature) are false. Instead, a factor is a cause by way of the relationship between it, its effect, and other factors in the contextual environment that serve the causal relationship both through increasing its likelihood, on the one hand, and not inhibiting the relationship on the other. Upon ultimate analysis, there is no clear, *a priori* distinction between the cause and the effect; rather a factor is a cause when, given a certain context, its capacity to be a cause is expressed. The same is true of an effect. This is Dupré's capacity model for causation.

Dupré illustrates this capacity model further with an example that serves our purposes by bridging the gap between this theory of causality and its usefulness for understanding the free will question. In a baseball game, many players come together in one location and engage in a process that generates one noticeable effect: a winning game. Each player in the game is a factor in the ultimate effect, and each factor has a

varying capacity for contributing to the causal process. Moreover, these factors interact in ways that are difficult to predict. A batter with, say, a batting average of .327 faces a pitcher with an ERA of 2.2 and the result of the interaction between these two player-factors (a hit, a walk, or a strike out) is part of the larger interaction that involves the other players on the field, the fans in the stands, the weather, the playing conditions, and everything else.

According to the Dupré capacity model, the causal relationships that are identified in order to explain some effect depend largely on the question that is asked. This is not a new concept to Western causation; indeed Aristotle recognized four unique ways in which a cause can be efficacious. As a rule, then, the more general the question, the weaker the causal explanation is likely to be. In our example above, an observer may ask "Why did the Mariners win 4-2 over the White Sox?" This question could be answered any number of ways. "They won because Ichiro Suzuki hit a clutch two-run homerun in the bottom of the eighth inning" might be one answer; of course, this same causal event might be expressed as, "They won because Freddy Garcia threw a fastball that was just what Suzuki wanted." Or, perhaps they won because in the top of the ninth inning, when the White Sox had their chance for a comeuppance, it started to rain, which threw off Chicago's batting. Still another explanation would be that the Mariners won because the rules of baseball state that after nine innings the team with the most number of runs is deemed the winner. We could continue on with these causal explanations, and for each one offered we could continue to ask "Why?" until we wish the game had never occurred. The point is that causal relationships cannot be isolated from the context in which they emerge and that the identification of an interaction as "causal" depends largely on the framing of the question.

In bringing all this to bear on the question of free will, Dupré says that human beings are "bundles of capacities" (1993, p. 217) who stand out from their environment as entities who create order in an otherwise disordered world. The question of whether human beings have freedom is not as interesting or useful to him as the question of how human beings interact with their surroundings, shaping and being shaped by the multitudes of factors that comprise experience as a whole.

Toward a Capacity Model of Human Behavior and Agency

The empirical picture of human functioning presented in section one of this paper relates dialectically to Dupré's capacity model in the same way that Dupré's own baseball analogy does: it provides empirical reason to suspect that the capacity approach really does describe processes in the world, namely, human functioning, and in turn, offers a theory of causation from which to generalize about human agency without relying on notions of free will. In the interest of this second advantage, I would like to expand upon Dupré's model as it might apply to human behavior and interaction.

Human beings clearly do have causal power—the typed characters on this paper represent my ability to effect change in the computer. Our causal power, though, does not come from some mysterious will, but from the complex and dynamic relationship that occurs between internal states and external states. In the same way that a fire cannot burn without wood, our agency is dependent upon the factors with which and among which we interact. Similarly, the "bundles of capacity," which are the stuff of our agency, are interdependently related to their objects and the context in which a causal relationship occurs. That is to say that the very fact that we can be capacious

suggests that we must also be receptive. Human agency, then, is a dynamic relationship of interdependence within the context of our life-space. We are efficacious agents only inasmuch as we are receptive products of causes and conditions ourselves.

One of the difficulties faced by research psychologists is that human beings exhibit both equifinality and multifinality tendencies. On the one hand (multifinality), two people can be exposed to the same stimulus and exhibit very different responses. Similarly, one person can be exposed to the same stimulus in two separate contexts or at two distinct times and respond differently in either case. On the other hand (equifinality), if we expose a person (or a sampling of individuals) to two completely different stimuli, we may see the same observable response. Why is it that on one day I can buy flowers for a loved one and be met by an enthusiastic and loving response, but on another day get little more than a half grin and a meager "thanks" for the same attempt? The intuitive answer suggests a capacity-like explanation: because the context in which the stimulus is introduced has changed from one experience to the next, thus altering the causal capacity of the act of giving flowers. This non-uniformity is one of the reasons that, historically, behaviorism was deemed incomplete as an explanation of human behavior: we do not demonstrate one-to-one linear matches between stimuli and responses.

This disconnect has led us to the realization that neither environment nor cognitive-emotional state is sufficient in and of itself for explaining human functioning. In social psychology, especially, the back-and-forth dance between the individual and the social and physical environment describes a rich and complex interaction that fits well with Dupré's capacity model. In dyad interactions, for example, the level of attention that is allocated to a shared task can actually bring about an emergent property

of synchrony, in which interactors' behaviors become entrained in the social interaction. Interactors may, for example, non-consciously mimic each other's gestures or adopt each other's speech patterns in keeping with the flow of a shared, rather than autonomous, behavioral stream (Bernieri & Rosenthal 1991).

One resulting experience is that of a complex fractured self that responds uniquely in given contexts. Social psychologists identify a number of different "selves" with which a person may identify, and posit that the favored identification at any given time is driven by the social context (Kunda 1999). Between situations, people show stunningly low trait consistency. How Steve presents himself at the office tends to correlate very little with who Steve seems to be at home or who he is when engaged in a competitive game of basketball. Within situations, though, traits seem to hold. That is, that watching Steve at work one day is a good predictor of his behavior at work on other days, even though it tells us little about Steve in other situations. Indeed, our self-concept, the idea of who we are, has been shown to shift not only in terms of trait association ("I am a hard worker on most days, but today I am lazy"), but also in terms of breadth (the things we are willing to accept under the label of "me") (Wegner & Vallacher 1980). Viewed from a capacity model, our capacities for the expression of certain traits are in constant interplay with our spatial and temporal context orientation.

In addition, this capacity theory of functioning is useful as a theoretical framework from which to view the varying schools of psychology. When a person acts, psychologically-minded observers may generate any number of descriptive causal explanations. We may refer to motivation theory and see action as need-based. Or we may appeal to personality theory, positing that the behavior is an expression of internal traits. We may look to social psychology and claim that social judgment factors were

involved in bringing about action. We could also use cognitive or neurological explanations. All of these are useful and indispensable, according to a capacity model of functioning. Indeed, such a model is useful for a multi-systems analysis. Capacity theory does not require an either/or explanation, but rather seeks to understand and explain the interaction between these systems.

As discoveries in psychology continue to uncover a more comprehensive and detailed picture of human behavior, a shift to a capacity-like approach will be necessary for conjoining the research findings and theories from the various areas of focus within psychology. In addition, these further developments will render the notion of free will even less tenable than it seems in light of the research offered in section one. If for no other reason, the scientific explanation of human behavior will usurp traditional philosophical and religious theories of will on the grounds of parsimony: psychological explanations of human functioning present a more satisfying and scientifically-valid theory based on fewer assumptions than do the arguments offered by libertarian philosophers. And then there is the matter of evidence. Whereas philosophers in defense of free will offer hypothetical scenarios intended to evoke logical conclusions about the state of will, psychologists and scientifically-informed philosophers offer empirical answers based upon observations of actual people in action, and draw causal conclusions and theories about human behavior.

Eventually, the question must at some point emerge for any honest inquirer, "What good is free will?" The preponderance of evidence coming from psychology leads us to the empirical conclusion that human beings are complex interactors with their environment. We are not "free" in any sense of the idea, but dynamic, receptive, and involved with our surroundings. Philosophers tend not to challenge these empirical

data, but rather to reconsider notions of freedom, defining it in such a way as to plausibly show its existence (Dennett 1984, e.g.). In certain philosophers' minds, there is something intrinsically valuable to the idea of free will that warrants vigilant defense.

Based on the empirical evidence that science provides and the further evidence that is bound to emerge, I have assented to Daniel Wegner's claim that conscious will is illusory. This does not mean, however, that the idea of free will is unimportant. It also does not immediately disqualify the concerns raised by defenders of the idea of free will. For in their objections lies the answer to the question of why the notion of free will has persisted throughout philosophical and folk consciousness. To this end, in what follows I will consider philosophical objections to attacks on free will and from these extract the answer to the question of why the idea of free will is important to philosophers. I present what I find to be the most compelling and perhaps fundamental responses to this question: that what belief in free will affords us is not nearly as important as what the rejection of free will denies us. I then turn to the question that logically emerges from the first: are we really denied the things libertarian philosophers claim we lack if we rebuff free will? I argue that we can still retain important metaphysical constructs (like responsibility) in the absence of free will, and that it is the role for philosophy to unpack the contents of the scientific backpack and to adjust our metaphysical constructs accordingly *or* to offer well-founded justification for not doing so.

Why (some) Philosophers Just Won't Let Go

Free will is an illusion. If 95 percent of all of our mental processing is non-conscious and if the remaining five percent is largely automatized and/or consciously misattributed to a priori causal schemas (likely explanations), then this is an informed

concession. Libertarian philosophers do not offer counterclaims to these empirical

findings, and yet they continue to argue for the coherency of the idea of free will as a

causally-efficacious construct.

In fact, these philosophers are acting like other humans in this regard. Most

people tend to be mixed in their response to the issue. We want credit when we

complete a task well (that is, we want recognition for our agency), and we also want to

be let off the hook when less favorable circumstances arise. Consider a person sitting in

traffic at 4:45 on her way to an engagement at 5 o'clock. If she is late to the meeting she

will arrive in frantic mood saying, "There is terrible traffic on the road today. Can you

believe how bad our roads have gotten?" If, though, she arrives just before the meeting,

she may rush in saying, "I drove as fast as I could to get here and beat the traffic." In

one outcome, she is the pawn of the traffic gods who have cursed her with tardiness; in

the other she is the master of her fate. In reality, the woman sitting in traffic has

absolutely no control over when she will arrive once she finds herself in the jam. This is

the philosophical expression of a well-supported psychological phenomenon: the self-

serving bias. We tend to attribute positive outcomes to our own performance or abilities

while attributing negative outcomes to the circumstances. We do this all the time (Miller

& Ross 1975). In the same manner, philosophers and laypeople alike have what appears

to be an inconsistent attitude concerning free will: we want predictability in human

behavior and the universe so that we can theorize philosophically, walk safely down the

street, but we also, in some sense, want freedom. We want to be more than the product

of circumstances (Searle 2004).

There is an important distinction that emerges. Philosophers arguing in favor of

free will might be arguing from one of two platforms, or both. Platform 1 consists of the

philosophical position that we *have* free will. That is, that free will actually exists in an

efficacious way. Platform 2, though, states that we *want* free will. That the idea of free

will affords us something that is desirable. Note that these two propositions are

independent of one another, and although they may be compatible, they do not

necessarily need one another. Also consider that whereas platform 1 can only be

satisfied if its claim is true, platform 2 is instrumental: free will allows us X. This raises

the question of whether or not X can only be satisfied by assuming that free will exists.

Is it feasible that X could be satisfied by some other circumstance, we may ask? I

consider both of these platforms separately.

Why Do (some) Philosophers Think We Have Free Will?

Possibility One: The power of an illusion

       With regard to platform 1, why is it that philosophers argue for the actual

existence of free will, despite evidence demonstrating its illusory nature? One reason

may be that it seems to be there; that the scientific understanding of human behavior

undercuts the very fundamental and raw experience of *willing*. As Thomas Nagel

comments: "Something peculiar happens when we view action from an objective or

external standpoint. Some of its most important features seem to vanish… Actions seem

no longer assignable to individual agents as sources, but become instead components of

the flux of events in the world of which the agent is a part" (Nagel 1986, p. 229). Even

the most skeptical of scientists (Wegner 2002) grants that the experience of will is a very

powerful and convincing illusion. To doubt the concept of will, libertarians might

argue, is to reject of one of our most personal and direct experiences.

Doubting one's own free agency presents, in a very real way, skepticism about one's own existence: indeed, if we cannot be assured of something as fundamental as the act of willing, how can we be assured of anything proved by experience, including science itself? Echoes of Descartes sitting by his fire might resound in the heads of philosophers attempting to make sense of the endless causal chain that determines action in the absence of the free agent. It is plausible, then, that some philosophers posit free will because, to use Cartesian language, it is perceived "clearly and distinctly"—it seems to be there. Their writings try to establish the legitimacy of this intuitive position (see, for example, Reid 1792/2001) by arguing that free will is not only an apparent phenomenon, but an actual one as well.

In fact, the distinction between appearance and reality has many iterations in the free will debate. Nagel (above) distinguishes between the internal, subjective perspective, which clearly feels like will, and the external, objective perspective that reveals agents and actions to be "components of the flux of events in the world" (Nagel 1986). Wegner (2002) categorizes the distinction as the experience of will on the one hand and the (non-existent) force of will on the other. The question emerges, then, why might we be fooled into believing we have free will?

This is a question that skeptics of free will must take seriously. Answers by Wegner and Efrans, et al., have been offered in section one, and I would like to add an additional possibility. Perhaps the idea of free will is actually a post-hoc proposition that emerges from an ability to look back and wonder (passively) if we could have done otherwise. Essentially, the idea of free will is dependent upon a fixed past, which resides in memory, and counterfactual reasoning. Because we are able to think of alternative possibilities for any event in the past ("if only I had turned left at that traffic

light instead of right"), we may be led to believe that there was more than one option at the time of decision-making, even if there in fact were no other available or likely options. That is, we make a faulty induction: we generalize our thought experiments about the past, which are passive insofar as no amount of thinking about the past can change it, to conclusions about (active) agency in the present. This process has its advantages: because we are able to remember and to think counterfactually, we may be able to learn from perceived mistakes in the past such that we are driven away from similar situations and responses in the future and driven toward other suitable options. However, this is not to say that this conscious process, this idea of free will, has actual efficacy in the present. The claim, ironically, is that those philosophers who argue that consciousness or self-reflection gives us free will are correct—just not in the way they would like to be. Consciousness, the repository of memory and imagination, gives us the *idea* of free will, but nothing more.

Possibility Two: The Power of Suggestion

Another, related, answer is that the social and cultural institutions in which we live and operate reinforce the idea of persons as free agents and in fact depend upon this idea. Judeo-Christian dogma, for example, which permeates Western (and some of non-Western) society, claims free will as the underlying mechanism of choice and behavior. Free will, or rather the choice we make on its behalf, is the basis of punishment and reward in the eyes of a god-figure. And in fact, this belief, that human beings are isolated, self-caused decision makers, is one that has been extended to philosophy generally (even those philosophies which deny the existence of a god) as well as to Western theories of politics and law.

The twentieth century saw the rise and tapering-off of existentialism, a philosophy that emphasizes the strength of the individual over her environment, life circumstances, and social existence. Indeed, it could be argued that the fundamental maxim of existentialism is to subject all events—mental and otherwise—to the microscope of conscious processing; to abandon automatic thoughts, actions, and systems in favor of individual meaning and efficacy. Jean-Paul Sartre (1957), one of the recognized founders of existential thought, reflects:

> *What is meant here by saying that existence precedes essence? It means that, first of all, man exists, turns up, appears on the scene, and only afterwards, defines himself… Thus, there is no human nature, since there is no God to conceive it. Not only is man what he conceives himself to be, but he is also only what he wills himself to be after this thrust toward existence… Man is nothing else but what he makes of himself.* (p. 345)

Within psychology, as well, an existential bias emerged during the mid-1900s, mostly due to the experiences and writings of Viktor Frankl. Logotherapy emerged on the wings of the psychological humanism, a movement that held unconditional positive regard and self-fulfillment as the cornerstones of the therapeutic approach. Frankl's existential therapy was a process of self-actualization. Echoing the words of Sartre, Frankl's approach sought to instill in the client feelings of personal control, willfulness, and self-identity. Controlled, effortful reflection would lead to insight, thus creating a self-actualized, full human being (see, for example, Frankl 1962).

But the appeal to will does not limit itself to academic theory. Law making and standards of guilt in the American judicial system echo this faith in free agents. An act is willful, according to the law, when it is intentional, conscious, and purposeful. The standard of guilt for defendants rests on the willfulness associated with the offense. The presumption of the law is that human beings operate within domain of controlled processing by exercising self-regulation, self-control, and conscious decision-making.

These examples coming from philosophy, psychology, and legal policy demonstrate the tremendous extent to which a belief in free will is not only favored, but assumed and required by institutions within this society. The presumption of the autonomous, free willing agent is ubiquitous, especially among individualist cultures like our own (Efran, et al. 1990). That people are the result of their own free choices is more or less a given understanding within our cultural context. So this is another possibility: we claim free will because we are told we have it and because our societal structure assumes its existence.

Why Do Humans Want Pits? Our desire for free agency

Are there other possible reasons that philosophers think we *have* free will? Arguably, no. Philosophers who argue for the existence of free will must either do so because they are certain of their experience of will and/or because they can observe vis-à-vis institutions in our society that free will is presumed. Notwithstanding, there are a number of other reasons to argue in favor of free will even though its existence may be dubious. To examine these, though, we shift to the second platform, those philosophers who *want* free will. As I stated above, there can be variation in how any given philosopher relates to these two platforms: many philosophers who argue that human beings *want* free will do so because they also believe that human beings *have* free will (Reid, Kane). Others, especially compatibilists, recognize the desire or even need for free will and therefore attempt to define it in such a way so as to show that we also, conveniently, have it (Dennett). A third group of philosophers recognize a certain desirability of free will, but are not at all certain that we have it (Nagel). This distinction

is a helpful analytical tool for a fourth possibility as well. Perhaps free will is an illusion *and* we do not want it.

In order to understand this position, we must first understand the instrumentality of free will. There are certain metaphysical concepts afforded to us by free will that we desire, or even need, in order to function as people in a society. When one reads the arguments of libertarian philosophers, the most prevalent motivation behind their arguments is that without free will there can be no responsibility, and thus no accountability and no morality. This is precisely the claim that will direct my continuing investigation. In what follows, I will examine some of the very real libertarian concerns were the executive agent of free will to fall. But then I will ask the question, is free will really the only and best basis for metaphysical necessities like morality? Does free will really do the job philosophers suppose it does? Or, can a new ethic emerge by the giving up the illusion in favor of a capacity approach to human agency?

The Threat to Personal Identity As We Know It

I made reference in section one to the analogy of the peach and the onion, and claimed that defenders of free will must have pits. In other words, for free will to exist in any meaningful way, there must be a free agent who is at once isolated from the overwhelming behavioral and environmental stream (and thus uncaused by these factors) but also able to interact with it in order to exert efficacy. The cognitive onion, which has no pit, reveals itself to be a collection of layers of thoughts and emotions activated by internal brain processes and external environmental cues, and if we accept this, we must recognize that there is no such free agent. Personal identity, the question

of "what is a person," lies closely in hiding when the question of free will emerges. Gary Watson (2003) sums up a very prominent concern for philosophers and others when faced with the possibility of determinism: "…if a deterministic explanatory scheme is valid, it looks as though I do not originate my actions; I am merely a confluence of causal processes that were initiated long before my existence… Autonomy then is an illusion" (p. 2).

There is an implicit judgment in Watson's summation when he says "I am *merely* a confluence of causal processes…" (emphasis added), that captures a major concern for Western libertarians. Somehow, in the mind of Watson and others, the interrelatedness that determinism (or something like it) requires is a matter of inadequacy. Upon reflection, we are right to wonder if this appeal to self-importance seems hasty—why are autonomous, self-caused, *free* agents preferable to other alternative conceptions of persons?

Indeed, the view of a person that has been offered by social psychologists and reinforced by a capacity model of human functioning is not skimpy by any means. We are beings in constant contact with our environment. Equipped with brains that are able to sift through millions of stimuli at any moment, we make sense of the world in a way that allows us to function well within our surroundings. We are sensitive pattern-seekers and meaning-makers. We are storytellers. We do philosophy. In being "confluences of causal processes," human beings are inextricably linked to objects, ideas, and circumstances in the "outside" world. Our lives are intimately involved with the past and present, including its events, ideas, and people. And despite the dubious nature of consciousness and its ability to deceive, we are reflective and aware. What is so meager about this?

It is indeed true that the concept of a person fundamentally changes if we accept the consequences of people as receptive bundles of capacities. For one thing, it is hard to separate definitively the boundaries between the person and the environment, insofar as the two are in constant dynamic interaction. Our individualist culture, which includes Judeo-Christian ideas of personhood (including the soul) and capitalist ideals of self-determination, strictly enforces boundaries of personhood such that there can always be a discernable and responsible agent for each attribution of sin, praise, success, or failure.

However, these assumptions of individuality and subsequent divisions are culturally-determined and not necessarily statements about reality. Examination of classical Eastern conceptions of personal identity, for example Buddhist hermeneutics, reveals a complete ontology of interrelatedness that in many ways predicted the conclusions of western science hundreds of years prior. Consider the following quote in Joanna Macy's Mutual Causality in Buddhism and General Systems Theory, which combines western systems theory analysis with Buddhist causation:

> *We must do away with the subject-object distinction in analyzing experience. This does not mean that we reject concepts of organism and environment, as handed down to us by natural science. It only means that we conceive of experience as linking organism and environment in a continuous chain of events from which we cannot, without arbitrariness, abstract an entity called 'organism' and another called 'environment'. The organism is continuous with its environment and its experience refers to a series of transactions constituting the organism-environment continuum.* (Laszlo 1969 in Macy 1991)

Strikingly, even within our own individualistic society, philosophers of science (like Laszlo) and scientists themselves are challenging Western conceptions of self in a way that echoes the old wisdom of Eastern philosophy. The capacity theory of human functioning offered in this paper suggests the same general conclusion: in the absence of free will, people are contextually interdependent with their environment. The above quote diverges somewhat from the capacity theory of human functioning in the sense

that it takes issue with the very concept of the individual, while the capacity theory views individuals as receptive bundles of capacity. The basic point, though, is that our western assumption of personhood is a culturally-situated and conventional means of understanding the world that does not hold for all contexts and cultures.

Collectivist cultures, for example, have been shown to have distinctive notions of person (foreign to many Westerners), which include lineage, family and friends, and social group membership as components of identity. Kunda (1999) offers a summation: "American examples point to a sense of the self as distinct from others, and highlight the value of being different and special. In contrast, the Japanese examples point to a sense of the self as closely interrelated with and dependent on others, and highlight the value of harmoniously fitting in…" (p. 516).

It is interesting to consider our reactions to the possibility of an interrelated description of ourselves. To the extent that we value the ideals of rugged individualism and self-determination, an attack on free will becomes all the more irksome. We seem to want pits, for one, because this is how we relate to others and ourselves. It is embedded into our cultural worldview that human beings are individuals who operate according to their own executive functioning. However, there is no definitive reason to believe this cultural myth, and in fact, there is mounting empirical evidence to show its falsity. In fact, people are embedded within the context of their life-space to an inseparable degree. Personal identity, when viewed from a capacity model must focus on relational interactions between the organism (person) and the environment and the contexts that bring about shifting boundaries and ideas of the self.

The Feeling of Control

The issue of free will can also be thought of as a question of control. For any action, we may inquire, who or what was the cause? When something is regarded as "out of our control" we commonly see it as an inevitable circumstance. Likewise, when something is deemed "within our control," the result of the action is typically assumed to be intentional and we are more likely to entertain the possibility of other outcomes. Suppose I am driving down the street when the driver ahead of me quickly slams on his brakes and I rear-end the back of his car. Say there are two possible reasons for this: the first is that the brakes on my car suddenly stopped functioning; the second is that my mind was distracted for just enough time so that I could not register the situation at hand and respond accordingly. Most people would agree that under the first circumstance, the accident was beyond my control. However, in the second circumstance, we are more likely to judge that I was in control and thus the accident *could have been* avoided. This is an odd, though typical, reaction. Firstly, in neither circumstance is the actual event changed—both lead to the same outcome, so our projections about what might have been are based only in imagination and counterfactual reasoning. Moreover, with a little reflection we may see some absurdity in our conclusion about the second circumstance: if, by any means, the displeasure and annoyance of an accident could have been avoided, surely I would have avoided it. Nevertheless, we tend to stick with our assessment of the second circumstance and I am blamed for the accident. "I didn't mean to" is given very little recourse.

Was I in control when my mind was elsewhere? What if my mind was elsewhere because I had just learned of my grandfather's death? What if I had been given some hallucinatory drug prior to my drive home that had just started to take effect? And

what about the other driver? Is it important to know his reasons for stomping on the

brakes? All that should be conveyed here is that we tend to make hasty, fuzzy, and

heuristical judgments about control that do not necessarily reflect the actual facts, do not

usually dwell on the actual stream of events, and may not show consistency from

situation to situation. These judgments are efficient: in knowing who or what was the

controlling force for an event we can quickly fit our explanations into schemas about the

way the world works. But these judgments are also adaptive. Numerous studies have

shown that perceived feelings of control, the extent to which a person feels that she or he

can change the present circumstance, contribute to differences in functioning

irrespective of whether or not a person has actual influence over the situation (Lefcourt

1973; Langer 1975). Individuals who feel as though an adverse situation is within their

control, if only partially, show longer persistence when trying to complete a task (Glass,

Singer, & Friedman 1969), and people who are able to predict certain circumstances, in

some sense to exert control vis-à-vis cognitive preparedness, also show benefit (Glass,

Reim, & Singer 1971). The feeling that one has mastery over one's environment is

associated with feelings of competency and encourages effort, persistence, and

engagement with activities and life in general (Reeve 2001). In sum, perceived feelings

of control are a necessary part of psychological adaptation and well-being.

　　　　To be clear, then, it is our perceptions of control, and not necessarily control

itself, that leads to positive outcomes. Whether or not the feelings are based upon an

illusion does not appear to matter. Lefcourt summarizes: "Illusions do have

consequences, and… the loss of the illusion of freedom [and control] may have

untoward consequences for the way [people] live" (Lefcourt 1973, p. 417). Indeed, our

relationship to this construct suggests a true reticence to give up perceptions of control under most circumstances.

To what extent are these perceptions accurate reflections of an actual ability to control? How might we talk about control from the perspective of capacities? Clearly our colloquial understanding of control is incoherent given the relational model used here. When most people talk about control, they most nearly mean the ability to manipulate or dominate a situation or thing. From the perspective of free will, manipulative control can be seen as the power of a person to exert her will upon something else. But from a capacities perspective, this kind of control, in which a person tries to overexert one set of capacities to the dominant exclusion of other contextual factors, is prone to create unexpected or undesired effects.

In fact, ironic or ineffectual consequences for forced manipulation have been demonstrated empirically. Daniel Wegner's (1994) research on thought suppression, for example, demonstrates that such an effort yields ironic reversals: in trying to suppress the thought of a white bear, participants actually made the thought hyper-accessible and intensified its presence. Current theory about ironic reversals in thought control suggest that the act of over-exertion involved in trying vigilantly to guard against a thought reduces cognitive resources, and therefore has the unintended effect of impeding our ability to bring up replacement thoughts.

Consider another example from clinical psychology. What might happen if a father drops a young daughter into a swimming pool without preparing her for the experience? In effect, the father has manipulated the situational context in a way that lessens the responsive capacity of the child. It would not be unlikely that the child, not knowing how to swim, would leave the experience with an aversion to swimming pools.

Despite the father's best effort to encourage his daughter to face her fears and learn to swim, his forceful manipulation actually causes a reverse effect.

Since causation and control are relational events having to do with the interaction between factors in a system, and because factors become causes only when the causal context, an amalgamation of the capacities of all involved factors, provides for the causal relationship, then it stands to reason that if manipulation of the causal context occurs, if one bundle of capacities is prematurely or forcefully activated, as in the above cases, something other than the intended effect will result. Because causes have no inherent or *a priori* ability to effect change, effects are more likely to occur when matched with the appropriate factors and causal context. Control, then, in the capacities model, has more to do with fitting certain capacious factors with appropriate contexts than with exertion or domination of one factor over all others. Just as trying to jam a square block into a circular hole will result in either failure or destruction of the hole, so too does manipulative control damage the system in unexpected ways.

From the perspective of libertarian philosophers, though, feelings of control and control itself may serve more than a practical, adaptive function, and are in fact a requisite for establishing responsibility. There are certain metaphysical reasons to want free will, and according to these philosophers, the constructs granted to us by our theories of internal and external control can only exist in a universe where people are free choosers of their behaviors and in control of their own destinies. In his case for free will, Robert Kane offers the following:

> *Why do we want free will? We want it because we want ultimate responsibility. And why do we want that? For the reason that children and adults take delight in their accomplishment from the earliest moments of their awakening as persons, whether these accomplishments are making a fist or walking upright or composing a symphony.* (Kane 1996, p. 100)

In addition to the feeling of autonomy granted to us by free will, which reinforces our sense of selfhood and makes us the masters of our fate, the doctrine of free will, because it posits persons as final causes of behavior, allows for accountability and responsibility, merit and punishment.

Freedom and Responsibility

Arguably, the concern for individual responsibility and morality is the most common motivation for defenders of free will. Some philosophers are explicit about this concern, while others do not directly state its importance. However, the issue permeates the literature on freedom (Watson 2003; Smart 1961; Van Inwagen 1983; Strawson 1963; Lewis 1981; Frankfurt 1969; Kane 1996; and the list could virtually extend several pages). In one form or another, defenders of free will wrestle with and object to the following argument from anti-libertarianism:

> (1) A person can only be held responsible for those actions that she herself originates freely.
> (2) If the behaviors, thoughts, and actions, the personality, literally the person herself is nothing but a confluence of causal processes (that is, if the doctrine of free will does not hold), then she is not the originator of her actions.
> (3) Therefore, a person cannot be held responsible for her behaviors, thoughts, or actions.

From here, there is a domino effect. If a person is not responsible for her behaviors, thoughts, or actions, morality does not make sense. Neither do reward and punishment.

All of a sudden, the emperor has no clothes, and we are plunged into meaninglessness.

Am I overstating the issue? Consider Paul Russell's synthesis of Dennett:

> *…if determinism is true and 'every deed and decision is the inexorable outcome…of the sum of physical forces acting at the moment,' then the human condition would be a 'terrible' and 'frightening' existence… Freedom would be an illusion, and we would be reduced to 'awful' circumstances similar to those of individuals who find themselves imprisoned or paralyzed, or subject to (hidden) control and manipulation by others.* (Russell 2002, p. 230)

Such a statement instills fear in the skeptic in much the same way that Pascal's wager encourages belief in God: by asking us to visualize the horrible alternatives to the defended thesis and to concede its value on this basis alone. If this sounds more like advertising than philosophy or science, we see the urgency of the issue at hand.

There is something circular about the entire position: in order for a person to be held responsible for his actions, he must have freely chosen them, because if he has not freely chosen them, he is not responsible. The two concepts, freedom and responsibility, are applied reflexively—where there is one there must be the other. Together, they form the bedrock of other metaphysical constructs, namely blame, merit, and morality. An appeal to common sense is probably the most succinct and strongest argument for the constant pairing of the two: how can a person be responsible for something she did not freely choose to do? However, common sense arguments sometimes hide complex issues and are not necessarily tenable explanations. After all, we all know that birds of a feather flock together, except of course when opposites attract.

In a paper that has become a landmark in the current debate, Harry Frankfurt (1969), a libertarian philosopher, challenges the necessary connection between freedom and responsibility and shows that the former can be irrelevant to the latter. Using a definition of freedom as the availability of alternative possibilities, Frankfurt explains that in cases where a person has only one option (and thus no possibility of "free" choice in any meaningful sense), responsibility may still be assigned. He goes a step further, arguing that even if external forces are present to ensure only one outcome, there can still be an identifiable responsible agent. Frankfurt is not concerned with whether or not an action was necessary, but rather with the actual process of willing itself. Insofar as a

person subjectively experiences willing a behavior, she is responsible for its consequences.

What is useful for our purposes is that the Frankfurt model allows us to think about alternative conceptions of responsibility. It is not entirely obvious that freedom and responsibility are inseparable if we take a closer look. If a soldier, under the threat punishment for disobeying orders, kills another person, is she responsible for her actions, for example? We tend to say yes, we do want her to be held responsible, regardless of whether the decision was hers to make. If I make the (free) decision to go walking in a high-crime area of the city and get attacked, am I responsible for the beating I receive? To say yes would seem callous and unreasonable. We see here that even common sense can lead us to very different ideas about freedom and responsibility.

The inevitable question, then, is does the doctrine of free will do all that philosophers hope it does? And, secondly, is free will really the only (and best) premise for establishing these things? The answer to both of these is no. Free will, when examined, is not only difficult to establish, mysterious (Van Inwagen 2002), based on assumptions and biased perceptions, inconsistent in application, and contrary to our understanding of the rest of the natural world, it also simplifies the complexity and interrelatedness of life to the choices of egocentric agents. It is a more or less unsatisfactory view, when looked at empirically or even with informed common sense, which can lead to "righteous indignation" (Smart 1961) and misattribution of agency. Why is the man on the corner homeless? Because he made bad decisions and has not willed himself a better life. Is this really the best we can do?

Responsibility and morality are important constructs that cannot be abandoned. Libertarians argue that without the notion free will we cannot establish these in any meaningful way. I disagree with this hasty assessment. Certainly responsibility and morality *can* be established under libertarian framework, but this is not the only method for positing these things. In fact, philosophy might best progress with the free will debate by trying to establish tenable theories of morality and responsibility under other frameworks. There are, in fact, moral systems that do not rely on the premise of free will, and, in fact, deny that humans are free (Goodman 2002). Buddhist philosophy, for example, recognizes a complex interrelatedness of all life, and puts forth the doctrine of dependent origination, which states that everything –from people to trees to thoughts and attitudes—arises and passes based on causes and conditions. They argue that persons have no essential self, but are instead dependently arisen and will dependently cease to exist. Even without the postulation of an autonomous self, Buddhist philosophy advocates for a strong set of ethics that include non-violence and, overall, compassion for others. Goodman (2002) suggests that Buddhist ethics are really guidelines for living a meaningful life without free will. Interestingly, when we consider the Western libertarian argument that free will is the basis of responsibility and moral behavior and compare it with morality established under a different framework, we see distinct outcomes. For example, no wars have ever been waged in the name of Buddhism (Blumenthal 2002). This is a claim that Western moral theory, which stresses autonomous and free agents, cannot match.

A Thought Experiment in Responsibility Without Free Will: Agency, and not freedom, as a basis for responsibility

My intention is not to institute a code of morality or generate a new definition of responsibility that suits a capacity model of human behavior. Instead, the focus here will be to illustrate with broad strokes some of the possibilities that emerge once we challenge the assumption of free will. First, though, we must synthesize the human condition as it is laid out in section one with the capacity model of humans as receptive bundles of capacities, so that we can view this new framework on its own merits, rather than as the "absence of" free will.

People are dynamic and responsive to their environment. We are in constant relation to the world around us and have only the slightest crack of consciousness through which to view the complexity of these interrelations. Observable behaviors and thoughts can be activated from within the person and from outside of the person. If activation is internal, it is caused by interactions between conscious processes and automatized, non-conscious functioning, as well as physiological, neuronal, and chemical activity. The small amount of mental behavior we know about is influenced by the vast majority of activity we do not know about, and so there is ample room for skepticism about how accurate those conscious attributions actually are.

Activation that comes from the environment, from other people, from circumstances, is registered both consciously and non-consciously. We are bundles of capacities, such that we have real ability to effect change in our environment. By the same token, we are bundles of receptors and are forever shaped by information from world around us. The information that receives conscious attention is always affected by the processing that happens below the level of awareness. We do not have the ability

to recognize all the ways in which our internal processes are shaped and influenced by external causes, but through experiment, we can demonstrate how easily that manipulation can be made and the potency of its effect. We are the products of our genes, our environment, our society, our culture, and the schemas we use to make sense of the world. All of these combine to make us capacious agents who seek order. We are truly complex, and it is difficult, upon honest philosophical analysis, to draw a definitive line between the internal and external causes of behavior. We are much more interesting than lone, isolated, individual "selves" that will this or that based on our own desires.

Based on this model, our actions and behaviors have far-reaching consequences that we cannot ultimately track. Our behaviors shape the environment around us, and it, in turn, influences our ideas and subsequent actions. A mere gesture may have the capacity to trigger in another person an entire chain of cognitions whose effect resonates in surprising and unexpected ways. If anything, then, behaviors and actions become more important in this universe than in a universe of free will. In the latter, isolated agents respond to stimuli in more or less their own fashion. It is the difference between behavioral chunks and a behavioral stream.

What role does the individual play in this model, and how do we establish responsibility? To answer this, let us consider that there is a fundamental difference between a free agent and an agent. A free agent is one who acts according to his own free will. In contrast, let us define an agent as one who acts. According to our model, persons are simply agents; free agents, in fact, do not exist. To go further, an agent is a bundle of capacities that brings about an action.

Let us make one further distinction with regard to agency. An active agent is one who has the capacity to influence the causal stream, whereas a passive agent is one who is hopeless to have any effect. Libertarians who argue for the depressing "cog in the machine" interpretation of determinism are supposing that the result of determinism is passive agency. This, though, is really a form of fatalism, rather than the kind of complex interactionism assumed in this paper. A fatalist is one who claims that human actions are irrelevant and inefficacious because whatever will be, will be (Bernstein 2002). From the fatalist perspective, it makes no difference whether or not I finish this sentence or finish college, because whatever is *supposed* to happen will happen. Fatalism assumes people to be passive agents of destiny.

There is a key distinction, though, between the fatalistic view of persons as pawns and the more optimistic view of persons as bundles of capacities. Insofar as agents are in constant interaction with their environment, according to this thesis, our actions *do* matter, and, in fact, do influence the causal stream. Moreover, under this model, we, the collection of capacities that constructs our identity, are active contributors to the environment of which we find ourselves a part. The effects of our actions might be subtle or far-reaching, but they are not sucked into the vacuum of predestination; rather, they are actively involved in the complex causal process of the moment. We are not free agents, and our ability to cause change in the world is mediated by the contextual factors of the life-space at the time of acting. Thus, our capacity to act is interdependent with the environment. However, we are active nonetheless. We are dynamic organisms interacting in a larger context, shaping and being shaped by the causal stream of the environment and by our own internal

processing. We are not "in control" of the situation, but we influence and help to create the circumstances of the interaction.

Under this capacity view, then, we can still identify active agents and from this foundation begin to build a theory of responsibility. However, it is a responsibility that is circumstantial and situated within a context. As such, this form of responsibility may be more pragmatic than metaphysical, and it would force us to call into question many of our current assumptions about and uses for the concept of responsibility. We would need to situate our claims of responsibility in appropriate contextual frameworks.

Take, for example, a judgment of responsibility for criminal behavior. Ted Bundy was prosecuted and convicted for the vicious killing of a number of people. A classic libertarian retort to an assault on free will might be that if there is no free will, we cannot hold Bundy responsible for his actions because he did not freely choose them. This claim assumes that responsibility is a metaphysical construct that is deeply interwoven with the psychological act of choice. However, our decision to put Ted Bundy in jail really has little to do with his choices, and has much more to do with his repeated actions and the fact that he has violated a social contract. Our judgments of responsibility typically have more to do with pragmatism than with metaphysics.

Bundy may not be a free agent, but he is an active one. The fact is that as an agent he is held responsible for his behavior insofar as he is the bundle of capacities that interprets the world through anti-social biased perceptions, which bring about repeated anti-social behavior. When we say that Bundy is "responsible" for his actions, what we mean, on the most basic level, is that he has engaged in behaviors that carry with them their own socially-proscribed consequences.

At some point, we come face-to-face with the question of why we assign

responsibility. If we are defenders of the concept of free will, we may assign

responsibility so that we can adequately reward or punish the person for his choices.

Anti-libertarians, though, might use responsibility as a reference point from which to

identify and analyze patterns. To return to our criminal justice example, then, the point

of imprisonment from a libertarian perspective may be to punish the individual for his

atrocious choices. Of course, due to our outstandingly high rates of recidivism in this

country, we might question the effectiveness of such a strategy. From an anti-libertarian

perspective, though, imprisonment might serve a number of purposes. Firstly, it may

remove individuals like Ted Bundy (with recurring patterns of anti-social behavior)

from society in the interest of the greater good of society. Second, it may serve a

rehabilitative purpose insofar as an individual can be psychologically rehabilitated.

That is, we may analyze recurring causes of behavior and attempt to introduce new, pro-

social internal processing with a more conducive environmental (external) setting. Our

goal, then, is not to punish the individual's character but to determine the extent to

which his behavior is malleable and then to make an informed decision taking into

account the context of the situation, which, in this case, would be the welfare of society

as a whole (for a more thorough review of treatment of criminal behavior that does not

appeal to the idea of free will, see Pereboom 2001).

In addition to situating our claims of responsibility in given contexts, an ethic of

social responsibility emerges when we give up the illusion of free agents. Under a

capacity model, people's behaviors are involved with the environments, social cues, and

circumstances of the time, which are, in turn, reinforced by people's behaviors. When

we do see recurring patterns of behavior, we must look for broader environmental

causes. If homelessness is an epidemic, it is not enough to place responsibility on the homeless, for, according to this model, the causes of homelessness have much less to do with the individuals who are homeless and much more to do with a larger, more complex fabric of social causal processes. If violence is a problem, do we see it as something whose causes and effects restrict themselves to the perpetrators and victims of violence or do we look more deeply at the web of thoughts and actions that are a part a larger phenomenon of violence?

With respect to morality, then, although we may still posit individual responsibility in a universe without free will, many claims about morality are better understood as circumstance- or action-claims rather than claims about an autonomous agent. It is important to note that judgments of right and wrong are not absent in such a universe. However, since the concept of a truly independent, autonomous agent is incoherent, these attributions may be most accurately applied to moral and immoral circumstances or actions, rather than judgments about the individual actors themselves (Pereboom 2001).

I have offered the above thought experiment not as an answer to the problem of responsibility, but as a possible conceptualization of responsibility without an appeal to free will. There is, without a doubt, a shared responsibility that comes with the rejection of the notion of free will. Since we are in constant interaction with others and together influence the larger interactive and dynamical stream, we have responsibility that extends beyond our immediate needs and desires. While free agency is rejected, active agency is still maintained by the thesis of capacity, and therein can reside responsibility. As Frankfurt's (1969) argument suggests, freedom is not the necessary condition for responsibility, unless of course we define freedom and responsibility reflexively.

Rather, responsibility is better conceptualized if it is situated within a context and used as a jumping off point, from which to discern a more complex, descriptive factor analysis of behavior.

CONCLUDING REMARKS AND A REFLECTION ON THE VALUE OF INTERDISCIPLINARY WORK

Philosophy holds a unique place among academic disciplines. It is, according to some, the oldest discipline, and plays a fundamental role in every other academic endeavor. Indeed, when one has reached the highest point of knowledge in any area, she receives a Ph.D., a doctorate of philosophy, that symbolizes her understanding of both the facts of the discipline and the reasons for those facts, the relevant epistemology. Finding commonalities between philosophy and any other discipline is, in some sense, a circular endeavor, insofar as the experts in any field are operating within its philosophical framework.

Still, psychology and philosophy have a unique link, since many of the first psychologists were themselves philosophers by trade. Both disciplines, of course, make claims about reality and offer theories and explanations for observed phenomena in the world. Both are seeking some kind truth about the human condition. And this, I think, is the first shared value: both disciplines look very closely at the experiential component of human existence and try to situate that subjective experience into a larger framework of reality as a whole. Psychologists and philosophers alike dwell in the complex interaction between the individual and her environment and try to provide explanations and paradigms for understanding that vast intersection.

Second, psychology and philosophy are equally concerned with the inner life of the individual. A person's capacities, her thoughts, emotions, beliefs, values, personality, and predispositions, are given particular importance and are viewed as significant factors in the human experience. The ways in which this inner life can be

transformed into outward behaviors and deeds is the fascinating domain of both disciplines.

In addition, social psychology and philosophy operate in a distinctly egalitarian fashion. That is, both take as their object of study the common experiences of otherwise scientifically-unremarkable persons. The everyday functioning of "typical" people is of particular interest to these specialists in a way that it is not in other social sciences. History, for example, (and here I am speaking in general terms) takes as its subject the important events and persons of the past. It focuses on the remarkable. Anthropology looks at the values, customs, and practices of civilizations of past and present and speaks globally about their significance. By comparison, social psychology is concerned with the more mundane nuts and bolts of day-to-day human functioning. And philosophy, while it surely can be classified among the most erudite of disciplines, does seek an understanding that is based in the common experience of human existence.

Finally, the shared goal of commenting upon the human condition places practitioners of both fields into an abstract and complicated role. Presumably if I were to ask a fish swimming in a stream to comment on the nature of water, she would have a difficult time doing so since water, as the primary substance of the fish's life, cannot be experientially distinguished from not-water. In the same way, philosophers and psychologists, in their attempt to analyze, describe, and reveal the functioning of human beings, cannot take a perspective that is distinctly non-human. Psychologists and philosophers place themselves in the inimitable position of using their own human faculties to understand human faculties generally. And so it is that we are potentially bounded by the object of our study and require multiple perspectives to make coherent and accurate claims about the human condition.

Because these two disciplines share common core values, the work done in each realm can be synthesized to advance knowledge. In more concrete terms, our investigation of the problem of free will is one that requires the work of both philosophy and psychology: philosophical thinking is needed to raise the question and to provide the framework under which we may propose responses; psychology, in turn, because it shares the common goal of explaining and understanding the human condition, can contribute its empirical justification for the philosophical claims that are made. In addition, the philosophical perspective challenges us to examine the consequences of our possible responses. The two perspectives working together help to refine shared conclusions, and encourage us as researchers to question the "usual" answers and presumptions that guide our thinking about a topic.

Most people tend to accept as true the idea that problems are best solved when considered from a variety of perspectives. In academia, there seems to be, at once, an overt affirmation of this principle and a less obvious denial. University curriculum values the liberal education, which exposes students to a number of ways of making sense of the world. At the same time, though, exclusivity is promoted within the disciplines themselves. Psychologists should not write like philosophers; philosophers should not sound like sociologists. Collaboration between the disciplines, at least in my view, is not nearly what it could be.

There are two primary advantages to interdisciplinary work. The first comes in the form of multiple hypotheses for understanding phenomena in the world. By looking at an academic question using the tools, concepts, and theories of more than one domain, investigators are more likely to come across similarities and differences in perspective. When commonalities are found, these can be seen as assurances that the

combined investigation is on the right track, and that the conclusions are not merely the idiosyncratic "plugging and chugging" of applying the usual disciplinary formulas to new problems. When differences emerge, like the now-tired story of the blind men and the elephant, investigators can apply these as additional pieces to the puzzle, or disregard dissenting viewpoints (binary conceptions of the problem, for example) in favor of a new paradigm as the next step in the study.

The second benefit a multi-disciplinary approach allows us is an outsider's look into the values and assumptions of the fields in which we work and the ability to reduce, to the extent possible, the biases of our own viewpoint. We may see that the different conclusions generated by researchers in a specific domain are actually best explained in terms of the implicit values used to perceive and interpret reality. From an interdisciplinary standpoint, then, we gain a clearer view of the specific questions each field can answer, and how they may be combined into a shared understanding about reality.

The question of free will, like any question that is bound up in the experience of people, cannot be merely a philosophical inquiry. Viewed from one perspective only, the questions asked by researchers tend to suggest the desired answer. A relational approach to problem solving, in which the expertise of multiple fields can be combined to one effect is preferable. Answering the question of free will without the input of psychology, it seems to me, is akin to men asking about women's issues without consulting women and just suspecting that men's intuitions are reasonably grounded. Such an approach is presumptive and likely to arrive at the wrong answers.

When investigated from multiple disciplines, calling upon the wisdom of metaphysicians, philosophers of science, and psychologists in this case, we see that on

balance free will is both an inaccurate articulation of human functioning and an unnecessary assumption for social functioning and personal identity. Instead, I have urged that scientific analysis of the question provides support for the philosophical claim of persons as beings of capacity, not unlike the rest of the world. As such, we live in interaction and interdependence with our life-space. We are humbled by our reliance upon contextual factors for agency, but impressed by the capacities that we as humans do, relationally, possess.

The idea of free will is not useful to our understanding our place in the world. Rather, it is an impediment to our self-perception. So, too, does linear Laplacean determinism fail to embrace the complexity of life as we live it. These concepts, it seems to me, are wrapped in values of dominance, autonomy, and independence that miss the mark of real life. Philosophy and psychology are at their best when describing interactions and relationships between organisms, factors, persons, and systems. We do better investigating the ways in which we are involved with the world around us than we do trying to separate ourselves as autonomous and free agents in a world of clockwork. Rejecting these extremes leads us to a more thorough and careful look at our own experience.

References

Anderson, J. R. (1983). *The Architecture of Cognition.* Cambridge, MA: Harvard University Press.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer (Ed.), *Advances in Social Cognition (Vol. 10)* (pp. 1-62). Mahwah, NJ: Lawrence Erlbaum Associates.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54(7)*, 462-479.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71(2),* 230-244.

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin, 126(6),* 925-945.

Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology, 43,* 437-449.

Baumeister, R. F., & Sommer K. L. (1997). Consciousness, free choice, and automaticity. In R. S. Wyer (Ed.), *Advances in Social Cognition (Vol. 10)* (pp. 75-81). Mahwah, NJ: Lawrence Erlbaum Associates.

Beck, A. T. (1967). *Depression: Clinical, Experimental, and Theoretical Aspects.* New York: Harper and Row.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders.* New York: International Universities Press.

Berkowitz, L. (1997). Some thoughts extending Bargh's argument. In R. S. Wyer (Ed.), *Advances in Social Cognition (Vol. 10)* (pp. 83-94). Mahwah, NJ: Lawrence Erlbaum Associates.

Bernieri, F., & Rosenthal, R. (1991). Cordinated movement in human interaction. In Feldman & Rime (Eds.) *Fundamentals of Nonverbal Behavior* (pp. 401-432). New York: Cambridge University Press.

Bernstein, M. (2002). Fatalism. In R. Kane (Ed.), *The Oxford Handbook of Free Will* (pp. 65-84). New York: Oxford University Press.

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin, 99,* 229-246.

Blumenthal, J. (2002). Personal interview. October 2002.

Bruner, J. (1986). *Actual Minds, Possible Worlds.* Cambridge, MA: Harvard University Press.

Clore, G., & Ketelaar, T. (1997). Minding our emotions: on the role of automatic, unconscious affect. In R. S. Wyer (Ed.), *Advances in Social Cognition (Vol. 10)* (pp. 105-120). Mahwah, NJ: Lawrence Erlbaum Associates.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82,* 407-428.

Corey, G. (2001). *Theory and Practice of Counseling and Psychotherapy.* Belmont, CA: Wadsworth.

Damasio, A. R. (1994). *Descartes' Error: Emotions, Reason, and the Human Brain.* New York: Avon Books.

Dennett, D. (1984). *Elbow Room.* Cambridge: MIT Press.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5-18.

Dupré, J. (1993). *The Disorder of Things: Metaphysical Foundations of the Disunity of Science.* Cambridge, MA: Harvard University Press.

Efran, J. S., Lukens, M. D., & Lukens, R. J. (1990). *Language, Structure, and Change.* New York: W. W. Norton and Company.

Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin, 117,* 39-66.

Frankfurt, H. G. (1969/2003). Alternate possibilities and moral responsibility. In G. Watson (Ed.), *Oxford Readings in Philosophy: Free Will, Second Edition* (pp. 167-176). New York: Oxford University Press.

Frankl, V. (1962). *Man's Search for Meaning: An Introduction to Logotherapy.* Boston: Beacon Press.

Gazzaniga, M. S., & Le Doux, J. E. (1978). *The Integrated Mind.* New York: Plenum.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23,* 183-209.

Gleick, J. (1987). *Chaos: Making A New Science.* New York: Penguin.

Glass, D. C., Reim, B., & Singer, J. E. (1971). Behavioral consequences of adaptation to controllable and uncontrollable noise. *Journal of Experimental Social Psychology, 7(2)*, 244-257.

Glass, D. C., Singer, J. E., & Friedman, L. N. Psychic cost of adaptation to an environmental stressor. *Journal of Personality and Social Psychology, 12(3)*, 200-210.

Goodman, C. (2002). Resentment and reality: Buddhism on moral responsibility. *American Philosophical Quarterly, 39(4)*, 359-372.

Hefferline, R. F., Keenan, B., & Harford, R. A. (1959). Escape and avoidance conditioning in human subjects without their observation of the response. *Science, 130,* 1338-1339.

Hume, D. (1738/2001). *A Treatise of Human Nature.* D. F. Norton & M.J. Norton (Eds.) New York: Oxford University Press.

Ingram, R. E., Smith, T. W., & Brehm, S. S. (1983). Depression and information processing: Self-schemata and the encoding of self-referent information. *Journal of Personality and Social Psychology, 45(2),* 412-420.

Jacobs, G. D. (2003). *The Ancestral Mind.* New York: Penguin Books.

Jaynes, J. (1976). *The Origin of Consciousness in the Breakdown of the Bicameral Mind.* Boston: Houghton Mifflin.

Johnson, G. (1995). *Fire in the Mind: Science, Faith, and the Search for Order.* New York: Random House.

Jones, E.E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology, 3,* 1-24.

Kane, R. H. (1996). *The Significance of Free Will.* New York: Oxford University Press.

Kunda, Z. (1999). *Social Cognition: Making Sense of People.* Cambridge, MA: MIT Press.

Lambert, A. J. & Raichle, K. (2000). The role of political ideology in mediating judgments of blame in rape victims and their assailants: A test of the just world, personal responsibility, and legitimization hypotheses. *Personality and Social Psychology Bulletin, 26(7),* 853-863.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32(2),* 311-328.

Lefcourt, H. M. (1973). The function of the illusions of control and freedom. *American Psychologist, 28(5),* 417-425.

Lewis, D. (1981/2003). Are we free to break the laws? In G. Watson (Ed.), *Oxford Readings in Philosophy: Free Will, Second Edition* (pp. 122-129). New York: Oxford University Press.

Locke, J. (1690/1974). An essay concerning human understanding. In *The Empiricists* (pp. 7-134). New York: Anchor Books.

Mackie, J. L. (1974). *The Cement of the Universe: A Study of Causation*. Oxford: Clarendon Press.

Macy, J. (1991). *Mutual Causality in Buddhism and General Systems Theory: The Dharma of Natural Systems*. New York: State University of New York Press.

Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology, 35,* 63-78.

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist, 44,* 1469-1481.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, *82*, 213-225.

Mischel, W. (1997). Was the cognitive revolution just a detour on the road to behaviorism? On the need to reconcile situational control and personal control. In R. S. Wyer (Ed.), *Advances in Social Cognition (Vol. 10).* Mahwah, NJ: Lawrence Erlbaum Associates, 181-186.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244,* 933-938.

Nagel, T. (1986/2003). Freedom. In G. Watson (Ed.), *Oxford Readings in Philosophy: Free Will, Second Edition* (pp. 229-256). New York: Oxford University Press.

Narvaez, D., Getz, I., Rest, J. R., & Thoma, S. J. (1999). Individual moral judgment and cultural ideologies. *Developmental Psychology, 35(2),* 478-488.

Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84(3),* 231-257.

Nørretranders, T. (1991). *The User Illusion*. New York: Viking Books.

Ornstein, R. (1991). *The Evolution of Consciousness: The Origins of the Way We Think.* New York: Touchstone.

Pereboom, D. (2001). *Living Without Free Will*. New York: Cambridge University Press.

Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. Rabbitt & S. Dornic (Eds.), *Attention and performance: Vol. 5* (pp 669-682). San Diego, CA: Academic Press.

Reeve, J. (2001). *Understanding Motivation and Emotion, 3rd Edition*. New York: John Wiley and Sons.

Reid, T. (1792/2001). Of power. *Philosophical Quarterly, 51(202)*, 3-12.

Rosch, E., Mervis, C. G., Gray, W. D., Johnson, D. M., & Bayes Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382-439.

Russell, P. (2002). Pessimists, Pollyannas, and the new compatibilism. In R. Kane (Ed.), *The Oxford Handbook of Free Will* (pp. 229-256). New York: Oxford University Press.

Sartre, J.-P. (1957/2004). Existentialism. In G. Marino (Ed.), *The Basic Writings of Existentialism*. New York: The Modern Library, 341-367.

Schachter, S. & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological review, 69(5)*, 379-399.

Scher, C. D., Segal, Z. V., & Ingram, R. E. (2004). In R. L. Leahy (Ed.), *Contemporary Cognitive Therapy: Theory, Research, and Practice* (pp. 27-44). New York: Guilford Press.

Searle, J. (2004). *Mind: An Introduction*. New York: Oxford University Press.

Shane, H. (1993). The dark side of facilitated communication. *Topics in Language Disorders, 13(4)*, ix-xv.

Shiffrin, R. M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and general theory. *Psychological Review, 84*, 127-190.

Skinner, B. F. (1971). *Beyond Freedom and Dignity*. New York: Alfred A. Knopf.

Smart, J. J. C. (1961). Free will, praise, and blame. *Mind, 70*, 291-306.

Smith, E. E. (1990). Categorization. In D. N. Osherson and E. E. Smith (Eds.), *Thinking: An Invitation to Cognitive Science* (pp. 33-53). Cambridge, MA: MIT Press.

Smith, E. E., & Medin, D. L. (1981). *Categories and Concepts.* Cambridge, MA: Harvard University Press.

Strawson, P. (1963/2003). Freedom and resentment. In G. Watson (Ed.), *Oxford Readings in Philosophy: Free Will, Second Edition* (pp. 72-93). New York: Oxford University Press.

Van Inwagen, P. (1983/2003). Argument for incompatibilism. In G. Watson (Ed.), *Oxford Readings in Philosophy: Free Will, Second Edition* (pp. 38-57). New York: Oxford University Press.

Van Inwagen, P. (2002). Free will remains a mystery. In R Kane (Ed.), *Oxford Handbook of Free Will* (pp. 158-177). New York: Oxford University Press.

Watson, G. (2003). Introduction. In G. Watson (Ed.), *Oxford Readings in Philosophy: Free Will, Second Edition* (pp. 1-25). New York: Oxford University Press.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review, 101,* 34-52.

Wegner, D. M. (1999). The seed of our undoing. *Psychological Science Agenda, January, February 1999*, 10-11.

Wegner, D. M. (2002). *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.

Wegner, D. M., & Vallacher, R. (1980). *The Self in Social Psychology*. New York: Oxford University Press.

Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist, 54*: 480-491.

Wilson, T. D. (2002). *Strangers to Ourselves: Discovering the Adaptive Unconscious*. Cambridge, MA: Belknap Press of Harvard University Press.

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophesies in interracial interaction. *Journal of Experimental Social Psychology, 10,* 109-120.

Zillman, C. J. R. & Bryant, J. (1975). Enhancement of experienced sexual arousal in response to erotic stimuli through misattribution of unrelated residual excitation. *Journal of Personality and Social Psychology, 32(1)*, 69-75.