

UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Belief and Responsibility: An Essay on Control

Permalink

<https://escholarship.org/uc/item/1xp1c1cf>

Author

Bright, Joshua Eric

Publication Date

2010

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Belief and Responsibility: An Essay on Control

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Philosophy

by

Joshua Eric Bright

August 2010

Dissertation Committee:

Dr. Peter J. Graham, Chairperson

Dr. John Martin Fischer

Dr. Dion Scott-Kakures

Copyright by
Joshua Eric Bright
2010

The Dissertation of Joshua Eric Bright is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGMENTS

Only one person gets to be listed as the author of a dissertation, which is a base tragedy for something that is so fundamentally a collective effort. My list of debts is long and will of necessity be incomplete, as I'm sure that I've failed to recollect all those who have helped me along the way.

I want to express my profound thanks to the members of my committee. To John Fischer, who has been unfailingly encouraging, accessible and insightful throughout the process, and whose philosophical theory provided me with the framework for the core of my argument. To Dion Scott-Kakures, who has given generously of both his time and his copious philosophical talents, and who has repeatedly demonstrated the intellectual virtue of charity in both his reading of and response to my work. And most of all to Peter Graham, who has been an intellectual light and an immensely practical guide throughout not only my dissertation, but my entire graduate program. From Peter I learned the crucial insight that philosophy is not just a subject but also a practice and an academic field that must and can be skillfully navigated. His philosophical insight, unfailing responsiveness despite his many other commitments, and endless patience with my project are responsible for the majority of the virtues my dissertation possesses. The remaining faults are of course solely mine.

A variety of other philosophers provided me with helpful feedback and inspiration on portions of this project at conferences, in reading groups, and in individual conversation. My thanks to Jonathan Adler, Jack Crumley, Andrews Reath, Eric

Schwitzgebel, Dan Speak, Matthias Steup, Matt Talbert, Kevin Timpe, Chris Yeomans and Howard Wettstein. Three contributions in particular stand out. David Owens provided me not only with the inspiration for Chapter 2, but also with a generous gift of his personal time, both in conversation at the Pacific APA and in feedback on my analysis of his work. Gary Watson's class on Belief and the Will gave me the inspiration for my project, and his feedback and guidance on my dissertation were invaluable. And Paul Hoffman, through his selfless leadership of the UCR Dissertation Workgroup, and through his trenchant critiques of nearly every chapter of my work, provided me with exactly what I needed at the toughest part of this process. His recent and untimely death leaves the world without both a tremendous philosopher and a wonderful person.

My graduate colleagues and friends also supplied great assistance with their feedback as well as the sharing of the experiences, both large and small, that make up the strange world that is graduate school. Jacob Affolter, Chris Franklin, Joshua Hollowell, Felipe Leon, Dave Meyer-Reith, Nathan Placencia, Robert Sanchez, Chip Steers, Nathan Westbrook, Dwight Yao and Jeff Yim all provided help along the way. Two such individuals deserve special thanks: Matt Heard and Neal Tognazzini. Both Matt & Neal have proceeded me in the path (which I hope to follow) into the professoriate, and I consider them not only wonderful friends but also intellectual mentors. Both also spent countless hours with me swapping stories, providing feedback, and sharing life. Without their dear friendship, for which I am eternally grateful, there is no possible world in which I would have finished this work.

A variety of personal debts remain. Thanks to the kind staff of the Canyon Crest Starbucks, who have provided me with the one thing every dissertator desperately needs: a warm, comfortable and inviting place to write (oh yes, and caffeine!!).

A much greater thanks to the management of UCR Computing & Communications, where I have worked throughout my graduate career. Thanks especially to Chuck Rowley, David Gracey and Larry McGrath, who have exemplified kind and caring leadership, who have allowed me to provide for my family and to develop personally and professionally, and who I am proud to count as friends.

I hope it will not seem odd if I also pause at this way-station on the vocational journey to thank God. I'm not quite sure how it functions a piece of natural theology, but completing this process is a tremendous testimony to me of God's faithfulness and kindness, manifested largely in the persons I mention here, but also in variety of other ways great and small.

Finally, thanks to my family. My beautiful children Jacob, Grace & Katelyn have loved their Daddy unconditionally through years of sleep-deprived crankiness. And my wife Dianne has quite simply been the foundation of love that made it all possible. This dissertation and my graduate studies, which have been such huge parts of our lives for such a long time, represent but a sliver of the debt of gratitude and love that I owe her. Many of my debts I might consider paid with these comments, but hers I recognize with joy I will spend a lifetime repaying, and still not balance the scales.

This work is dedicated to three persons:

To Dad, who always told me to do what I love,

To Mom, who always believed I could do anything I wanted to do,

And to Dianne, who actually made it possible.

ABSTRACT OF THE DISSERTATION

Belief and Responsibility: An Essay on Control

by

Joshua Eric Bright

Doctor of Philosophy, Graduate Program in Philosophy
University of California, Riverside, August 2010
Dr. Peter J. Graham, Chairperson

My dissertation attempts to establish the plausibility of a thesis that I call moderate direct control over belief. This thesis tells us that a significant percentage our beliefs, or more specifically their formation, retention and elimination, are subject to our direct control. Moderate direct control is in direct conflict with the dominant view in philosophy, which claims that we lack direct control over belief. I make my argument first by appealing to common sense, including our standard responsibility-attributing practices and experiences. I suggest that an intuitive and defensible interpretation of these factors provides a basic presumption in favor of direct control which the skeptic must overcome. I then argue that such a refutation is not available, first by offering theoretical arguments in support of direct control, and then by examining and criticizing various responses offered in the relevant literature on doxastic control and responsibility.

In this latter task of providing theoretical support, I try to proceed in as neutral a manner as possible. It is impossible to operate without a theoretical framework of some kind, so I instead operate by identifying the major distinctions in the relevant literature

(intellectualist, voluntarist, libertarian and compatibilist), and then in turn adopting as assumptions the various positions available within the debate over doxastic control. In each case, the broad structure of my argument is to examine relevant criteria of control over action offered within the given theoretical space, and then to show its applicability to the formation, retention and dissolution of belief. I thus do not attempt to defend a particular account of the necessary and sufficient conditions of control, but rather argue that those persons who embrace the reality of control and responsibility over action within one of these conceptions (which together comprise the majority of viewpoints on control) are thereby committed to a similar view regarding belief.

TABLE OF CONTENTS

Table of Contents	x
Introduction.....	1
I.1. In Defense of Common Sense	2
I.2. Assumptions & Clarifications	5
I.3. Dissertation Outline.....	9
Chapter 1 – Common Sense and Direct Control Over Belief.....	15
1.1 Responsibility for Belief.....	15
1.2. The Language & Phenomenology of Belief	20
Chapter 2 – Moderate Direct Control Over Belief on Intellectualism.....	38
2.1. Setting Aside the Will.....	38
2.2. Intellectualism & Voluntarism.....	40
2.3. Intellectualism and Belief	45
2.4. Owens on Belief.....	58
2.5. The Idle Wheel of Reflection.....	66
2.6. Reflection and the Intrinsic Authority of Judgment	72
2.7. A Unified Conception of Judgment.....	75
2.8. Pragmatic Factors and Reflective Control	83
2.9. Conclusion	87
Chapter 3 – Moderate Direct Control on a Voluntarist Framework	89
3.1. Returning to the Will	89
3.2. The Normative Structure of Belief Determination	92

3.3. Conceptions of Control: Guidance & Regulative	103
3.4. Guidance Control Over Belief Determination	111
3.5. Regulative Control Over Belief Determination	118
3.6. Intention: A Potential Objection	131
Chapter 4 – Voluntarism: Conceptual Arguments Against Direct Control	137
4.1. Conceptual & Psychological Objections	137
4.2. Conceptual Arguments.....	138
4.3. Williams.....	139
4.4. Scott-Kakures.....	149
4.5. Adler	155
4.6. Audi.....	171
4.7. Conclusion	175
Chapter 5 – Voluntarism: Psychological Arguments Against Direct Control	177
5.1. Incoherence vs. Implausibility	177
5.2. Psychological Arguments	178
5.3. Audi Redux	181
5.4. Alston.....	192
5.5. Conclusion	198
Conclusion	200
Bibliography	205

INTRODUCTION

“That’s the way it’s done” the Queen said with great decision: “Nobody can do two things at once, you know. Let’s consider your age to begin with – how old are you?”

“I’m seven and a half exactly.”

“You needn’t say ‘exactly,’” the Queen remarked: “I can believe it without that. Now I’m going to give You something to believe. I’m just one hundred and one, five months and a day.”

“I can’t believe That!” said Alice.

“Can’t you?” the Queen said in a pitying tone. “Try again: draw a long breath, and shut your eyes.”

Alice laughed. “There’s no use trying,” she said: “one Can’t believe impossible things.”

“I daresay you haven’t had much practice,” said the Queen. “When I was your age, I always did it for a half-an-hour a day. Why sometimes I’ve believed as many as six impossible things before breakfast...”

--Alice & the White Queen
Through the Looking Glass
Lewis Carroll

1.1. In Defense of Common Sense

While the above passage from Lewis Carroll's classic¹ is surely a bit of fanciful hyperbole and fun, it nevertheless articulates a certain conception present in recent philosophical literature on the subject of belief. The majority of philosophers writing on the topic view beliefs as things not subject to direct control, and they sometimes seem to imply that those who defend direct control are adopting a view similar to the White Queen's.² The putative images of direct control are of persons exerting themselves by sheer willpower to believe, if not something impossible, at least something deeply implausible. Defenders of direct control are implied to be defending a position that appears dubious and even a bit epistemically pernicious.

As a defender of direct control, I have no such desire to defend the impossible, or even the implausible. Instead, in this work I take myself to be offering a defense of a common sense view of belief. It is our uniform and common practice to use normative language regarding beliefs. Had Alice & the Queen's dialogue continued –instead, the Queen turns into a sheep on the next page– we might have naturally expected Alice's implicit rebuke of the Queen in the dialogue quoted to be made explicit. This is because we regularly offer praise and blame to ourselves and others over the beliefs that we form, hold, and/or discard. But are we right to make such evaluations? Does it make sense to praise or blame someone for their beliefs? Are we in fact responsible for what we believe, properly blameworthy or even praiseworthy for our beliefs? Surely such questions are

¹ Carroll (1897), pp. 102-103.

² See, for example, Alston (1988), p. 263.

significant. Unless we can find a sufficient ground for such statements, one of our common and universal linguistic practices stands on shaky footing. Moreover, in answering these questions, we gain further insight both into belief as well as our general responsibility attributing practices.

Much of the philosophical reflection on these questions has occurred under the rubric of “doxastic voluntarism”. This phrase’s etymological roots tell us that it is concerned with the question of whether belief (*doxa*) is subject to or closely connected with the will (*voluntas*). The general (though not univocal) philosophical consensus in current literature is that though it is occasionally appropriate to hold one another responsible for beliefs, beliefs are nevertheless subject to the will in at most an indirect and tenuous way.

I question this philosophical consensus in two distinct ways. First, I take issue with the claim that belief is only tenuously connected with the will. But secondly and perhaps more importantly, I question the approach or framework in which this philosophical discussion has generally occurred. The role of the will is certainly significant in answering the question of doxastic responsibility, but only because of the will’s role as a principal mechanism of control for human agents. Control is the fundamental issue for responsibility, and the primary questions in need of answering include “Are our beliefs under our control, and if so, how?”, and “Is this control, if it exists, primarily indirect, or does it also include direct control?” I argue in this work that we do possess doxastic control, and that it is often direct. My claim is that arguments about control over belief have largely gone wrong not because of misunderstandings

about belief, but rather because of misunderstandings about control, and the applicability of notions of control to belief.

This differing focus leads to a new approach, which I outline in the following sections. Much of the literature on doxastic voluntarism has an epistemological bent, taking its cues from the categories of epistemology and concerning itself with the epistemic norms governing belief. This often leads to a focus on the dissimilarities on between belief and action, which I believe contributes strongly to the consensus against voluntarism for belief. But if the fundamental issue is control, as I argue, then significant resources may be available to us from the philosophy of action. This approach leads to an emphasis not on the norms of belief but rather on the mechanisms through which these norms are expressed, and highlights not the differences between belief and action, but rather the similar way in which we exercise control over both our beliefs and our actions. To be clear, there is no strong analogy between belief and action. Too many dissimilarities exist for this to be so. But by taking as paradigmatic the accounts of the necessary and sufficient conditions of control offered in the philosophy of action and applying them to the realm of belief, we gain great insight into the direct control that we regularly exert over our beliefs.

Of course, using the phrase “the direct control we regularly exert over our beliefs” is likely to set off warning lights on the philosophical dashboard of many readers, who may think that such control is neither possible nor desirable. Indeed, the arguments against doxastic voluntarism often seem to be substantially motivated by the admirable (though misguided) conviction that to possess control over our beliefs would be a

dangerous, epistemically pernicious power. Let me again assure readers that I write only in defense of what I take to be common sense. As I note above, attributions of responsibility for belief are part of our daily interaction with each other. Surely the most straightforward (though far from the only) way to be responsible for something is to possess direct control over it. I believe that philosophers have been far too quick to dismiss this possibility for belief. Direct control over belief is surely not the whole story of doxastic responsibility. But I write from the conviction that it is a part of the story that has been significantly neglected and misunderstood. In the following chapters I defend a reasonable form of direct control over belief, first by explicating the common sense conviction that motivates my inquiry, and then by searching to see if there exists a plausible theory or theories to support the deliverances of our everyday experience.

1.2. Assumptions & Clarifications

All philosophical works proceed under the aegis of a variety of assumptions, and here I should clarify mine. I make certain assumptions about control. First, I assume that control is sufficient for responsibility; if one possesses control over some event or state of affairs, one can be properly held responsible for it. I take this assumption to be wholly uncontroversial, and am not sure how I would even go about defending it, were I to wish to. But it is worth noting that I do *not* assume the related claim that direct control is necessary for responsibility. Much of the critical discussion concerning doxastic control seems focused on disabusing this notion. While my discussion focuses on control as the

ground of responsibility for both belief and action, it does not in fact depend on any assumption about the necessity of control. If we in fact lack control over our beliefs, perhaps we still retain some form of responsibility for them. Certainly this is true if we lack the direct control that I am principally concerned with, it may be true even if we lack the indirect control that most (myself included) agree that we possess. This is an interesting philosophical question in its own right, worthy of the discussion that it has received, but it is not the question I attempt to address. My argument simply proceeds with the idea that the most straightforward way to establish responsibility is through direct control.

Secondly and more importantly, I work from the assumption that we in fact possess control over our actions. This assumption is somewhat more controversial, but I take it to be shared by the vast majority of philosophers and nearly all non-philosophers. As I noted just above, while I do not attempt to make any kind of strict analogy between action and belief, I do provisionally embrace different accounts of control over action as paradigmatic for evaluating claims of control over belief. This limits the appeal of my argument; those philosophers, for example, who believe that we lack control over our actions because no notion of control is ever satisfied (we may call them *hard determinists*) are unlikely to be convinced by my arguments concerning belief.³

A third assumption of my project is neutrality. Throughout the following chapters I articulate several distinct notions of control, but do not attempt to defend one particular

³ For an example of hard determinism, see Pereboom (2001).

conception over against the others. This neutrality is not a philosophical assumption; I certainly have a view on the nature of control over both belief and action. It is rather a dialectical assumption. My claims about control over belief are, I believe, generalizable across a variety of different conceptions of control, and I attempt to demonstrate this in the following chapters. My goal is not to convince my readers that one particular conception of control is correct, but rather to convince advocates of a variety of prominent conceptions of control that their preferred conception of control is satisfied in broadly the same way by both belief and action.

I also should say something here about belief. First, I need to very briefly explain the notion of belief I am working with throughout the dissertation. There are two principal conceptions of the nature of belief in philosophical literature. Belief is typically conceptualized as either an occurrent mental state or a dispositional property of an individual, or some combination of the two.⁴ While I will make no effort to defend the idea here, it seems to me that some combination view must be correct. My beliefs definitely include occurrent mental states which involve the adoption of a distinctive attitude toward a proposition (precisely what attitude this is another issue that I will generally avoid, though I will touch on the difference between acceptance and belief). But it seems probable to me that my beliefs are also dispositional in some sense. Surely only a tiny fraction of my beliefs are occurrent at any given time, and yet I don't cease to believe in those things not present before consciousness. My main claim regarding belief,

⁴ For one example of an account of belief that attempts to combine both dispositional and occurrent mental states, see Schwitzgebel (2002).

however, is that my arguments hinge not at all on the resolution of this debate. Whether belief turns out to be principally an occurrent or dispositional phenomenon, or to involve both of these in equal measure is largely immaterial to my thesis.

My account of control over belief is focused not on the mental or dispositional states that are our beliefs, but rather on the events that initiate, modify and terminate these states. I accept Robert Audi's point (which I discuss in Chapter 5) that *behavioral voluntarism*, the view that beliefs themselves are actions which are subject to our control, is a non-starter. Beliefs are states of a person, rather than actions. Instead, whatever control we possess relates to the events that initiate, modify, retain or terminate our beliefs. This view Audi calls *genetic voluntarism* (Audi speaks specifically only of the initiation of belief, but I think his point applies to the other actions mentioned as well, and my discussion concerns all of these different actions). While I don't adopt Audi's labels for my broader account, I nevertheless think that his assessment is accurate. And given that control over belief is focused not on the resulting states but rather on the events that initiate, modify, retain and/or eliminate them, the precise character of these states is of relatively little importance to my argument. Both standard dispositional and occurrent accounts of belief accept that similar types of events (deliberation, judgment, decision, etc.) are paradigmatically involved in inaugurating, changing, preserving and terminating our beliefs. The question then becomes whether these events are such that they can be said to be under our direct control. So, while the nature of belief is another question of great independent philosophical interest, I will have relatively little to say about it here.

In light of the above discussion, I also want to clarify the way in which I will speak about belief throughout the remainder of this work. In the paragraph above I've attempted to clarify the precise focus of my upcoming discussion of control, namely that it centers on the acts or events that initiate, modify, retain and terminate our beliefs. Control, insofar as we possess it, is over belief formation, modification, retention and termination. However, as the preceding paragraph also shows, speaking in such a precise way is exceedingly cumbersome and awkward. Were I to consistently use the formulation "control over the events that initiate, modify and terminate our beliefs", I would likely drive both myself and the reader to distraction. In hopes of avoiding this outcome I will use the phrase "control over belief determination" as a kind of shorthand for the more lengthy expression. Since belief formation, modification, retention and termination are all ways of immediately determining one's belief state, this abbreviation seems warranted. So I will typically discuss direct and indirect control over belief determination, and only use an alternative more precise locution occasionally, when the circumstances seem to require it.

1.3. Dissertation Outline

I begin my dissertation with an intuitive or common sense argument for direct control over belief formation. I have two principal concerns in Chapter 1: our standard linguistic practices with respect to belief, and the phenomenology of belief. I argue that both the ways in which we speak about belief and the ways we experience belief provide

support for the thesis of moderate direct control, which I define here. In this first chapter my goal is quite modest; I seek only to establish that there is a basic but significant presumption in favor of direct control over belief formation, one which the critics of control must overcome to make their case. Common sense supports the idea that we have direct control over at least some of our beliefs. Theory may undermine or support common sense, but it is important to establish the state of the dialectic at the outset of theoretical discussion.

Of course, I believe that theoretical reflection in fact vindicates common sense, and showing this is the task I begin in Chapter 2. I begin by laying out my general plan for the following chapters. Since my principal focus is on control, a central concern is to continue to clarify this concept. However, as I've mentioned above, I take no substantive position on the nature of control. Instead, I plan to draw the principal distinctions in the debate concerning control over action, and argue that those who hold any of the majority of the possible positions on control are committed to supporting direct control over belief formation. In each case, I will attempt to identify the necessary and sufficient conditions of control articulated with respect to action, and then to show how these conditions are satisfied with respect to belief formation.

Clearly outlining a theory requires the drawing of such distinctions, and in Chapter 2 I begin by attempting to clarify the difference between intellectualism and voluntarism, which will figure prominently in my project. This nettlesome distinction,

which sometimes goes by other names,⁵ concerns the appropriate characterization of control, and its locus within our mental activity. The intellectualist argues that control is in essence conformity to judgment, while the voluntarist contends that control necessarily involves the activity of the will. I examine this distinction through the use of several examples, setting up the discussion for the next several chapters.

Following this, I spend the bulk of Chapter 2 considering the case for direct control over belief formation on intellectualism. For the intellectualist, control is principally exercised through reflective judgment, and the appropriate measure of control is the efficacy of judgment in determining our actions or beliefs. I argue that it is clear that our reflective judgments often directly control our actions, and that *prima facie*, the same seems true of our beliefs. I offer several examples in support of this contention. I then consider the objections of David Owens, one prominent intellectualist who argues that we possess reflective control over our actions, but lack the same control over belief formation. Owens' arguments fail, I assert, but their failure nevertheless helps us to substantially clarify the account of intellectualist moderate direct control over belief formation.

In Chapter 3 I turn to control within a voluntarist framework, which I take to be the harder case. The essential role of the will poses a challenge for the doxastic voluntarist because, as many philosophers have noted, it is not clear that belief formation is subject to control by the will. Again, answering this question with any precision

⁵ See Chapter 2, n. 20.

requires getting clear on what we mean by control. I argue that while belief and action are governed by different norms, they nevertheless are similar with respect to their normative structure, by which I mean the way in which both belief and intention are governed by their respective norms, and the function of the salient kinds of reasons in the production of each of these mental states. I then look squarely at the characterization of control. To examine the issue further, I employ the distinction between *guidance control* and *regulative control* developed by John Martin Fischer in his prominent work in the philosophy of action. Guidance control (Fischer's view) is the position that control requires responsiveness to reasons and ownership of the mechanism that produces an intention, but importantly does not require access to alternative possibilities, the ability to do otherwise than one actually does. Regulative control, as a broad category of views, adds the additional libertarian requirement of alternative possibilities to these requirements. I argue that whether one accepts the general requirements of guidance control (as do many compatibilists) or regulative control (many libertarians), one must acknowledge the plausibility of direct control over belief formation given these requirements. Doxastic voluntarism is clearly plausible given guidance control, and if one accepts the possibility of libertarian (regulative) control generally, there is little reason to believe that it does not obtain for belief formation as well.

A substantial literature exists on doxastic voluntarism, and I spend the final two chapters of my work examining and responding to several representative arguments from this literature. The present philosophical consensus is against doxastic voluntarism, so there is no shortage of critics to choose from. The critical literature generally divides into

two basic types of arguments: conceptual and psychological. My fourth chapter is focused on conceptual arguments, which claim that our inability to exercise direct control over our beliefs, to modify our beliefs at will, is a conceptual truth grounded in the nature of belief, the nature of control, or some related necessary base. I consider the arguments of Bernard Williams, Dion Scott-Kakures, Jonathan Adler and Robert Audi. I argue that despite the challenges posed by these critics, the case for moderate direct control made in Chapter 3 escapes nearly unscathed.

In my final chapter I consider psychological arguments against direct control. These arguments, though varied, are unified by the admitted contingency of their claims, and their focus on human psychology. These critics claim that while the impossibility of direct control over belief may not be a conceptual truth, it is nevertheless the case that we, normal human agents, can exercise no such control. Here I recall the general presumption established in Chapter 1 in favor of direct control. This presumption has particular importance in these discussions because it means that the critic cannot make his or her case simply by offering alternative interpretations of cases that intuitively support direct control. Further argumentation against direct control is clearly required. However, Audi and William Alston, the critics whose psychological arguments I consider, seem to rest much of their case on precisely this method of alternative interpretation. I argue that their interpretations are unconvincing, and that the critical arguments they use to buttress these interpretations fail.

With the failure of these critics' arguments, the case for moderate direct control over belief formation is well established. This result, while a blow to philosophical

consensus, is a victory for common sense. It provides a clear and straightforward theoretical grounding for our ordinary forms of speech, our regular and universal practices of holding each other responsible for our beliefs. While as I've mentioned the question of whether control is in fact necessary for responsibility is philosophically interesting in its own right, we are not forced into this discussion in an attempt to rescue our common practices, as we have vindicated such responsibility talk with the establishment of control. Whether or not control is necessary, it is clearly sufficient for this purpose, and in showing the plausibility of direct control over belief formation we have provided a basic and clear-cut ground for our speech about and experience of belief.

Or so I say. The introduction of the above arguments is certainly easier than their execution, to which I now proceed. The reader will of course decide for herself whether I have made my case. Or perhaps, if I am mistaken, she will not.

CHAPTER 1 – COMMON SENSE AND DIRECT CONTROL OVER BELIEF

1.1. Responsibility for Belief

My topic is responsibility for belief. There is perhaps no more famous case of an individual being held responsible for his beliefs than that of Martin Luther. The German monk and professor of theology, born in 1483, sent tremors through Germany and changed both the face of western Europe and the history of Christendom by nailing his 95 *Theses* to the door of the *Schlosskirche* in Wittenburg on October 31, 1517. Luther's grievances concerning certain practices of the Catholic church attracted widespread interest and support, causing great aggravation for church officials. These and Luther's other writings led ultimately to his excommunication by Pope Leo X in 1520. The following year Luther was summoned by the Emperor Charles V to the Diet of Worms, to stand trial for his teachings. At the Diet, Luther was presented with copies of his written works and asked to publicly repudiate his teachings. After taking a day to consider his answer, Luther stood before the emperor, assorted nobles and officials of the church, and gave his famous reply:

Since your majesty and your lordships desire a simple reply, I will answer without horns and without teeth. Unless I am convinced by Scripture and plain reason – I do not accept the authority of popes and councils, for they have contradicted each other – my conscience is captive to the Word of God. I cannot and will not recant anything, for to go against conscience is neither right nor safe. Here I stand, I can do no other. God help me!
Amen.⁶

⁶ Adapted from Bainton (1995), p. 144. See also Marty (2004). It is worth noting that both Bainton and Marty indicate that there is some doubt about the authenticity of the famous portion "Here I stand, I can do no other" of Luther's speech. However, neither scholar concludes that the sentence is a later redaction of

Luther was found guilty of heresy by the Diet, and the emperor declared Luther's life forfeit, though he was granted safe passage to return to his home. On his way home, Luther's supporters arranged for his apparent abduction, placing him in hiding for a year in a nearby castle. He emerged in 1522 to continue his leadership in what ultimately became known as the Protestant Reformation, until his death in 1546.

The above account is of course an exceedingly brief history. The details of even this sliver of Luther's life are highly complex and to unpack them adequately would move us far afield of our main topic. But surely this much can be said with confidence: Martin Luther was tried, convicted, sentenced and would have been killed (had his enemies been able) in large part for the beliefs he held. And of course, though we would like to resign such proceedings to the faded memories of the 15th century, we cannot do so. One need only read the regular reports of organizations like Amnesty International or Human Rights Watch to discover that prosecution, punishment and even death for holding (or failing to hold) particular beliefs continues into the present day with alarming regularity.

I am not concerned here, however, with the political standing of the "freedom of conscience", matters of human rights or the like, important though they are. My interest in responsibility for belief is rather in its metaphysical and psychological foundations.

And while life & death examples are useful for gaining attention, one need not consider

Luther's speech. While this sentence is useful for its dramatic effect, clearly the primary significance of the case for my purposes, that Luther was being held responsible for his beliefs, is not dependent on the historicity of this sentence.

such dramatic circumstances to see the significance of this topic. As we will see repeatedly in the pages to follow, our common forms of expression indicate a conviction that we are responsible for many of our beliefs. We regularly utilize strongly normative language concerning belief. When a mother says to her teenage child, “You shouldn’t believe everything you see on TV!”,⁷ the clear implications of the statement are that the child is blameworthy for the beliefs formed on the basis of her viewing, and that the child has some measure of control over the formation of these and other similar beliefs. Our daily lives are sprinkled with such statements. Partly on this basis, I will argue that we clearly believe in both doxastic responsibility (from the Greek *doxa*, meaning “belief”) and doxastic control, at least in what philosophers term our “folk” or everyday account of our psychology.

And yet we admittedly have conflicted intuitions about our beliefs. Beliefs often feel more like things that happen to us than things that we do, things over which we have direct control. Recall Luther’s response to the injunction to repudiate his beliefs. “My conscience is captive to the Word of God”, Luther replied, and then “Here I stand, I can do no other.” This is the language of passivity or submissiveness rather than action. Clearly Luther was reflecting in these statements a commonplace view of belief as not simply subject to the dictates of the will, not something capable of being changed at a whim, and on this basis at least *prima facie* importantly different from actions.

⁷ Perhaps at present the more appropriate expression would be, “You shouldn’t believe everything you Google!”, or “You shouldn’t believe everything you read on Wikipedia!”

Reflection on intuitions and incongruities such as these has spawned a notable philosophical literature. This literature centers primarily on the issue of doxastic voluntarism, the idea that belief is directly subject to the will. The broad consensus in contemporary philosophy is against doxastic voluntarism, though there are notable exceptions.⁸ Most philosophers who reject doxastic voluntarism, however, have attempted to accommodate both of the above intuitions about belief, rather than rejecting either out of hand. Belief is thus commonly regarded as something which we are appropriately held responsible for, but not because we possess direct control over it. Advocates of this position are thus, as we shall see, often at pains to explain the meaning and grounding of our regular attributions of responsibility for beliefs.

I am convinced that these discussions have gone astray, both with respect to the conclusions that have been reached as well as the framework in which they have occurred. These discussions have frequently assumed doxastic voluntarism as the only account for control over belief formation, and they have often seemed to regard epistemology as the principal source of philosophical resources for this discussion. I will argue that we do have direct control over beliefs formation, but it is not my specific goal to defend doxastic voluntarism. It is my conviction that the first question that must be answered is a broader question of control that does not presume the will as the only relevant mechanism for shaping beliefs. And I believe that in answering this question of

⁸ For recent criticisms of doxastic voluntarism, see Williams (1973), Scott-Kakures (1994), Alston (1998), Audi (2001) and Adler (2002), all of whom I discuss in Chapters 4 and/or 5. For prominent examples of defenders of doxastic voluntarism, see Meiland (1980), Steup (2000), (2001) and Ryan (2003).

control we ought to take advantage of the resources offered by philosophy of action, which concerns itself with precisely this topic.

Throughout this essay, I will be concerned principally with the defense of a thesis that I call *moderate direct control*. Moderate direct control is the claim that the formation, retention and dissolution of beliefs are subject to direct control by normal human persons in much the same way as actions are. This thesis is broad and in certain ways ambiguous. As I characterize it, the thesis of moderate direct control unfolds into the following two corollaries:

- (1) Individuals' beliefs are responsive to reasons about what they ought to believe, in ways similar to how individual's actions are responsive to reasons about what they ought to do.
- (2) In certain circumstances, individuals have the ability to decide what to believe.

I will argue in support of this thesis in several related ways over the coming chapters. In my later chapters I will attempt to articulate a theoretical framework in support of my thesis. I begin in this chapter, however, with an intuitive or common sense argument for direct control. I claimed above that our daily lives are infused with statements that imply responsibility for and control over belief. In support of this claim, I will consider a number of individual examples, and argue that our normal ways of speaking about belief, as well as the phenomenology of belief, provide *prima facie* support for moderate direct control.

1.2. The Language & Phenomenology of Belief

I begin with a discussion of the ways in which we talk about our beliefs, and the ways that we experience the acquisition, retention and elimination of beliefs. At the outset, I wish to be clear about the modest aims of my discussion in this chapter. My contention is that a survey of some of the common ways in which we speak about belief, as well as some of our common experiences of doxastic activity, establishes a presumption in favor of direct control. We regularly speak of beliefs as subject to our direct control, and we frequently experience our beliefs as things which we embrace or reject in a direct, active manner. Given the somewhat anecdotal nature of this survey, this presumption is of course *prima facie* and defeasible. But it is a significant presumption nonetheless, one which I will argue in subsequent chapters that the opponents of doxastic voluntarism often fail to acknowledge in their attacks upon direct control. This discussion will also provide us with a context and a stock of examples for the conceptual discussion to follow in later chapters.

The first kind of linguistic practice I consider is our tendency to speak of beliefs as something for which we are responsible, something for which we are subject to praise or blame.⁹ Consider the following statements:

⁹ Some recent discussions in the philosophy of action (such as Fischer (2006), pp. 231-234) have emphasized a distinction between the necessary and sufficient conditions required for responsibility and the conditions required for the application of praise or blame. Illuminating this discussion would take us somewhat far afield of the topic of this chapter, and as I do not believe that this distinction figures prominently in my argument, I will largely ignore it for the purposes of my discussion.

1. "I used to enjoy John's company, to enjoy talking with him about things. But now he's gone and become a Republican, even a conservative. I just can't forgive him for adopting such views."
2. "You ought not think such things of Mr. Thompson, sweetheart. He is a good man and deserving of more respect from you."
3. "I'm proud to have a son of such firm Christian conviction. After the experiences and challenges he had in college, for him to have held so tightly to his faith is really admirable."
4. "It's hard for me to imagine that someone could ever be a supporter of slavery. Even those living in the 19th century U.S. shouldn't have believed that Blacks were inferior or sub-human. The idea is obviously false and just deeply immoral on the face of it."

Each of these statements attaches normative language to beliefs. Examples 1 & 3 clearly fix praise or blame to a belief or beliefs. In the first example, the individual who is the object of the statements is being excoriated for the acquisition of certain beliefs. In the latter case, the retention of the religious beliefs by the individual in question is the object of praise.¹⁰

Examples 2 & 4 make the same connection between belief and praise or blame, and also clearly imply that the holder of the belief has violated or fulfilled an obligation in acquiring or failing to acquire a certain belief. The use of "ought" and "should" in these cases indicates that the subject feels that the individual in question needs to conform to this obligation. By the dictum "ought implies can", such statements also imply the subject's belief that the individual can do so, that he or she has control over the

¹⁰ It is noteworthy that it is a common feature of the world's major religions to speak of belief in the relevant deity or the articles of the faith as meritorious, something for which one is in some way responsible in the sense of deserving praise for the belief or for holding to it. This of course shows nothing more than that the vast majority of the world's population seems to believe in some level of control over belief, but given my objectives in this chapter, even that point is quite significant.

belief or beliefs in question. I contend that such language is identical to that which is regularly used concerning actions of various types that are commonly understood to be under our direct control.¹¹ Indeed, in both examples verbs that uncontroversially indicate action can be easily and naturally substituted for the doxastic verbs currently used, without any loss of intelligibility (“You ought not *do* such things *to* Mr. Thompsen, sweetheart”, etc.).

We can multiply such examples *ad infinitum*, but this is unnecessary for our purposes. I take it to be relatively uncontroversial that we commonly use such language. The practice is nearly universal in the sense that nearly everyone speaks this way, and does so with some regularity. However, despite this fact, such language may offer only oblique support to the thesis of this chapter. It may be that such language reflects direct control over belief, but it is possible that it may instead reflect something less. Attributions of responsibility for some state of affairs do not necessarily imply direct control over it. One may possess only indirect control over a set of circumstances or even an action and yet be reasonably held responsible for it; this too is our common experience. In certain circumstances someone may hold me responsible for your actions, even though I had only a very indirect form of control over them (think of managerial responsibility in the workplace, or the responsibility of parents for their young teenage children). So perhaps all we mean by our responsibility language regarding belief is that

¹¹ As I note in my Introduction, I do not attempt to make the case for control generally, and assume that “hard determinism”, the thesis that both metaphysical incompatibilism and determinism are true, is false. For an example of hard determinism see Pereboom (2001).

we possess a kind of indirect control over it, one that may be effectuated through our other actions. This alone would not be enough to offer support for the thesis of this chapter, that our control over our beliefs is substantially similar to our control over our actions. The vast majority of us clearly think that we possess not only indirect, but also direct control over our actions.¹²

Alternatively, and of more concern for my thesis, perhaps such language reflects no conception of control at all. The use of evaluative and normative statements extends far beyond the realm of control. Normativity simply implies the presence of a norm. The failure or success of some events or states of affairs to conform to this norm may be subject to no one's control. The norm in question may, for example, be statistical or biological in nature. As Phillipa Foot notes in her book *Natural Goods*, we may call a cheetah or other non-human animal "good" even though we mean nothing more than that it exemplifies certain qualities (possessing four legs, being capable of running fast) that evaluate favorably against the biological norms (by which she means something like statistical averages) of its species.¹³ We may call another cheetah "bad" or "deficient" in virtue of the fact that it possesses only three good legs. In neither case do we mean to imply that these are things over which either animal had any control whatsoever, at least

¹² Indeed, it is not clear how one could straightforwardly possess indirect control without possessing the direct control as well, at least in the case of action.

¹³ See Foot (2003), especially pp. 25-37. Foot is concerned to argue that there are different sorts of normative statements made about non-human creatures, some of which concern our interests in these creatures, and some which concern their own flourishing (which she says is defined in terms of the species) independent of human interests. My claim is simply that these statements are typically examples of normative language having nothing to do with control.

in any sense of the word “control” as we usually apply it to human beings. Similar examples could be concocted with plant life or even geological formations, making the irrelevance of control to this language even more stark. Normative statements do not necessarily indicate or imply control, and perhaps the normative statements applied to beliefs in the examples above and our normal linguistic practices are nothing more than that which we apply to cases where control is clearly absent.

I want to concede immediately that the first concern raised above clearly blunts the impact of this specific type of normative language regarding direct control. While the possibility of indirect control does not show that we lack direct control over belief, it does make clear that more will have to be said to establish the presumption that is the goal of this chapter. However, the second possibility mentioned, that the normative statements in question have nothing to do with control, is less plausible. I will attempt to briefly explain why.

It is certainly the case that we employ normative statements in many circumstances where control is irrelevant, as noted in the examples just above. However, a distinction needs to be made between normative statements *simpliciter* and what I will call *responsibility statements*. The broad category of normative statements simply picks out all statements which imply or state the presence of a norm against which some state of affairs or event is being or may be evaluated. The category of responsibility statements, which is a subset of the former, identifies circumstances where not only is a norm present, but also in which some subject is being held responsible for their failure or

success in conforming to that norm. Normative statements may or may not implicate control, but responsibility statements essentially imply control on the part of the subject of the statement. Examples concerning animals, plants and inanimate objects typically fall under the broader category of normative language. But the examples 1-4 initially laid out above are, I allege, examples of the use of responsibility language, closely connected with control. This seems clear enough from close examination of the cases. Examples 1 & 3 use obviously active language concerning the acquisition and/or retention of beliefs. The statement in case 2 takes the form of a command or entreaty which the speaker clearly hopes to have some effect on the hearer (hence the offering of reasons in the second sentence). Example 4 is perhaps the most ambiguous, but, like the others, it clearly involves language that implies responsibility on some person or persons for their beliefs. In this particular case the responsibility charged is obviously moral responsibility, for which some form of control is a prerequisite.

The distinctive character of such responsibility statements about belief is also seen in their limited scope of application. I spoke just above about the universality of such language (its usage by nearly everyone), but in another important way it is not universal at all. We clearly apply such statements only to a limited subset of our beliefs. All of the examples above involve belief characteristically formed by some kind of deliberative process. Deliberatively formed beliefs (and analogous instances of action) offer paradigmatic examples of control. Our perception that many such beliefs are in fact up to us as believers, I suggest, helps to explain the use of responsibility statements with

respect to them. Contrast this with beliefs formed directly through perceptual experience. Imagine that you and I are standing outside on a brilliantly sunny day. You remark on how wonderfully sunny it is. I turn from looking straight up into to the blue sky, look at you incredulously, and blurt out, “Sunny? What are you talking about? This is a cloudy day if ever I saw one!” Now, insofar as I was sincere, and you recognized my sincerity (which admittedly stretches the imagination a bit), it is unlikely that you would respond with some kind of responsibility statement like those we’ve considered. You would not say, “You ought not believe such things,” or “How foolish of you to think so. Consider the evidence.” Instead, you would likely question the functioning of my visual system (“is there something wrong with your eyes?”), or my sanity, (“are you feeling OK? Do you want to sit down for a bit?”). This difference in approach reflects the fact that we don’t typically see such perceptual beliefs as subject to our control, and thus don’t generally see them as matters of responsibility. This is contrasted with the responsibility statements in our original cases. The active, control-implicating character of responsibility statements, as well as the recognition of their limited scope of application, provides us with sufficient resources to distinguish such statements from “mere” normative language, and thus to reject the claim that they imply nothing at all regarding control.

We have established, then, the widespread belief in at least indirect control over belief, reflected in our standard patterns of speech. The question, then, is whether there is any language which offers more (forgive the pun) direct support for the thesis of

moderate direct control. Happily, there is. Along with our common practice of holding ourselves and others responsible for our beliefs is the less common but still widespread practice of describing our beliefs as subject to our direct control. Consider the following examples:

5. "I've decided to believe him," John said.
6. "I understand your position, Lieutenant," replied Mrs. Jones, "but until you give me good reason to think otherwise, I refuse to believe that my husband is dead."
7. "I reviewed the evidence, and in the end I simply decided that the theory was not plausible."
8. "The treasure is a myth", Ben's father concluded eagerly. "I refuse to believe that," Ben replied. "You can believe what you want", his father retorted with frustration, "you're a grown person!"¹⁴

As these and many other common examples show, the language of decision and choice is not foreign to doxastic discussions; to the contrary, it is part and parcel of our normal ways of speaking about belief.

The above are examples of what I call *control statements*. Control statements offer the most straightforward and substantial support to the presumption of direct control over belief arising from our regular linguistic practices. While normative statements imply only the presence of a norm, and responsibility statements require the existence of only indirect control (though they still may imply direct control), control statements clearly indicate a conviction of the existence of direct control over one or more beliefs.

To say that I decided to believe someone is, intuitively, to say that I deliberated and made

¹⁴ This example is taken from the 2004 movie *National Treasure*. Interestingly, even though the Ben's father is dismissive about the refusal to believe that Ben articulates, his response nevertheless implies that Ben has the ability to believe something voluntarily.

a decision, and belief was direct result of this decision. The focus on the decision to believe (as opposed to the more compact locution, “I believed him”) often implies the relative indeterminacy of the evidence related to the belief, and the role of the decision in directly acquiring the belief. To say, as Ms. Jones does above, that one refuses to believe something is to indicate that one feels some significant inclination towards a certain proposition, but also that one actively, directly and successfully resists the attraction of that belief. The third example similarly indicates the role of choice in directly determining belief states, in this case the rejection of a particular proposition or the elimination of a particular belief. The fourth example says at least as much as the others.

This type of speech is, as noted earlier, less common than the simple normative language of responsibility statements, but what it loses in ubiquity it gains in inferential significance. It is far more challenging to explain away such language, to explain how it does not indicate at least a widespread embrace of direct control over belief. Some opponents of doxastic voluntarism nevertheless attempt to offer such explanations, though others seem to feel no such burden.¹⁵ We will consider such attempts in Chapters 4 and 5, but here it is worth noting that the simple fact that the critics of doxastic voluntarism take themselves to need to offer alternative explanations for such language shows the intuitive support that it offers to the idea of direct control.¹⁶ This is precisely

¹⁵ Audi (2001) seems to keenly feel this dialectical challenge, while Alston (1988) though aware of it, is more dismissive of its significance. Steup (2000), p.26 also notes this form of argument on the part of the critic.

¹⁶ In reality, the same is true of simple deontological or responsibility language about belief. Critics of the deontological conception typically argue in the following way: We lack responsibility for our beliefs, or the

the presumption which I have been arguing for. The common sense interpretation of such statements is that they point to our direct control over at least some beliefs, hence the need for the involuntarist to provide an alternative explanation of their significance. Such statements, normative, responsibility, and control, are data for the philosophical discussion of control over belief. They help fix the starting point for such a discussion, and establish the need for skeptic concerning direct control to make it clear why we ought not think that we possess such control over belief, just as we do over action.

I have argued that our patterns of speech illuminate the widespread belief that we possess direct control over at least some of our beliefs, and thus establish a dialectical presumption in favor of direct control. A second factor in support of such a presumption comes from our phenomenal experience. Just as we regularly speak of belief as subject to direct control, we also regularly have experiences of beliefs as subject to direct control in a manner very similar to the way we experience certain voluntary actions. This type of experience may even be the foundation of the patterns of speech noted above, though I will not attempt to establish that claim here.

The circumstances and ways in which such control is exerted and experienced are diverse. Again, I will focus principally on circumstances of deliberation, in part because these provide the clearest examples of control. One common experience of direct control occurs in instances of belief retention or elimination, when a subject is confronted with

deontological conception of justification is false, *because* we lack control over our beliefs. This form of argument will become readily apparent in Chapters 3 & 4.

propositions that either threaten an existing belief or threaten to initiate a belief that is for some reason unwelcome. Consider the following examples, the second of which will figure prominently in the next chapter:

9. *Suspicious Spouse* – Gerard is a husband who, despite his best intentions, finds himself regularly doubting his wife’s fidelity and worrying that totally innocent actions (trips to the grocery store, etc.) indicate unfaithfulness. He believes that his wife is faithful and would affirm so honestly to anyone who asks. There is no evidence of such infidelity that an objective third party would find suspicious, and he is aware of this fact, having discussed it with friends. Yet he still struggles with such doubts. When he finds himself in such patterns of thought, he actively resists the thoughts occurring to him, often telling himself things like, “I know that isn’t true”, and truncating his undesired deliberations with an internal affirmation of his wife’s love. In such circumstances, he has the experience of directly and actively retaining his beliefs, and of these beliefs being subject (though with some difficulty) to his direct control.¹⁷
10. *Black Widow*: John has a significant fear of spiders, owing to a frightening childhood incident. One evening, as he is climbing into bed, he notices an old cobweb in one corner of the room, near the ceiling. John immediately and non-reflectively forms the belief that there is a highly venomous black widow in the room. His heart rate accelerates, his eyes dart nervously about the room, he reaches for a hefty philosophy tome to smash any spider that he sees. Then John stops and reflects, asking himself if the old cobweb is really a good reason to believe that there is a black widow in the room. He judges that in fact it is not a good reason at all, and that his initial reaction was in reality quite silly. Though his fear doesn’t entirely dissipate immediately, he ceases believing that there is such a spider in the room. He rolls over in bed and opens the book he is holding, deciding to pursue his preferred method of inducing sleep by reading long passages of Kant’s *Groundwork*.

¹⁷ This example is not entirely unlike the example advanced by Meiland (1980), p. 16. One difference is that in Meiland’s example the subject exercises control over belief against the preponderance of the evidence (though in pursuit of what Meiland sees as a noble goal), whereas in my example the subject attempts to exercise control in favor of the evidence, but against strong non-rational factors influencing his belief. This is consistent with my conviction that direct control over belief can be a positive force in the regulation of our cognitive lives.

The form of Gerard's and John's deliberations (though hopefully not the content) is very familiar. We've all had the experience of considering or deliberating about some proposition, feeling the tug of certain evidence in its favor, but then expressing (and experiencing) an internal "No! I won't/can't/don't/shouldn't believe that." The expression sometimes even spills over from our private internal speech into audible statements. I suggest that this experience of deliberation and the willful resolution of the deliberation closely mirrors the resolution of many of our practical deliberations. It is I who resolves the dispute, makes the decision, and determines my belief. The experience is of the belief being up to me, just as I experience many of my decisions pertaining to action as up to me. And such experiences offer substantial intuitive support to the idea that belief is under our direct control.

We also may exert and experience direct control over belief in contexts of belief formation, rather than belief rejection or elimination. In such cases, we may form our beliefs in the context of the experience which they concern, or via deliberation after the fact, as the next two examples indicate:

11. *Failing Friend*: Luis asks a favor of Kendra, and then is dismayed to find that Kendra fails to fulfill his request as she had agreed to. When confronted she tells him a long and somewhat dubious story, but relates it with great sincerity. Luis mulls over Kendra's story, weighing its implausibility against their good friendship and her history of honesty. He finally decides that he believes her, and indicates that he accepts her apology. Luis' experience is of his own decision terminating his deliberation and initiating his belief.
12. *Rude Interlocutor*: Sarah is at a party, chatting with an acquaintance named Kathy. During the course of the conversation Sarah, while not reflecting on them, nevertheless registers an awareness of many things

about Kathy's physical mannerisms, which guide Sarah in her interaction with this Kathy. Later that night at home, Sarah reflects on her conversation, and in particular about several minor cues from Kathy, which she had not previously reflected upon or formed any beliefs about. She considers certain slightly exaggerated gestures, a laugh that upon reflection seemed somewhat forced, an eyebrow raised at one of Sarah's comments, a brief downward glance Kathy made that was probably at her watch. Through her deliberation, Sarah judges that these considerations together are good reason to believe that Kathy was actually feigning friendliness and being rude to her. In forming this judgment, Sarah comes to have the belief that Kathy was being rude.

These, too, are common phenomena. We regularly encounter circumstances of theoretical deliberation that are complex and whose outcome is not altogether or immediately clear based on the evidence, whether they occur during the initial experience or involve retrospective deliberation after the fact. Yet we regularly terminate such deliberations with a decision one way or another, and these decisions directly determine whether or not we believe in the propositions in question. The reasons for which we terminate our deliberations, whether practical or evidential, are relatively immaterial at this point in our discussion. What is important is that we experience these beliefs as determined by decisions that are ours, are up to us, and made by us. These experiences contribute to the conviction that belief formation, modification, retention and elimination are subject to our control.

One objection commonly raised in discussions of doxastic voluntarism is germane to our discussion here. A key difference between belief and action, it is sometimes alleged, concerns the kind of introspective awareness I necessarily possess when making a practical decision. The claim is that in cases of practical decision-making I experience a

moment of decision with a particular phenomenal quality, where I resolve the deliberation in favor of one of the available alternatives. Critics claim that such an introspective awareness is a necessary condition of direct control, but is lacking in our experiences of belief. This shows that we lack direct control over belief.

I wish to respond to this objection in two ways. First, in light of the examples just above, I clearly disagree with the claim that instances of belief formation, retention and dissolution lack these phenomenal qualities. It seems to me quite obvious that we frequently experience (and make) decisions that directly determine our beliefs, and moreover that these experiences lack no relevant phenomenal quality possessed by similar acts of practical decision. Arguing fully for this latter point would require much more space than I have available here, but I take the cases above, and the many others that could be produced from our experience, to offer significant evidence in its favor.

But while I do think that we often possess the relevant introspective awareness with respect to belief formation/retention/dissolution, I also follow Matthias Steup, who has argued convincingly that such awareness of decision is an inappropriate condition on both theoretical and practical decision-making.¹⁸ While such introspective awareness is often present, it is also not necessary for control over belief or action. Steup's point is intuitive; just as we often fail to experience an awareness of doxastic decisions (the formation or elimination of a belief), so too do we often fail to experience an awareness of practical decisions we've made. The employee who blows up at her boss, leading to

¹⁸ Steup (2000)

her firing, may only realize retrospectively that she had decided that she no longer wanted or planned to retain her job, precipitating just such an incident. More starkly, I may only realize when I've opened the fridge that I wandered out into the kitchen precisely to get another piece of the chocolate cake found there. Similarly, a jilted lover may drive around town blindly, lost in his sorrow, only to discover as he arrives that he had been aiming all along at his and his love's favorite overlook. And perhaps most simply and commonly, I make a thousand small decisions every hour, without any present introspective awareness. I voluntarily open the car door in the process of loading up my screaming children, without any introspective awareness of a decision to do so. Practical decisions like these which are clearly under our control do not always present themselves for consideration before the seat of consciousness. They often pass by in a stream of events of which we have no present awareness to speak of.

As Steup argues, a more reasonable condition upon practical decision is the availability of a retrospective awareness of my decision to perform a particular action. Here it should be understood what this awareness amounts to, however. When I think back (if I have occasion to do so) upon my opening of the door, I discover no period of deliberation and moment of conscious decision, no internal discussion resolved by internal speech of the form "I will do Φ ." Instead, I recognize that I decided to open the door because I am able to reflect and see that it was I who performed the action, perhaps in the course of a larger plan of action (getting the kids home and to bed), without any undue external influence or manipulation. It seems to me that I could have chosen not to

open the door, had I wished or had reason to do so. The retrospective awareness available to us is thus often not of an intellectual process of deliberation or decision.¹⁹ But if such clear mental acts of decision are not a condition of control over action, they are similarly not a necessary condition of voluntary belief formation, retention, or elimination. I need not, simultaneous with my formation of a belief, have an “it is I who now am forming this belief” experience of my belief formation. Rather, I only need possess a similar available retrospective awareness of the fact that it is I that has come to a belief, a condition which we frequently can satisfy. Thus this objection fails to establish a meaningful difference between action and belief with regards to our phenomenal experience of decision.

At this point, I take myself to have established that we speak as if some beliefs are under our direct control, and we experience the formation and dissolution of beliefs as frequently under our direct control. But what exactly does this prove? I surely do not wish to argue that in virtue of these considerations we ought to simply close our inquiry and conclude that we possess such control. Rather, as I noted at the beginning of this chapter, I simply claim that such language and experience establishes a presumption in favor of direct control. In virtue of these widespread and regular phenomena, it is incumbent upon the opponent of doxastic voluntarism to explain why it is that we lack such control. The burden of proof with respect to direct control over belief is on the critic.

¹⁹ It may be that such a process occurred, the point is that I frequently don't have access to it in my retrospective awareness.

Of course, many criticisms of precisely this nature have been made; I will spend the remainder of this work attempting to buttress the presumption in favor of direct control with further argument, as well as responding to specific criticisms aimed at direct control.

My claim in this chapter may in one sense be understood as a point about philosophical method. I am suggesting a philosophical analogue of the common sense principle, “If it waddles like a duck, and it quacks like a duck, then it is a duck.” If there exists a widespread and regular human phenomenon with a *prima facie* straightforward and intelligible explanation, that this explanation ought to be accorded a certain deference. This presumption is of course *prima facie*, and pre-theoretical. It may be that the allegedly straightforward explanation is in the end unintelligible, as some philosophers claim in the case of free will and direct control over actions. My claim is simply that in such circumstances the burden of proof is upon the person who would make such a charge of unintelligibility.

In connection with this point, we might ask ourselves why after all we believe that we possess direct control over and are responsible for our actions. Certainly some philosophers hold this position in virtue of their rehearsal of the relevant theoretical arguments in favor of libertarianism or compatibilism. But nearly everyone else believes so because of the way in which they and everyone else speak about actions and their regular experiences of their actions as seeming to be under their control. The philosophical debate over free will and control with respect to action has become immensely complex; some minimal part of that complexity will, I trust, be replicated in

the upcoming chapters. But surely nearly all of the interlocutors in that debate agree that it starts with a presumption in favor of control. I would argue that this presumption is based on our experiences of action as well as our universal practices of attributing responsibility, praising and blaming each other for our actions. The debate may move forward by attempting to unpack and explain the control alleged, or by the advancement and consideration of arguments against such control, but in either case it starts with the presumption of control. I am suggesting that some approximation of the same presumption is present with respect to belief, based on the considerations raised above.

So, I claim, common practices and common experience tell us that we do possess control over our beliefs. But philosophical reflection cannot stop with common sense; it should rather begin with it and see if the intuitions found there possess any plausible theoretical support. In the following chapters I will engage in such an inquiry. In the next chapter I will critically examine and depart from the assumption, noted at the outset of this chapter, that doxastic voluntarism is the only model for direct control over belief, and provide an account of control over belief that does not focus on the activity of the will.

CHAPTER 2 – MODERATE DIRECT CONTROL OVER BELIEF ON INTELLECTUALISM

2.1. Setting Aside the Will

In Chapter 1 I laid out an argument supporting a presumption in favor of direct control over belief based on our common and intuitive experiences and practices. As I noted previously, however, what we call common sense must be capable of surviving in the light of reflective scrutiny if this presumption is to be vindicated. In this chapter I begin the task that will occupy the rest of this work, that of examining the theoretical grounding for direct control over belief. In the course of this pursuit I will make and examine distinctions concerning both the nature and necessary conditions of control. Here we will see the first of these, the distinction between intellectualist and voluntarist conceptions of control.

As I noted in the Introduction, one of the cardinal assumptions of much of the recent discussion of control over belief has been that this discussion is best conducted in terms of *doxastic voluntarism*, the idea that belief is under one's direct control by being subject to the dictates of the will. In some ways, this assumption may fall out of the considerations we examined in the last chapter. As I observed, the language we use about belief is often strongly analogous with the ways that we talk about action. Many philosophers concerned with action conceptualize this control in terms of the will, though there remains much debate over precisely how both the will and this control ought to be characterized.

In this chapter, however, I wish to question the assumption of doxastic voluntarism as a model of control, and examine the potential implications for control over belief. Not all philosophers accept that the will is essential for control, either for belief or for action. Alternate models of control exist, including the model I will consider here, sometimes known as intellectualism.²⁰ The intellectualist focuses on judgment and argues that control consists in conformity to it. One possesses control just in case one's action or belief follows one's judgment about what one ought to do or believe. The voluntarist, on the other hand, places the locus of control in the will, the mechanism of intention formation, and posits no necessary connection between the will and judgment. One has control just in case the resulting state of belief or intention formation is responsive to one's willing or decision.

In what follows, I will articulate an intellectualist conception of control over belief, and defend it from certain objections, primarily from within the intellectualist framework. As will be my practice throughout this work, I will adopt the intellectualist conception for the purpose of showing its commitment to direct control over belief, rather than to defend it as the correct or singular account of control. I will begin by offering a more extended sketch of the distinction between intellectualism and voluntarism, leading to my description of an intellectualist picture of control that I will argue accords with our

²⁰ See Owens (2000), pp. 77-87. Owens appears to get the distinction from Kent (1995). For a closely related distinction concerning agency and control over belief and action, see Watson (2004). Watson employs different terminology to describe the distinction, as he notes, his "internalist" is roughly the intellectualist, and his "externalist" is roughly the voluntarist. Another similar distinction is found in Smith (2005). Smith uses the terms "rationalist" and "volitionalist", respectively.

analyses of various common instances of belief. I will then contrast my view with that of David Owens, a prominent intellectualist who holds that we lack direct control over our beliefs. I will argue that Owens is mistaken in this claim, and that in his argument he equivocates on central notions of control and rationality. Examining Owens' argument will additionally be helpful to us in distinguishing precisely what intellectualist control consists in, for both belief and action.

2.2. Intellectualism & Voluntarism

The intellectualist/voluntarist distinction goes by a variety of names, and though a minority position in the contemporary field, it has a lengthy history within philosophy. These particular terms come from Bonnie Kent's *Virtues of the Will*,²¹ where she applies them to the debate over freedom and responsibility that occurred between a number of late 13th-century thinkers operating in and around the University of Paris. Kent uses the term "intellectualist" to describe the Thomistic/Aristotelean conception of action, focused on the intellect, judgment and the essential rationality of free action, which was held by the followers of this school of thought. "Voluntarist" is the name she gives to their interlocutors, such as Duns Scotus, who doubted the adequacy of the intellect & judgment in explaining human responsibility, and focused instead on the power & liberty of the will. Owens follows Kent in using this terminology in his *Reason Without*

²¹ Kent (1995).

Freedom.²² As my own thought on these matters is significantly influenced by Owens, I utilize his terminology, though other philosophers have given the positions other names. Gary Watson uses the term “internalist” to describe the basic intellectualist position, and “externalist” for the voluntarist.²³ Angela Smith uses “rationalist” and “volitionalist” for the same distinction.²⁴ Despite the varying terminology, however, all of these philosophers are concerned with the same basic issues and the same fundamental distinction about the roles of judgment and the will in exercising control.

In examining the distinction between intellectualism and voluntarism, it is helpful to initially focus upon actions.²⁵ Consider an uncontroversial case of free action. Imagine that Eric is hungry. He stops working on his lecture notes and reflects upon how he might satisfy his hunger. He considers, amongst other things, going to the campus commons, or visiting the sandwich shop across the street from his office. The commons is less expensive and tends to have good food. However, the sandwich shop is much closer, Eric must teach in 25 minutes, and he has not finished prepping for class. Eric weighs the various considerations, and comes to the judgment that he ought to go to the sandwich

²² Owens (2000).

²³ Watson (2004).

²⁴ Smith (2005).

²⁵ Since I will be only offering a sketch of the two views to set up the discussion of intellectualism and belief, certain subtleties will be glossed over or ignored, and I will draw the distinction between the two views in a somewhat stark & simplistic fashion. For further resources on the distinction, see the references above in n. 20.

shop across the street. He gets up from his desk, leaves his office, and goes to the sandwich shop, ordering turkey on rye.²⁶

It seems plain that if any action is free, Eric's was. I take this to mean (at least) that he was in control of his action. Moreover, the basic account of what is involved in his action is relatively uncontroversial. Eric reflected on various reasons for and against different courses of action, formed a judgment about what was best, this judgment led to the formation of an intention to Φ , and his action followed from his intention.²⁷ There are, however, at least two differing accounts of what Eric's control over his action consisted in, or, to put it another way, where precisely in Eric's various faculties we ought to locate his control over this action. As noted above, the intellectualist focuses on practical judgment and argues that control consists in conformity to it. For the intellectualist, Eric's action exhibited control and thus freedom because the formation of his intention followed the dictates of his practical judgment that going to the sandwich shop was best.²⁸ The voluntarist, on the other hand, locates control in the faculty of the

²⁶ I thank Peter Graham for pointing out the need to further clarify the distinction between the voluntarist and intellectualist positions.

²⁷ Of course, as we started to press for greater details about the various part of this action, significant disagreement would likely emerge between different philosophers. But hopefully the broad outlines of the account can garner significant agreement.

²⁸ To be precise, for the intellectualist, Eric's action must be in conformity with his judgment, *and* it must flow causally in the proper way from his judgment. Without this latter qualifier, the intellectualist account would be open to analogues of the worries about deviant causal chains raised concerning various accounts of intentional action, such as Davidson's famous mountain climber example (Davidson (1973)). My thanks to Dion Scott-Kakures for pointing out this worry. I will not attempt here to specify precisely what the causal connection must be, that is beyond the scope of my project.

will and its free decision. According to the voluntarist Eric controlled his action because it followed from his formation of an intention to go to the sandwich shop. The connection between Eric's judgment of what he ought to do and the subsequent intention formation is comparatively immaterial on the voluntarist account of control.

Initially, it's possible for this to seem like a distinction without much of a difference. The difference between the two accounts becomes clearer if we imagine a case of *akrasia*, where Eric's judgment of what he ought to do and his intention formation come apart. Consider a case where Eric judges that he ought to go to the sandwich shop, that this would be best. But in this case Eric forms a different intention, rises from his desk, and heads to the commons for Chinese food. Here there is discord between Eric's practical judgment on the one hand, and his intention formation and action on the other. The intellectualist and voluntarist disagree about the implications of this discord. The intellectualist argues for a tight connection between the will and practical judgment in free action. She claims that in this latter case Eric lacks control over his action, because it fails to conform to his judgment of what is best for him to do. For the intellectualist, this incoherence between practical judgment and the will is damaging to Eric's agency. She is concerned principally with the attributability of free actions in the sense of their coherence with the agent's normative commitments.

The voluntarist disagrees about the assessment of Eric's action. She conceives of no tight connection between the will and practical judgment, in part because it is not in this connection that control consists. It may be the case that we typically form intentions

in line with our judgments of what it best, but this is not necessary for controlling one's actions. The voluntarist conceives of the attributability of free actions in terms of whether an action is up to the agent, meaning connected appropriately with the free decision of the will. For the voluntarist, the counter-normative power²⁹ of the will exhibited in this case exemplifies agency. Eric is in control of his action, because in his action he successfully carries out his intention, which results from his free choice. The disconnection between Eric's judgment and his intention does not undermine agency, it rather shows the distinctive role of the will in voluntary action.³⁰

As Watson notes, such cases of *akrasia* bring out the motivation for both intellectualism and voluntarism.³¹ Our intuitions in such cases pull us in two directions. We recognize in these actions certain characteristic features of agency, such as the intentional nature of the action, that push us to ascribe responsibility in accordance with the voluntarist analysis. If Eric were late to class because of his decision, we would not simply excuse him because of the failure of his action to conform to his own judgment. No one forced him to go to the commons; he went there on the basis of his own decision.

²⁹ I take this term from Watson (2004), p. 131

³⁰ At this point the distinction between the intellectualist and the voluntarist may begin to sound similar to the distinction between the compatibilist and libertarian incompatibilist. The two distinctions are certainly closely related, but they are nevertheless different. The former is a distinction about the proper characterization of freedom within the agent, while the latter is of course a distinction about the necessary metaphysical conditions of freedom and responsibility. Certain pairings of the views tend to predominate, but not exclusively. Audi is an example of a compatibilist voluntarist (see Audi (2001)).

³¹ See Watson (2004), pp.129-134. Watson also notes that cases of *akrasia* help to show us why neither pure intellectualism nor pure voluntarism are adequate views.

On the other hand, we call such actions “weak-willed”, reflecting concerns about the disagreement between practical judgment and the will. As Watson says, it is not at all clear on the purely voluntarist account how this label applies. And yet, despite the fact the choice is made by the agent, something seems to have gone wrong. We recognize in such actions an incoherence in the actor that seems to mitigate against “full-blooded” agency. This latter instinct aligns with the intellectualist account.

This brief sketch illustrates that there are *prima facie* reasons for both the voluntarist and intellectualist conceptions of belief. The motivations for and virtues of intellectualism and voluntarism are complex and worthy of extended discussion and inquiry. But this is not my project here. My goal is not to adjudicate between the two accounts, but rather to introduce the distinction and to highlight certain characteristics of the views, in preparation for a discussion of an intellectualist conception of belief. This is what I turn to next.

2.3. Intellectualism and Belief

Briefly clarifying the intellectualist and voluntarist views in this manner helps us to infer the central features of an intellectualist account of control over belief. To bring out these features, I will focus on three examples of doxastic control, two of which were raised in the previous chapter (for convenience’s sake, I briefly restate the two examples here, adding the third just below):

10. *Black Widow*: John has a significant fear of spiders, owing to a frightening childhood incident. One evening, as he is climbing into bed, he notices an old cobweb in one corner of the room, near the ceiling. John immediately and non-reflectively forms the belief that there is a highly venomous black widow in the room. His heart rate accelerates, his eyes dart nervously about the room, he reaches for a hefty philosophy tome to smash any spider that he sees. Then John stops and reflects, asking himself if the old cobweb is really a good reason to believe that there is a black widow in the room. He judges that in fact it is not a good reason at all, and that his initial reaction was in reality quite silly. Though his fear doesn't entirely dissipate immediately, he ceases believing that there is such a spider in the room. He rolls over in bed and opens the book he is holding, deciding to pursue his preferred method of inducing sleep by reading long passages of Kant's *Groundwork*.
12. *Rude Interlocutor*: Sarah is at a party, chatting with an acquaintance named Kathy. During the course of the conversation Sarah, while not reflecting on them, nevertheless registers an awareness of many things about Kathy's physical mannerisms, which help to guide Sarah in her interaction with Kathy. Later that night at home, Sarah reflects on her conversation, and in particular about several minor cues from Kathy, which she had not previously reflected upon or formed any beliefs about. She considers certain slightly exaggerated gestures, a laugh that upon reflection seemed somewhat forced, an eyebrow raised at one of Sarah's comments, a brief glance downward Kathy made that was probably at her watch. Through her deliberation, Sarah judges that these considerations together are good reason to believe that Kathy was actually feigning friendliness and being rude to her. In forming this judgment, Sarah comes to have the belief that Kathy was being rude.
13. *Desperate Househusband*: Ned is recently married, and he and his wife are expecting the birth of their first child. Unemployed, Ned longs to give his family the kind of financial security and lifestyle that he sees many of his college friends enjoying. While listening to the radio in his car one day, he is intrigued to hear of the opportunity to purchase materials that can train anyone to be successful buying and selling homes for no money down. Ned is feeling particularly desperate on this day and with hopeful excitement he calls the number offered. He listens to a compelling five minute presentation of the basic system, followed by a series of testimonials by users of the system. Buoyed by the passionate and authentic-sounding testimonials, Ned begins to believe that this program is their opportunity to become financially successful and independent. He heads home to talk with his wife, Krista. He excitedly explains what he

has learned, and communicates his belief that this system will lead them out of their current financial difficulties. Krista, a former philosophy major, arches one eyebrow and simply responds, “Are you sure you’re being critical enough about what you’ve heard? A radio infomercial? You’re sounding awfully gullible to me.” Ned is taken aback. He reflects on his manner of thinking, and after a short time concludes that Krista is right. Believing such things and getting so excited on the basis of one marketing phone call is not reasonable, particularly on matters of such importance. He concludes that he needs to be more skeptical and not allow himself to be sucked in so easily. He vows to be more critical of such offers in the future. In forming this negative judgment of his evidential standards, he dismisses his earlier belief and irrational excitement. Krista helpfully reminds him that he also ought to stop buying lottery tickets.

Each of these examples highlights features essential to an intellectualist account of control over belief. As I’ve repeatedly noted judgment is central to any intellectualist account, as are reflection & deliberation, from which judgment typically arises. Control for the intellectualist is reflective control.³² Paradigmatic cases of control over action on an intellectualist account, such as the example in the previous section, involve deliberation culminating in a practical judgment that guides the formation of an intention. In the case of belief, theoretical judgment, judgment about what I ought to believe, stands in for practical judgment, and belief replaces intention as the state resulting from judgment.

All of the examples above share the same basic structure, analogous to the intellectualist account of action. In each case a subject has an initial cognitive state (belief or lack of belief) relative to a specific proposition. The subject engages in reflection, and this reflection culminates in a normative judgment about whether or not the subject ought

³² See Owens (2000), pp. 9-22.

to believe the proposition in question. This judgment, in the three examples above, results in a different cognitive state relative to the proposition than the subject possessed at the beginning of reflection. Thus, reflective judgment is the subject's mechanism of control over the relevant belief, the mechanism by which he or she comes to believe or disbelieve in each case.

Reflection and judgment not only serve as the causal genesis of the belief, they also establish the subject's normative ownership over the belief. The control exercised through reflective judgment, the intellectualist contends, makes a belief properly attributable to the subject. Though, as I've noted, Owens ultimately rejects reflective control, he offers a nice articulation of this point. When we believe on the basis of reflective judgment, Owens says, "Then belief is no longer being determined solely by the impression the world makes on us: our view of the quality of our reasons, our judgment of the normative significance of our psychological situation, is what commands conviction."³³ The connection of a belief to one's normative commitments through one's reflective judgment makes it reasonable to hold the subject responsible for the belief.

Of course, many beliefs are not formed with the level of reflection found in these examples. As Kent Bach notes, much of our practical and theoretical reasoning is "default", utilizing rules of thumb, shortcuts and other mechanisms to economize our scarce cognitive resources as we navigate the world forming beliefs and acting.³⁴ In the

³³ Ibid., p. 12.

³⁴ Bach (1984).

doxastic sphere, we may frequently operate under the guidance of ideas such as Bach's

"Take for Granted" (TFG) rule:

TFG: If it seems to me that p , infer that p , provided no reason to the contrary appears to me.

By contrast, the conception of belief formation, modification and retention that I focus on throughout this work involves explicit reflection and deliberation. Such a conception may seem somewhat idealized. While I've already discussed the fact that my focus on deliberately formed beliefs follows the pattern found in relevant literature on the topic, two further responses are worth mentioning. First, while the picture of belief I focus on is perhaps somewhat idealized to allow for clearer discussion and focus on the different elements involved in belief formation, it is not unrealistic. The types of situations and responses noted in the examples above and in Chapter 1 are ones we have all experienced (thus their intuitive quality), though the details of the examples and specific beliefs may themselves be unique. All of us have jumped to conclusions and through reflection come to realize that the belief is unmerited, and all of us have had new insights dawn on us when thinking about previous experiences. While we may typically economize our deliberation out of necessity, circumstances sometimes demand or gain our full attention, and we frequently engage in reflection and decision about what to believe or what to do as explicit as any found in my examples.

Secondly, even those cases where our reasoning is default are not problematic for the general argument in favor of moderate direct control. Recall that my claim is that the

control we enjoy over belief is strongly similar to that which we enjoy over action.

Bach's theory of default reasoning is general, applying equally to practical and theoretical deliberation. Corresponding to his rule TFG is the practical rule "Not Worth Considering" (NWC):

NWC: If it occurs to me to do A, do A unless there occurs to me the thought of a reason to the contrary or of an alternative to A.

NWC parallels TFG, and Bach clearly sees strong parallels between theoretical and practical reasoning. But unless the implication is that default reasoning undermines control over action, it is not clear why default reasoning should be a problem for control over belief formation. For his part, Bach makes no implication that control over action is threatened by default reasoning. Control is not his principal concern, but the framework of reasoning that Bach describes for both action and belief is strongly normative and implies (if not explicitly supports) the possibility of control.³⁵ So the mere fact of default reasoning patterns is not enough to motivate a concern for control over belief formation, modification and retention, some further difference between the practical and theoretical cases is required.

³⁵ See Bach (1984), pp. 37-38. One implication of Bach's ideas may be that our control over both our beliefs and actions is largely negative, involving suspending/terminating certain actions or belief forming processes in light of conflicting alternatives or reasons. Bach makes no such claim, but we may see it implied by the passivity of some of his language about reasoning & decision. Of course language of passivity and/or mechanism and negative conceptions of control are not uncommon in discussions of action, and, as I discuss in this chapter and the next, pose no threat to my argument.

Beyond cases of limited or default reflection, however, many beliefs come to us without any reflection at all. For example, in the Black Widow (BW) case above the subject comes to his initial belief without engaging in reflection (though subsequent reflection led to its modification). As Owens notes, any plausible intellectualist view must include a counterfactual element to account for this fact of our doxastic lives. To be credible any intellectualist position must hold that not only when beliefs *are* the product of reflective control, but also when beliefs *would be open* to reflective control, can we be properly held responsible for them.³⁶

This is not to imply that the level of responsibility for beliefs that involve no reflection is the same as those for which reflection is explicit and extended. Certainly our ascriptions of responsibility will be more common in the latter case. Our judgments of individuals and their responsibility for their beliefs seem to track the intellectualist account here. Consider two subjects who both hold the same unjustified belief. The first has acquired the belief through largely non-reflective processes, perhaps by simply passively imbibing the beliefs of her wider community. The other has reflected on the evidence in question and consciously come to the same conclusion. We would tend to hold the former individual less responsible for such a belief than the latter. The latter person seems the more appropriate target of judgments of fault for his belief, while the former seems more appropriately the object of sympathy or related attitudes. The differences in judgment in the two cases will in part be determined by how much

³⁶ See Owens (2000), pp. 13-14

reasonable opportunity for reflection and change seemed to be available to the former individual.

The examples above share not only a similar structure, they also share the common feature of illustrating the important role of control over belief determination in our intellectual lives. In each example, the subject initially fails to connect up appropriately with the facts of the matter. It is only through reflection and the control afforded by it that the subject comes to an epistemically correct position. As noted, we form beliefs and fail to form beliefs all the time without conscious reflection. But given the fallibility of our cognitive faculties, it is often the case that we do so mistakenly, and it is one of the critical functions of reflection to allow us to correct these errors. In doing so, we exercise reflective control over our beliefs. Of course, reflective control is no guarantee of good judgment, in theoretical or practical reasoning. But it nevertheless offers us an important means of correcting our frequent cognitive errors.

Though the examples described above share a common basic structure, I contend that each highlights a different form or modality of reflective control over belief which we possess on the intellectualist conception. The first case is an example of what I will call *negative control*, the second a case of *positive control*, and the last a case of *pragmatic or normative control*. I will briefly discuss each in turn.

Consider first Black Widow (BW). In cases of negative control, reflection functions not to motivate one towards belief, but rather to prevent or eliminate belief. In BW John's formation of the initial belief *p*, that there is a black widow in the room, is wildly

irrational, though perhaps through no fault of his own. But through reflection he is able to bring the purported evidential relations between his evidence and his belief to light. In this reflection he comes to the judgment that p enjoys almost no justification at all. This normative judgment concerning his evidence functions to disabuse John of the belief in p . John does no further inquiry; he garners no further evidence which operates against his belief. It is simply his reflection and judgment that the evidence supporting his belief is of an exceedingly poor quality that eliminates the belief.

Negative control over belief is in certain ways analogous to the control a spillway on a dam possesses over the water which passes through it. The spillway may normally be left open to allow excess water past, just as our reflective faculties are often not engaged in analyzing the beliefs we are constantly forming. The spillway's control over the water is purely negative, it can do nothing to make the water to flow over the dam, or to flow faster than it would simply through the force of gravity alone. Similarly, the control mechanism specified here does not involve the initiation of belief, and is in this way dependent on the deliverances of experience for certain types of belief formation. But the spillway still clearly controls the flow of water, in that the water is responsive to the state of the spillway. When the spillway is shut, water immediately stops flowing. Similarly, an agent exercises negative control over belief via theoretical reflection when her belief is responsive to her normative judgment that the proposition in question lacks sufficient justification.

Positive control is somewhat similar in structure, but in this type of case reflection leads to the initiation of belief. In Rude Person (RP), Sarah forms no particular belief regarding Kathy's treatment of her during their initial discussion. She does subconsciously note or record a number of things about Kathy's behavior in the course of their conversation, as we all do when talking other persons. Later when reflecting, however, she recalls certain things about Kathy's behavior that together support the proposition that Kathy was being rude to her. Again, she does no further inquiry into the world, she doesn't call a friend and ask about her impression of the conversation, she simply reflects upon her memories and their evidential import. And the judgment that is the culmination of this reflection initiates the belief in question. Here we see another modality of reflective control. Positive control is what we regularly exert when we reflect upon our experiences, memories and beliefs, and via reflection and judgment formulate new beliefs concerning topics about which we previously had drawn no conclusions.

Pragmatic control differs from the other two types. In exercising pragmatic control, an agent reflects upon and makes a judgment concerning her standards of evidence operative in the formation of her belief. As several philosophers have pointed out, the formation of a belief involves more than simply a body of evidence.³⁷ Belief formation comes about through the interaction of evidence with some kind of normative standard of sufficiency. If in the judgment of the individual the evidence rises to or above the standard in question, then belief results. If it does not, then belief is withheld. The two

³⁷ See for example Helm (1994), as well as Owens (2000), pp. 23-35.

previous modalities of reflective control involved an assessment or judgment concerning the quality of the evidence in question, on the basis of the relevant evidential standard. Pragmatic control involves an assessment or judgment of the standard of sufficient evidence itself, on the basis of certain relevant features of the circumstances, such as the importance of the issue in question, the time that the subject has to decide, etc.³⁸

In *Desperate Househusband* (DH), Ned becomes convinced that his operative standard of justification governing his belief about the real estate scheme is erroneous, given the stakes. Ned's judgment has a different content than the judgments in the previous two cases. Ned does not judge that his evidence in fact fails to meet his current standards of justification. He instead judges that his standards of justification are inappropriate for the circumstance, and need revising. As I characterized the example, it is Krista's input that prompts Ned's reflection, and one might object that this is new evidence that influences his ultimate rejection of his belief, not simply reflection.³⁹ But

³⁸ The claim that we possess pragmatic control over our beliefs, while not flowing directly from epistemic contextualism or the related thesis of subject-sensitive invariantism (SSI), is certainly connected with these theses. Contextualism and SSI are in part normative theses about the appropriate standards for knowledge and/or justification and their variance across different contexts. My discussion of pragmatic control implies that subjects may believe that the relevant standards of justification vary given different contexts. Of course subjects could be mistaken about this. But the two ideas are certainly related. It seem to me that contextualism and SSI support pragmatic control.

³⁹ Whether the simple introduction of a piece of evidence into the reflection concerning a belief can undermine reflective control is a complex question. There are a variety of interesting related questions in this area, concerning the nature of reflection versus inquiry, Owens' distinction between first and second-order reflection and content (see pp. 62-63 below), etc., which are beyond the scope of my discussion here. I am inclined to think that new evidence or input from the world does not automatically undermine claims that a belief was formed/eliminated via reflective control; the important issue is how it functions in the reflective process and the resulting judgment.

Krista plays no essential role in DH, we can just as easily imagine Ned reflecting on his own as he sits in the driveway of his home, before talking to Krista, and coming to the same judgment that he needed to be more critical. The same elimination of belief would result.

The distinction between pragmatic control and the other forms is subtle, but one that can (and should) be maintained.⁴⁰ Consider two hypothetical subjects who have come to the judgment that they ought not to continue believing some proposition p , and have in making this judgment stopped believing that p . One, when asked why she no longer believes that p , responds that given her reflection, she has realized that the truth of p is less probable than she thought. The other, when asked the same question, responds that she now believes that she failed to grasp the importance of the situation, and thus she now feels uncomfortable with her previous endorsement of p , given this fact. I would suggest that the former individual has likely eliminated the belief via an instance of negative control, while the latter has done so via pragmatic control.⁴¹

⁴⁰ Given that belief formation and elimination come about through the interaction between evidence and standards of sufficient justification, it may be that the judgment that the standard is mistaken is by itself not enough to result in the elimination or formation of a belief. It may be that the agent must make an additional judgment regarding the new standing of her evidence with respect to the revised standard in order for her doxastic states to change. Exploring this complex issue is beyond the bounds of this project. Even if the second judgment is required, though, I still think one can maintain the distinction between this kind of control and the other forms I note in this section.

⁴¹ It may be helpful here to speak here in terms of probability, though I am not suggesting that subjects typically reflect upon their beliefs in precisely this way. We can construe both subjects as initially believing that the evidence for p makes p .6 probable, and as both seeing the relevant evidential requirement for this proposition as .5 probability. The suggestion is that, in exercising negative control, the first subject came through reflection to something like the judgment that the evidence in question in fact only makes $p < .5$

The general picture offered by the various examples is of reflection serving in an executive position over our belief determining processes. Like a business executive, we often delegate the function of rational belief formation to lower levels of cognition, in this case largely unreflective or subconscious cognitive processes, and only exercise actual influence in the case of (perceived) malfunction. In such cases, we make through reflection the judgment that we have misinterpreted the evidential weight of certain reasons, or failed to utilize an appropriate standard of evidence, or something related. We exercise control over belief via this judgment when, as in the examples, the judgment results in the formation of a new belief, or the rejection of a belief previously held.

The three modes specified above do not exhaust the possibilities for reflective control.⁴² They do however offer the basic structure of an intellectualist account of control over belief. They also highlight the importance of this kind of control for our conception of doxastic responsibility. As noted above, reflective control may go awry. A subject may erroneously assess her evidence, or she may misidentify the relevant

probable. On the other hand, the second subject came through reflection to something like the judgment that, given the importance of the situation, the relevant standard of sufficient justification for this proposition was something $>.6$.

One additional significant difference between negative/positive and pragmatic control concerns the wider impact of the respective reflective judgments. As Helm (1994) discusses, standards of justification frequently have scope over a range of beliefs concerning a particular subject matter or context. If one's judgment is that the relevant standard of justification is mistaken and requires modification (becoming more or less strict, for example), this may have rippling effects on a wide range of beliefs in one doxastic structure. By contrast, the simple judgment that one's evidence for p is not as good as one thought it was will typically (though not always) have more limited or isolated impact.

⁴² For example, reflective judgment likely sometimes has a more subtle controlling effect on our beliefs, by causing us to revise our confidence in a given proposition, without leading to the adoption or rejection of a particular belief.

epistemic standard and end up eliminating beliefs she should have retained, retaining beliefs she should have eliminated, or initiating beliefs that are poorly evidenced. Each of these failures is nevertheless an affirmation of the concept of reflective control. Just as we may guide our actions poorly, we may manage our beliefs poorly as well. In either case, the intellectualist contends, they become our beliefs, states which we are properly held responsible for, through their connection with reflective control.

2.4. Owens on Belief

David Owens, as noted previously, accepts the intellectualist account of our direct control over our actions, arguing that it is for various reasons more plausible than the voluntarist alternative,⁴³ and more consistent with his modest compatibilism about freedom and responsibility.⁴⁴ However, he rejects the intellectualist account of direct control over belief. In this section and section VI I will briefly outline the main elements of his complex argument, and in sections V, VII, and VIII I will try to suggest both why Owens' argument fails to undermine the intellectualist account of control over belief just described, and how it helps us to see more clearly what this account amounts to.

⁴³ See Owens (2000), p. 78-85. Owens argues, amongst other things, that though the voluntarist says that intention formation is necessary to free action, intentions are in fact only *required* for long term planning and extended activities. Contrary to the voluntarist position, Owens claims, we sometimes act without the formation of an intention, on the basis of our judgment alone.

⁴⁴ As noted previously, while voluntarism and libertarian incompatibilism are not necessarily linked, the voluntarist conception of freedom, with its emphasis on the free decision of the will, seems more naturally congenial to libertarianism than does intellectualism. Owens admits that his picture of control may be unsatisfying for libertarians, but he rejects their demand that true control be "total and unconditional." He opts rather for a "modest compatibilist notion of freedom." (Owens (200), p. 11)

To set the stage for his criticisms, Owens first sketches an intellectualist account of control over belief largely in line with the one offered above. (Indeed, as I've noted, my account is indebted to Owens' in various ways). On this account, reflection and judgment are the mechanisms of control over beliefs, which properly ground ascriptions of responsibility for them. Borrowing from Christine Korsgaard, Owens characterizes reflection as the mechanism by which we achieve a certain critical distance from our desires and perceptions.⁴⁵ As Korsgaard notes, this creates a special problem for us:

It is the problem of the normative. For our capacity to turn our attention on to our own mental activities is also a capacity to distance ourselves from them, and to call them into question. I perceive, and I find myself with a powerful impulse to believe. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn't dominate me and now I have a problem. Shall I believe? Is this perception really a *reason* to believe? I desire and I find myself with a powerful impulse to act. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn't dominate me and now I have a problem. Shall I act? Is this desire a reason to act? The reflective mind cannot settle for perception and desire, not just as such. It needs a reason. Otherwise, at least as long as it reflects, it cannot commit itself or go forward.⁴⁶

At this critical distance desires, perceptions, beliefs and related mental phenomena are assessable as reasons. Through the process of reflection, we form normative assessments or reflective judgments concerning these inputs, and these reflective judgments determine

⁴⁵ Ibid., p. 9-11. Owens draws from the account in Korsgaard (1996), p. 92-93.

⁴⁶ Korsgaard (1996), p. 93.

our intentions or beliefs. Freedom and self-control are thus about “being able to do what you think you ought to do,” meaning the ability to conform to one’s practical or theoretical judgment.⁴⁷ Insofar as a belief is the result of one’s reflective judgment, one has control over that belief, and is responsible for it.

Owens suggests that a worry about the picture of reflective control arises at this point, a worry noted previously. Clearly I do not move around the world in a constant state of conscious deliberation and judgment. Most of my beliefs are formed quite unconsciously and automatically. But how does this square with the idea that reflective control is necessary for responsibility? To accommodate this fact, Owens introduces the distinction between *first-order* deliberation and content and *second-order* deliberation and content, along with the related concept of a *non-reflective awareness* of evidence. Mental states with first-order content are focused on features of the world, while those with second-order content make reference to other mental states like one’s desires & beliefs.

An example will help to illuminate these distinctions. Say that, in a paradigmatic case of perceptual belief, I look out the window and immediately come to believe that there is a green tree outside, with no intervening reflection. Here my belief has first-order content, it is about features of the external world. As in most cases of perceptual belief, here features of the world impinge upon the mind and seemingly directly determine my belief. When a belief is formed in this way, Owens suggests, I typically have only a non-

⁴⁷ See Owens (2000), pp. 78-79.

reflective awareness of the evidence. I am at some level aware of what motivates my belief, meaning I am at least in principle capable of articulating what evidence or grounds I have for my belief. When asked why I believe there's a green tree outside, I would reply that I saw one. When pressed further, I would perhaps refer to the general veridicality of such perceptual experiences and the connection of this particular perceptual experience with my belief. But this evidence, despite my latent awareness of it, played no role in a conscious inference leading to my belief. I engaged in no second-order reflection about the contents of my perceptual experience and what could plausibly be inferred from them, or what evidential relations they bear to the belief in question. I simply saw and believed. My non-reflective awareness of the evidence, however, does give me the capacity to engage in such reflection. And, Owens suggests, my general capability to engage in such reflection and for it be efficacious in determining my beliefs grounds ascriptions of responsibility concerning belief.⁴⁸

What has been articulated thus far, Owens says, is captured in his principle of Reflective Motivation:

Reflective Motivation (RM): If *R* is a *prima facie* reason to believe that *p*, reflection on *R* provides the rational subject with a motive to believe that *p*.

⁴⁸ Ibid., p. 15. In developing this picture of reflective control, Owens does not wish to suggest that we have the capacity to control every one of our beliefs. Circumstances may dictate that we form beliefs without any possibility of reflection. Owens mentions the person who seems to perceive a truck bearing down upon her, and forms such a belief before leaping out of the way. If it turns out that she was mistaken in her perception (say she was duped by a hologram), we would not hold her responsible for this faulty belief formation, in part because of the fact that in the circumstances there was no possibility of her exercising reflective control over this belief.

Owens suggests that only if RM is true, only if our reflection is capable of providing motivation to our beliefs, do we have reflective control over our beliefs. Owens' principle is formulated positively, in term of reasons *for* belief, reasons functioning in the formation of belief, but presumably it is intended more generally to apply to the determination of belief. More generally, then: Only if those things, which in the context of reflection are *prima facie* reasons for or against belief, can also function motivationally in the determination of belief, do we have reflective control.

Owens notes that by itself RM seems like "mere common sense."⁴⁹ However, he ultimately contends that both the principle and the general conception of reflective control dissolve under analysis. His initial argument against reflective control consists of two parts: an argument against the efficacy of reflective judgment, and a more circuitous argument against evidentialism. I will briefly sketch both claims.

Looking more closely at reflective control, Owens says that we can see that it analyzes into two related necessary components. The first is the *epistemic* component of control. This is the requirement that one's reflection properly capture one's reasons for believing. If in reflecting I surmise that my belief is based on certain pieces of evidence functioning as reasons, but it is in fact caused by a brain lesion, then I could hardly possess rational control over this belief via reflection and judgment.⁵⁰ The second component of reflective control is *motivational*. If I am to exercise reflective control over

⁴⁹ Ibid., p. 20.

⁵⁰ This assumes, plausibly, that the brain lesion bears no rational connection with these reasons.

my belief, then my reflective judgment must be able to motivate or determine my belief. Were I to reflect on my reasons and judge that they support or fail to support some belief that p , but my belief proved unresponsive to this judgment, then I would fail to possess any kind of reflective control over p .

This latter aspect of control Owens characterizes in an additional principle, Reflective Rationalization:

Reflective Rationalization (RR): If R provides (a) subject with a reason for belief, the judgment that they have reason R also provides them with a reason for belief.

Owens' attack against reflective control focuses on the motivational component of reflective control captured in RR. The difficulty becomes clear when we consider the typical nature of belief, alluded to above in the discussion of first and second-order deliberation. I commonly form beliefs on the basis of a non-reflective awareness of features of the world, which at the first order determine my belief, without the benefit of any conscious reflective judgment. But then, even if I do reflect upon such beliefs and the attendant evidence, Owens notes, my reflective judgment exercises no motivational force. If I stop to consider the evidence supporting my belief that there is a green tree outside, and judge that it is sufficient to support the belief, what is the effect of this judgment? The motivation has already been provided by my non-reflective awareness of the evidence. As the judgment that I have sufficient evidence for a belief is not an additional reason to believe, it provides no further motivation to that provided by my non-reflective awareness. As Owens says, "If you already have a non-reflective awareness of the

reasons which ought to motivate you, how does the judgment that you ought to be moved by them help to ensure that you are so moved? Such judgments look like an idle wheel in our motivational economy, whether we are perfectly rational or not.”⁵¹ But if reflective judgment is impotent towards belief, Owens says, then it seems doubtful that we possess reflective control over our beliefs.

Owens second argument is more roundabout. He considers evidentialism, the thesis that “what justifies belief in *p* is just evidence in *p*’s favor.”⁵² Owens points out that, at least in this strict form, evidentialism must be false. When we form a rational belief that *p* on the basis of reflection, we do so because we take ourselves to have not just evidence for *p*, but *sufficient* evidence for *p*. The evidence itself, however, cannot give us the relevant standard of sufficiency. This must be provided by something else, and thus evidence is not the only thing operative in the justification of beliefs.⁵³ The something else, Owens argues, consists of various pragmatic factors, such as the importance of the issue, the time and resources of the agent available to attend to it, the costs of various errors, etc.⁵⁴ These factors play a role in specifying what constitutes sufficient evidence for a particular belief.

⁵¹ Ibid., p. 18.

⁵² Ibid., p. 23.

⁵³ Paul Helm discusses the same point in Helm (1994). Interestingly, Helm takes this fact about justification as the basis for an argument *for* control over belief, albeit in a somewhat more indirect fashion than the control to be discussed below.

⁵⁴ See Owens (2000), pp. 23-35.

Owens notes that this admission of pragmatic elements into the justification of belief is distinct from pragmatism, the thesis that the benefits of believing *p* properly play a role in its justification.⁵⁵ Nevertheless, allowing pragmatic factors to play a role in any fashion is problematic for the thesis of reflective control. RM says that *any* reason for belief must also provide motivation for belief. But can pragmatic factors provide this motivation? Surely not, says Owens:

Ask yourself whether it is raining; what comes to mind is evidence for and against rainfall—the sunshine outside versus the watery blobs on the window pane. You don't decide that it is raining after contemplating how little time you've got to decide, how much the issue matters, what the consequences of making up your mind are, and so forth. Even if you do happen to reflect that time is running out, this reflection does not move you to belief in a way that it might move you to act. And you don't feel that your rationality is at all compromised by this fact.⁵⁶

Owens contends that these pragmatic factors are reasons for belief; they make belief rational in appropriate circumstances. And yet, as we see, reflecting on them can provide no motivation for belief. This shows us that RM is false, at least as currently formulated. But if RM is false, Owens argues, then we lack reflective control over our beliefs.

Owens claims that reflective judgment is inefficacious for beliefs, and argues that the core principle of reflective control is false. But if reflective judgment does not motivate or determine belief, then we lack reflective control over our beliefs. And if, as the intellectualist argues, responsibility requires such control, then we must reject

⁵⁵ Ibid., p. 29.

⁵⁶ Ibid., p. 32.

ascriptions of responsibility for believing. In seeking to avoid this counterintuitive conclusion, the intellectualist is left with a dilemma. She must either discard the intuitive connection between responsibility and control in favor of an alternative, looser connection, or find some further support for the idea of reflective control. Owens chooses the former option. In the remainder of this chapter, I argue that the intellectualist should opt for the latter.

2.5. The Idle Wheel of Reflection

I will first briefly respond to Owens' charge that reflection is an "idle wheel" in the determination of belief. As noted in the previous section, Owens' central claim regarding reflective control is that when a belief is formed via a non-reflective awareness of the evidence, reflecting later upon this evidence and forming a reflective judgment endorsing it can add nothing to the motivation for the belief, and thus can offer no source of control over belief formation, modification or retention. I offer two related responses. First, Owens supports his claim with cases where there is congruence between one's non-reflective awareness and one's later reflective assessment of the evidence. But we can better see the true role of reflection and judgment in determining belief in cases where non-reflective awareness and reflective judgment come apart or differ. Here it is clear that reflection plays a central role in determining belief. Secondly, these cases of disagreement show us that even in the types of cases Owens relies on, reflection ought not to be thought of as an idle wheel.

Black Widow, Rude Interlocutor and Desperate Househusband from section 2.3 illustrate the first point. Crucially, each of the examples highlights the potential gap or disconnect between the deliverances of non-reflective awareness and those of subsequent reflection. In each example, as noted previously, the subject fails to cognitively connect up with the facts of the matter in his or her initial process of belief formation, which occurs non-reflectively. In the subsequent process of reflection, it is precisely the subject's evaluation of the evidence he or she possesses, precisely his or her reflective judgment, which determines the subject's belief formation or dissolution. These examples show us that, far from being superfluous, reflective judgment is often absolutely necessary (and sufficient) in the determination of belief.

The propositions that comprise the evidence for our beliefs do not wear their evidential relations to other propositions on their sleeves, so to speak. Even if Owens is correct to characterize the evidence as directly determining one's beliefs at the first order, this determination occurs through the interpretive framework of the agent's psychology. Assessment of some kind is involved. This much is clear from the different reactions often made by different persons to the same experiences or basic set of evidence.⁵⁷ But this same kind of variation of assessment that occurs between different persons can also occur within the same person at different times. Owens is correct to say that one's initial assessment of evidence often occurs non-reflectively or subconsciously. But our initial assessments are often mistaken, and it is primarily through reflection that we are able to

⁵⁷ For an interesting discussion of this issue, see Cohen (2000).

correct these errors. When, as in the three examples mentioned, reflection plays this key role, it is no idle wheel. For the intellectualist it is the central mechanism of control over belief.

Examples where reflection clearly plays this determinative role help us to see both that Owens' analysis is significantly incomplete, and that the kinds of cases used to support Owens' claim are of a special kind that is not particularly useful in identifying the true role of reflection in belief formation. The initial plausibility of Owens' claim regarding the vestigial nature of reflective judgment may also be due in part to the way he characterizes the relevant cases of reflective agreement, as well as certain philosophical convictions about belief. In evaluating this claim, it is useful to consider a case of reflective agreement, similar to those used by Owens, but from the practical realm:

14. *Elderly Motorist*: Steve is driving to work one day when he notices an elderly man pulled over on the side of the road, struggling to change a flat tire on his car. Steve, a compassionate person, responds to this sight quickly and without reflection by forming an intention to help the man, immediately changing to the right lane and beginning to pull off the road to assist. As he pulls over, Steve reflects on the situation and asks whether he really ought to get out and help the man. After brief consideration, Steve confirms his instinctive decision as the best one, gets out, and helps the man to change his tire.

This example parallels the types of theoretical cases that Owens uses to support his claim. But it seems implausible to claim that in this case Steve's practical reflection and judgment was an "idle wheel" in the sense that Owens desires. Owens means by this that the judgment was ineffectual and incapable of motivationally contributing to Steve's intention and action. But Steve's intention is obedient to and determined by *both* his non-

reflective awareness and his reflective judgment, and the latter more than the former. Consideration of the closest counterfactual cases (where Steve, after his initial intention formation, reflects and decides that he *ought not* help the man) shows this. Steve's reflective judgment may have been an unnecessary addition to his intention formed via non-reflective awareness, but this is simply an accident of the relationship between the two in this particular case.

When we place them in a larger framework of the relationship between non-reflective awareness and reflection we are able to make a more accurate analysis of such cases. Rather than suggesting, as Owens does, that reflection is an idle wheel in the production of belief, it is more plausible to think that in these cases the role of reflection is masked by its agreement with the initial non-reflective assessment of the evidence. Steve's action in this case is over-determined by the outcome of his non-reflective awareness and his reflective judgment. But the fact that in such instances one's reflective judgment happens to coincide with the results of one's non-reflective belief-forming processes does nothing to undermine the control available through reflective judgment.

The relationship between non-reflective awareness and reflection in belief formation is complex and deserving of a fuller analysis than I can provide here. However,

a complete exposition of the different inter-relationships between non-reflective awareness and reflective judgment would include at least the following categories:⁵⁸

Reflective Correction: Reflection on the evidence leads to a judgment contrary to the deliverances of non-reflective awareness, and the subject adopts a propositional attitude (belief/non-belief) contrary to that formed via non-reflective awareness. The initial examples under discussion from section three are instances of reflective correction.

Reflective Adjustment: Reflection on the evidence leads to a judgment in line with the deliverances of non-reflective awareness, but one which does not accord precisely with the level of conviction established via non-reflective awareness. Cases of reflective adjustment would include both instances of reflective weakening (reflective judgment undermines, though not completely, the output of non-reflective awareness) and reflective “perma-coating” (reflective judgment offers further support to the belief state established via non-reflective awareness).⁵⁹

Reflective Over-determination: Reflection on the evidence leads to a judgment corresponding to the outcome established via non-reflective awareness. The examples used to support Owens’ claims are likely best understood as instances of reflective over-determination.⁶⁰

In cases of over-determination within the philosophy of action, the action is typically considered to have at least two sufficient causes, either of which is individually sufficient to determine the action.⁶¹ One of these causes may originate from within me,

⁵⁸ I am limiting consideration here to only those cases where the subject reflects only on the evidence that is the deliverance of the original non-reflective belief formation/rejection process. Greater complexity surely arises in cases where additional or different evidence is part of the reflective process.

⁵⁹ The brief taxonomy and some of the terminology here are both deeply indebted to Peter Graham.

⁶⁰ My thanks also to John Martin Fischer for clarifying for me the idea of reflective over-determination, and the connection with similar concepts in philosophy of action literature.

⁶¹ See, for example, the forest of such examples in the literature that developed in response to Harry Frankfurt’s nefarious intervener Black (see Frankfurt (1969)), such as those in Fisher (1998). I am not suggesting that cases of reflective over-determination of belief precisely parallel over-determination cases in the philosophy of action. Fischer notes two different types of over-determination, (1) “pre-emptive”

while another may be outside of my control. But just as in such cases we would not consider me “idle” on the basis that my cause was not strictly necessary for the occurrence of the action, we similarly should not consider reflective judgment idle in cases of reflective over-determination. And just as we do not take cases of over-determination of action to be indicative of our general capacity to govern our actions, so we should not utilize cases of reflective over-determination as our principal guides in settling the significance of reflective judgment.

Owens’ examples are thus not without value, but their actual significance must be correctly understood. They are part of the universe of possible relationships between reflective judgment and non-reflective awareness, but not the most useful for answering the question of control. Returning to an earlier (loose) analogy, a business executive’s agreement with or deference to the choices of her subordinates does nothing to diminish her control over the corporation. Such agreement may create the illusion of idleness on the part of the executive function, but this illusion is dispelled immediately by instances of disagreement. In such cases the executive’s influence and control over the organization is clearly seen. Likewise, cases of disagreement in belief determination clearly show us

(such as Fischer’s “Assassin”), in which a second cause, which typically does not in fact operate, nevertheless ensures that the outcome of the first cause would have happened anyways, and (2) “simultaneous” (such as Fischer’s “Joint Assassins”, in which two causes, each sufficient for the given outcome, operate independently and at exactly the same moment. The details of reflective over-determination do not precisely match either of these types of cases. I do not think, however, that any differences will be material to the claim I am making here, which is simply that the “charge” of idleness is equally inapplicable to most cases of over-determination in both realms.

the position and control enjoyed by reflective judgment over the belief determination process.

2.6. Reflection and the Intrinsic Authority of Judgment

Owens' initial idle wheel claim may seem implausible, but he has more to say on this matter. Owens briefly considers the issue of the potential conflict between one's normative judgments about what one ought to believe and the beliefs themselves.⁶² Even in cases where one's judgment is at odds with the deliverances of non-reflective awareness Owens argues, consistent with his earlier claim, that belief is immune to control via normative judgment. Below I briefly summarize his position, and then offer reasons why, in light of the kinds of cases discussed previously, I believe he is mistaken. For simplicity's sake, I will focus my discussion on BW and negative control, though the argument can be easily extended to include the other cases and modes of control mentioned in section III.

Owens claims that practical and theoretical deliberation differ regarding the *intrinsic authority* possessed by judgment in each respective sphere. It is in this difference in intrinsic authority that Owens grounds our rational control over action, and our lack of such control over belief. Consider first practical deliberation. When I am deliberating about a certain course of action, I may come to the judgment that I ought not

⁶² See especially Owens (2000), p. 101-113.

to pursue it. Because of the intrinsic authority of judgment over practical reason, Owens argues, that course of action becomes irrational for me *simply in virtue of my having made such a judgment*. That is not to say that I will, in the end, avoid this action, I may be weak-willed with respect to it. But, if I perform it, my action is nevertheless practically irrational because of its conflict with my practical judgment. This is consistent with the intellectualist position we ascribed to Owens previously. For the intellectualist, control consists in conformity to judgment. So we may say that, insofar as I am rational and in control, practical judgment will be motivationally efficacious over my actions.

Owens himself characterizes the intrinsic authority of practical judgment in the following way:

My views about what I ought to do have a right to control my actions which is independent of the cogency of those views: my practical judgments have an *intrinsic authority* over what I do, not merely an authority which derives from the beliefs and desires which should motivate some action. Of course, the practical judgment must make *some* sense in terms of the agent's perceived needs and interests, otherwise it would be hard to see it as a judgment of the agent at all. But the judgment's authority, its ability to determine the rationality of both intention and action, is not a product of its cogency.⁶³

A key feature of practical judgment's authority is brought out here, the fact that its authority is not connected with the actual quality of the reasons for which the judgment is made. I may in fact be mistaken, Owens says, about the quality of my reasons. But given

⁶³ Ibid., p. 104.

the intrinsic authority of practical judgment, an action contrary to my practical judgment is nevertheless irrational.

An example is helpful here. Sarah is considering whether to purchase a boat or a jet ski. She considers the fact that since childhood she has wanted a boat, that the boat she is presently looking at is on sale, and that a somewhat smaller version of this same boat was recently purchased by her neighbor, Ted, whom she wishes to both impress and upstage. On this basis, she decides that she ought not to purchase a jet ski, that she ought to instead purchase the boat. Assume she ignores, amongst other things, the fact that she can't really afford the boat, its additional purchase and maintenance costs over against the jet ski, her inability to store it at home or to afford the necessary moorage, and the fact that she's never actually driven a boat before. Owens nevertheless argues that in virtue of the authority of her practical judgment, insofar as she is rational, Sarah will not purchase the jet ski, and will purchase the boat instead.

Now we turn to the theoretical sphere. Here, Owens argues, judgment lacks the same kind of intrinsic authority that it possesses in the practical sphere. My normative judgments regarding my beliefs will not, Owens says, control my beliefs simply insofar as I am rational. Considering a case of belief analogous to the case of practical judgment just above, he writes:

If I arrive at a faulty judgment about what I ought to believe, there is no rational requirement on me to conform my first-order belief to this higher order judgment. If my first-order belief is well-grounded in various evidential and non-evidential considerations, that belief continues to be

rational, my higher order disapproval of it notwithstanding. I have no veto over the rationality of my own beliefs, my higher order judgments are not a source of rational motivation for belief, and thus I have no reflective control over which beliefs I adopt.⁶⁴

This conclusion is puzzling, and particularly so in light of the examples & discussion offered in section III and following. I contend that Owens comes to this conclusion only via a confusion or equivocation over the notion of “rational” at work in the two different domains of practical and theoretical reason. In the following section I attempt to offer a unified intellectualist account of the authority of judgment over action & belief, and will diagnose Owens’ equivocation through the use of a widely held distinction concerning rationality.

2.7. A Unified Conception of Judgment

The term “rational” is commonly understood in at least two senses, or at two levels of evaluation.⁶⁵ Rationality at the first level of evaluation, which I will call “L1 rationality”, refers to the relationship between an individual’s beliefs, intentions or actions and the facts in the world.⁶⁶ At this level rationality is a function of how a subject interacts with the reasons imposed by the state of the world. At L1, “rational” and related

⁶⁴ Ibid., p. 108.

⁶⁵ This distinction, which goes by various names, is made in a variety of areas of philosophy, including ethics, epistemology, philosophy of action and elsewhere. One such example is Parfit (forthcoming), pp. 67-72. Parfit explicitly acknowledges the two different forms of evaluation, though for his purposes he is principally concerned with L1 rationality and refers to L2 rationality as merely “apparent”.

⁶⁶ As will become more apparent in the next chapter, the L1 and L2 descriptors have nothing to do with relative importance or priority.

laudatory terms refer to circumstances where a subject responds well or appropriately to the facts, while “irrational” and similar terms imply a failure to respond appropriately to the facts. Rational at the second level of evaluation, which I will uncreatively call “L2 rationality”, refers to the relationship between an individual’s beliefs, intentions or actions and her other mental states. This second level of evaluation is concerned chiefly with issues of self-consistency, how a subject conforms to the requirements or constraints placed upon her by her own beliefs, desires, commitments, etc. Both of these forms of assessment are commonly employed both in philosophical literature as well as our common ways of speaking.⁶⁷

In part because of our imperfect and limited grasp of the world, these two forms of evaluation often come apart. Most commonly individuals are L2 rational while being L1 irrational, possessing at least a measure of consistency within their relevant beliefs, desires and intentions, but failing to connect with the facts available to them. Joe decides to buy flowers for Angela, based on his attraction to her, and his belief that she has given him signs of her affection for him. It is reasonable for Joe to form this intention and act in such a way (he is L2 rational), given his beliefs and desires. But if Joe has grossly misinterpreted simple gestures by Angela, and failed to understand her clear signals that

⁶⁷ Thanks to Matthias Steup for the suggestion of the L1 & L2 monikers, as well as for his feedback on a highly compressed version of this chapter. It might seem that the terms “internal” and “external” or “subjective” and “objective” would be more appropriate than the ones that I’ve chosen. While I utilized the former distinction in earlier drafts, I have dropped it here to avoid confusion and other issues associated with these philosophically loaded terms. One reason to resist applying the latter set of labels “subjective” and “objective” to the senses of “rational” here is that the external sense of rationality may not refer to an objective assessment of the state of an individual’s reasons, there may be no such assessment available. The relevant norms by which an individual’s reasons are to be judged may be only broadly inter-subjective.

she harbors no romantic interest in him, if his belief is in fact not true, then he is nevertheless L1 irrational in some significant way.

A less common failure involves L1 rationality with L2 irrationality. This is a subject whose belief-state or intention is consistent with the facts available, but inconsistent with her other beliefs, intentions, desires and other mental phenomena. Various cases of *akrasia* are instances of this type of irrationality, including our initial example of Eric and his akratic lunch choice from section 2. I will further describe and analyze this type of phenomenon just below.

While it may initially seem so, L2 rationality is not simply a first-person oriented concept, a notion which I would *only* apply to my own reasons. Both senses or levels of rationality are employed in our normal practices of assessing our own and others' actions and beliefs. A statement like, "Well, it certainly was rational of him to think that, given that he didn't believe anyone would be home now," is clearly a claim about L2 rationality despite the fact it is employed in the assessment of someone else's actions. On the other hand, "We can see, in light of all that has been argued here, that those who purchase bonds instead of stocks are acting irrationally," is likely an example of analysis in terms of L1 rationality. The suggestion is that the purchasers of stocks, whatever else may be said of them, have failed to connect up with the relevant facts in the world.

In alleging differences in our control over action and belief, Owens equivocates between these two different levels of analysis and corresponding uses of "rational". Concerning action, Owens is clearly utilizing "rational" in the L2 sense. We can see this

in his argument that, irrespective of the actual quality of my reasons, I am irrational if I act *for* these reasons, but *against* my practical judgment of what I ought to do. This is clearly a correct assessment at one level; the irrationality Owens alleges refers to a kind of incoherence in the agent, between her judgment and her resulting intention. For the intellectualist, this type of incoherence is what constitutes a lack of control over one's actions. Indeed, it is precisely this kind of irrationality that we attribute to at least one kind of *akratic* agent. Owens is right to claim that practical judgment has intrinsic authority over our actions. The ability to guide our actions by our judgment in this way is an essential feature of our agency. Insofar as we are rational in the L2 sense, we will possess control over our actions, meaning they will accord with our practical judgments about what is best to do. When they do not, something has gone wrong, and this is what we mean when call the agent in question "irrational" in this sense.

Nevertheless, there is still a perfectly intelligible and appropriate analysis of such actions according to which they are rational in the L1 sense of the term. After all, the agent in question acted for the best reasons, even if she was flawed in her assessment of them. Using the same reasoning, it is clearly appropriate to call our prospective boat owner Sarah irrational, insofar as she purchases the boat, and acts against what is clearly in her best interests. This is true despite the fact that in this case Sarah successfully guides her action via her practical judgment. Sarah is not incoherent, but she is irrational insofar as she fails to access the most relevant reasons for action, and thus harms herself.

This also is a common form of evaluation that we apply to both ourselves and other agents.⁶⁸

However, when Owens turns to analyzing theoretical judgment, he uses “rational” in the L1 sense. This is the only intelligible way to interpret his statement that if my belief remains “well grounded in various evidential and non-evidential considerations, that belief continues to be rational, my higher order disapproval of it notwithstanding.” Now again, this is a perfectly legitimate conclusion to reach, in terms of L1 rationality. Insofar as I hold a belief for objectively good reasons, I am rational in this latter sense. But this leaves many questions unanswered regarding the former, L2 sense of rationality in terms of which Owens articulated the intrinsic authority of practical judgment. Owens, for his part, is silent on these issues.⁶⁹ He seems to assume either that he is discussing the same notion in both cases, or perhaps that L1 rationality is the only one relevant for beliefs. But the former claim is clearly false, and I see no reason to accept the latter.

Indeed, the discussion of practical deliberation showed us that it is L2 rationality that is at the center of intellectualist considerations of control. From the intellectualist perspective, a closer examination of Black Widow (BW) shows a form of rational control

⁶⁸ I discuss these two forms of evaluation further in Chapter 3.

⁶⁹ On pp. 104-105, Owens discusses the famous case of Huckleberry Finn, discussed by Bennett (1974) and others. Owens’ discussion seems to indicate his denial that the external sense of rationality is relevant in the assessment of agents’ practical judgment and action. On p. 106, Owens makes this denial more explicit, saying, “But my action cannot be rational so long as it fails to conform to the verdict of practical reason, however unfounded that verdict may be, and however understandable the rebellion against it.” This only makes his claims about theoretical judgment more befuddling.

in the theoretical sphere analogous to the practical. Consider the content of the judgment in question in such a case of negative control. In BW a belief is formed through a non-reflective awareness of the evidence. However, upon reflection, the belief is eliminated by a negative normative judgment, a judgment that the subject ought not to believe the proposition in question.

There seem to be two broad possibilities for the content of this judgment. A subject might negatively assess his belief on the basis of non-evidential considerations. Consider the case of someone who reads a persuasive article denying the reality of objective moral norms.⁷⁰ If the person finds himself largely convinced by the article's claims, he might nevertheless upon reflection find his belief morally troubling, and judge that, for various reasons, he ought not to hold it. In such an instance, however, the agent in question may, despite his reflection, fail to conform his new disbelief in the reality of moral norms to his judgment.⁷¹ This failure results from the fact that, given what he has encountered thus far, and despite his negative judgment, the evidence *seems to him* against the possibility of objective moral norms.

But this type of case is clearly in the tiny minority. The standard case is one where the agent's judgment involves or arises from a reassessment of the evidential basis

⁷⁰ Robert Adams offers an interesting discussion of such a case in Adams (1995).

⁷¹ Of course, analogous cases can be constructed for action as well. In places where I have putatively strong self-interested reasons to pursue a course of action, but feel that I nevertheless ought not do so because of ethical considerations, I may nevertheless give into self-interest and act in this direction, all the while maintaining that I ought not do so.

for the relevant belief. This is clearly the proper description of BW. In this type of case, Owens is mistaken to think that, in the relevant L2 sense, one's belief might remain *rational* in spite of one's negative normative judgment. This is because the reflective judgment has undermined any rationalizing evidential basis that the belief possessed previously. The content of the judgment simply *is* that one has no good evidence for the belief. In this instance, an agent whose belief did not conform to his reflective judgment would be L2 irrational, just as in the practical realm. From the agent's perspective, the perspective that matters for control, there is no good reason to continue holding the belief.

It is this recognition that grounds our intuition that in BW John *would* be deeply irrational, that there would be something wrong with him, were he to recognize the poor support his evidence provides for his belief, judge that the belief has no grounding, and yet go on believing it nevertheless. We would think that John had in an important way lost control of himself. Moreover, it is telling that the agent in this type of situation would likely judge *himself* irrational and suffering from some failure of control over his beliefs. Of course none of this is meant to imply that a persistent conflict between belief and normative judgment never in fact occurs. Such a believer is incoherent, but not impossible. Beliefs, like intentions, have a certain cognitive inertia and are sometimes hard to dislodge,⁷² though this particular case stretches the bounds of credulity. On the account offered, the *akratic* believer is possible, just as the *akratic* actor is. There are

⁷² I thank Jack Crumley for pointing this out in his commentary on an earlier version of this chapter.

perhaps relevant differences that make one form of *akrasia* more likely than another, but exploring these is beyond the scope of this argument. The point is that such a believer suffers from the same internal incoherence that the L2 irrational actor does in the practical sphere. The intellectualist must argue that insofar as we are rational, our beliefs, like our actions, will conform to our normative judgments concerning their justificatory status. Theoretical judgment possesses intrinsic authority over belief analogous to that enjoyed by practical judgment over action.⁷³ The consistent intellectualist account, in terms of L2 rationality, can reach no other conclusion.

⁷³ A useful way of expressing the claims above is offered by John Broome's work on the logic of normativity (see Broome (1999)). Broome identifies three different types of normative relations central to our practical and theoretical reasoning: reason relations, ought relations, and normative requirements. Roughly, normative requirements are relations between one's beliefs or intentions and one's other mental states. In a given circumstance of theoretical or practical deliberation, my belief that *p* may normatively require me to believe that *q*. Say, for example, that it is abundantly clear to me that *p* implies *q*. Then insofar as I believe that *p*, I will, if I am rational, believe that *q*.

To say that I am normatively required to believe that *q* is distinct, however, from saying that I have a reason to believe that *q*. A reason relation holds between propositions, or the states of affairs that they denote. In the same circumstance of deliberation, I have a reason to believe that *q* just in case some proposition *p* is true, and *p* is a reason to believe that *q*. In such a situation there may be various reasons for and against believing that *q*. The sum of the relevant reasons in the circumstance tells me whether or not I *ought* to believe that *q*. So, what I ought to believe and have reason to believe is determined by features of the external world, while what I am normatively required to believe is in an important way dependent on my beliefs and other internal mental states, whether or not they accurately reflect reality. Each of these relations is an important aspect of correct reasoning, and each plays into our assessments of a subject's rationality.

From this description, we can see that reasons and ought relations comprise the L1 sense of "rational" discussed above, while normative requirements correspond to the L2 sense. We can recast the discussion above in Broome's terms, working from the assumption that judgments are attributable to agents as beliefs. Owens establishes the intrinsic authority of practical judgment in terms of normative requirements. An agent, insofar as she is rational, will act as she is normatively required to act by her beliefs (practical judgments) about what she ought to do. Through the influence of normative requirements, generated by practical judgments, we exercise reflective control over our actions.

In the theoretical sphere, Owens says that theoretical judgment is not a source of "rational motivation" for belief. Insofar as we interpret him to mean that a given theoretical judgment is not a reason to believe, he is correct. What I have reason to believe is determined by the facts of the situation. Even if I correctly access these facts, and form the appropriate normative judgment, this judgment is no additional reason, beyond the evidence that I already have. However, as we have seen, these considerations are irrelevant for the question of control. Reflective control is a matter of normative requirements, and the

For these reasons, I conclude that the alleged asymmetry in the authority of judgment between the practical and theoretical spheres rests on a confusion. The intellectualist analysis of reflective judgment must be in terms of L2 rationality. On this analysis, theoretical judgments possess the same authority over beliefs that practical judgments possess over action, and thus reflective control for the intellectualist is not limited to actions, but extends to beliefs as well. As I suggested previously, the execution of reflective control over beliefs formed via non-reflective awareness is one of the essential functions of theoretical reason.

2.8. Pragmatic Factors and Reflective Control

At least one serious concern remains for the foregoing account of reflective control. Owens contends that we possess reflective control over belief only if his principle RM is true. But RM says that *all* reasons for belief must function motivationally. As Owens has argued, pragmatic factors are *prima facie* reasons for belief, but cannot function motivationally towards belief. A complete account of reflective control must address this concern.

theoretical judgment that is the outcome of reflection does have an important impact on such requirements. Coming to the theoretical judgment that she ought not believe that *p* makes an agent both aware of and answerable to a specific normative requirement, namely that she ought not believe (or continue believing) that *p*. Insofar as she is rational, this belief and associated normative requirement will determine her doxastic state, and she will cease believing that *p*.

Broome's distinctions allow us a certain further insight into the mechanisms of reflective control. In addition, they help us to see why Owens' initial claim, that judgments are not additional reasons for belief, may seem to have a certain intuitive force, though it is ultimately irrelevant. My thanks to Gary Watson and Andrews Reath for making me aware of the similarity between Broome's classifications and my own L2/L1 distinction.

Desperate Househusband in section 2.3, which involves the influence of pragmatic factors in determining Ned's belief, seems to be a counterexample to Owens' claim regarding the motivational irrelevance of pragmatic factors in theoretical reflection, and I am amenable to this interpretation of it. I recognize that this example will be controversial, however, and, since pragmatic control is not essential to my account of reflective control, I will not attempt to defend it further here. Moreover, I am certainly not prepared to defend the idea that *all* pragmatic factors can function motivationally in deliberation about belief, even in principle. Owens' example of how one determines whether or not it is raining seems quite apt on this point. One deliberates on the evidence in such situations, not on how much time one has to decide. And in the absence of evidence for or against the proposition that it is raining, it seems implausible to think that realizing that one has no more time to deliberate can move one to believe either in the proposition or its negation.⁷⁴

However, even if we accept that pragmatic factors cannot always function motivationally in belief, I submit that ought to lead us not to reject the idea of reflective control, but rather to reject RM as an appropriate principle of reflective control. I am convinced by Owens' and others' arguments that strict evidentialism is false, and thus

⁷⁴ This seems to me separate from the issue of whether, when I do have a certain amount of evidence in favor of the proposition that it is raining outside, pragmatic factors can function motivationally in moving me to believe or to withhold belief by raising or lowering the relevant evidential standard. This strikes me as closer to the example DH, but again, exploring this further moves too far from the central issue of this paper.

that pragmatic factors must play some role in the formation of belief.⁷⁵ But why think that their failure to rationally motivate belief undermines one's reflective control generally? Why not think instead that this failure shows us that there are simply some considerations relevant to belief (those that are evidential) that can function motivationally, and some (those that are pragmatic) that cannot, or at least cannot always?

This claim gains plausibility when we think about the goals of belief and action, respectively. While not wholly uncontroversial, it is plausible to hold that belief in some way aims at truth, or is governed by norms closely connected with truth.⁷⁶ Action, on the other hand, is often seen as aiming at some conception of the good, or what is best for the agent. When I ask whether p , I am inquiring as to p 's truth. When I ask whether I ought to Φ , I am asking whether Φ -ing is best for me, or good for me, or something closely related. I will not attempt to defend these claims here, but if this basic account is accepted, then we have a principled basis for explaining why pragmatic factors cannot generally function directly as motivation for belief.⁷⁷ Pragmatic factors do not generally

⁷⁵ The precise nature of this role is a challenging issue. But I think there are additional questions here regarding in what sense pragmatic factors are "reasons" for belief. The account that Owens' gives does seem to establish that pragmatic factors are necessary conditions for the formation of belief, but it still seems an additional question as to whether they are reasons for belief. One natural way, it seems to me, to determine whether some consideration is a reason for adopting some propositional attitude just *is* whether or not the consideration can function reflectively as a rational motivation for the attitude in question. I take Korsgaard to be suggesting something similar in Korsgaard (1996), p. 93-94.

⁷⁶ For an argument that belief is closely connected with truth, see Velleman (2000). I will discuss this claim at greater length in Chapter 3.

⁷⁷ Helm's (1994) claim is that pragmatic factors of this sort can function indirectly to determine beliefs.

speak directly to the aim or governing norms of belief, they offer us no reason to judge that a proposition is true, and thus can offer the prospective believer no assistance in satisfying these norms.

In the practical realm, on the other hand, pragmatic factors can be directly relevant to the goal of action, and the relevant norms governing one's decision. Consider a circumstance of deliberation about some action Φ . In such deliberation, there are reasons in favor of Φ -ing, reasons against Φ -ing (perhaps in favor of some other option, Ψ), and reasons which are relevant to my decision-making or the question of what to do, though they are neither reasons directly for nor against Φ -ing. The fact that I am running out of time typically falls into this final category. However, this consideration may be directly relevant to the question of what is best, or what I ought to do, particularly if the ability to Φ , or the associated benefits to Φ -ing, may become unavailable at a certain point. If the reasons for Φ -ing are generally superior to the reasons for Ψ -ing, but nevertheless not determinative on their own (they don't immediately move me to Φ), the fact that I will lose the ability to Φ because of the time constraints on my decision may be a sufficient reason to bring me to the judgment that Φ -ing is the best thing to do, all things considered, and to motivate me to form the relevant intention.

Given this explanation of why pragmatic factors are directly motivationally efficacious for action but not belief, we should reject principle RM as overly demanding for the intellectualist. The function of pragmatic factors does mark a difference between belief and action, but not one that undermines rational control over belief. Even though

pragmatic considerations cannot typically move me to judgment regarding the truth of p , evidential factors can, through the mechanism of reflection. And insofar as my judgment concerning these evidential considerations can be efficacious in determining my beliefs, the intellectualist should accept that I possess reflective control over my beliefs, just as I do over my actions.

2.9. Conclusion

I have argued in this chapter that affirming control over one's beliefs is a natural outgrowth of the acceptance of intellectualism. While the adjudication between intellectualism and voluntarism as competing conceptions of control will have to occur elsewhere, I have attempted to offer a plausible intellectualist account of control over belief, one that accords with our intuitions about various cases of belief, as well as our common judgments of responsibility for belief. The central claim of this account is that when one's beliefs are actually or potentially responsive to one's reflective judgments about what one ought to believe, one possesses control over these beliefs. I have argued both that this type of control is essential to normal human cognitive functioning, and that it may be expressed through various modalities, including at least negative and positive control. Finally, I have suggested that the reflective judgments at the center of this account exercise control through an internal (L2) notion of rationality connected with self-consistency and other requirements imposed upon a subject by her beliefs and other mental states, rather than through the external (L1) sense of rationality that concerns a

subject's connection with the world. In this last point we see one of the chief limitations of reflective control. Though it may offer us the means to govern and own our intellectual lives, regrettably reflective control does not guarantee us a grip on the true reasons for belief or action. Surely though, this is a general problem of human cognition and reasoning, rather than any limitation in the idea of reflective control or in intellectualism as a theory of how we exercise control over ourselves.

As I noted at the outset of this chapter, intellectualism, though an important minority view, stands somewhat apart from the dominant thread of philosophical discussion concerning control over belief, where voluntarism is assumed. Beginning with the next chapter I turn to the examination of voluntarism. I will seek to answer the same questions: if one adopts the voluntarist outlook, what implications does this hold for the issue of doxastic control? Is the voluntarist committed by her principles to supporting doxastic control, or are there distinctive characteristics of voluntarism that make doxastic control impossible or implausible? These discussions will occupy the remainder of this work.

CHAPTER 3 – MODERATE DIRECT CONTROL ON A VOLUNTARIST FRAMEWORK

3.1. Returning to the Will

In this chapter I will continue the argument for direct control. In the last chapter I introduced the distinction between intellectualism and voluntarism about control over both belief and action, and argued for direct control over belief determination on an intellectualist conception. The intellectualist suggests that control is effectuated through judgment, while the voluntarist argues that control consists in subjection to the will. It may seem that embracing the voluntarist view seems to pose a special challenge to the advocate of direct control over belief. This is because the voluntarist, given her characterization of the nature of control, must show that belief is subject to the dictates of the will. I will attempt to make the case for direct control on voluntarism in what follows. Many authors have argued against the idea of direct control from within a voluntarist conception; I will concern myself with responding to several such arguments in Chapters 4 and 5. But in this chapter I will focus on advancing the positive argument for moderate direct control within a voluntarist framework.

Recall that moderate direct control claims that the determination of beliefs is commonly subject to direct control in much the same way as actions are. In stating the thesis previously, I also noted that it leads to the following two corollaries:

- (1) Individuals' beliefs are responsive to reasons about what they ought to believe, in ways similar to how individual's actions are responsive to reasons about what they ought to do.
- (2) In certain circumstances, individuals have the ability to decide what to believe.

While I mentioned these corollaries in my previous statements of the thesis, their significance will, I hope, become more apparent in this chapter. They represent related ways of unpacking the thesis of moderate direct control that should be agreed to by those with a wide array of philosophical outlooks on control. In this chapter I will discuss both libertarian and compatibilist conceptions of control, and suggest that both perspectives can affirm these two statements, though they will imbue them with somewhat different significance.

To make the case for moderate direct control on voluntarism, I will pursue a two part strategy. I will first consider what I call the normative structure of belief. Many of the arguments against direct control over belief, I suggest, arise out of the mistaken view that action and belief have very different normative structures, meaning that they are governed by different norms and subject to these norms in very different ways. I will advance a conception of belief and action below that emphasizes the similarities in their normative structures. I will further argue that this view supports the claim that we possess control over the determination of our beliefs similar to the control we possess over the determination of our actions.

I will then spend the rest of the chapter examining in greater detail this specific control that I allege we possess over belief, from within a voluntarist framework. As

noted, the thesis of moderate direct control is open to different interpretations in part depending on one's conception of the metaphysics of control. Continuing my practice of not attempting to adjudicate these sorts of disputes, I will instead argue in this latter section that however one comes down on the metaphysics of control, one ought to acknowledge the reality of direct control over belief determination. John Martin Fischer employs a useful distinction in his influential discussions of the metaphysics of control, and I will appropriate it here for my purposes. Fischer distinguishes between two different conceptions of the control required for responsibility: *regulative control* and *guidance control*. Fischer's distinction is raised in discussion of action, but one of my main goals in this chapter will be showing its relevance for discussions of belief. Regulative control, the type of control often discussed by libertarian incompatibilists, requires access to genuine alternative possibilities. Guidance control, the notion of control advocated by Fischer and many other compatibilists,⁷⁸ does not require access to genuine alternative possibilities, and delineates control in other terms, which I will describe. My argument will take the same form regardless of the notion of control; I will suggest that if one accepts the existence of the respective type of direct control for action, one ought to accept it for belief as well. By examining the conditions of control and

⁷⁸ Strictly speaking, Fischer considers himself a "semi-compatibilist", because while he believes that moral responsibility and acting freely do not require access to alternative possibilities, free will in the traditional sense of the term does require such access. Compatibilists like David Lewis and Keith Lehrer, on the other hand, argue that both free will and control over actions are compatible with a lack of such possibilities. These aspects of Fischer's theory need not concern us here; my argument will, I believe, be capable of generalization to other broadly compatibilist theories of control.

responsibility established for action we can see that identical or appropriately amended conditions are satisfied by belief determination as well.

3.2. The Normative Structure of Belief Determination

I turn first to examining the normative structure of belief determination. A norm is of course a standard, a guideline or rule. To say then that a process like the determination of belief has a normative structure is to say that the process has one or more norms which govern its operation and which provide for a standard against which the outcome of the process may be measured. As I noted in the previous chapter, in the case of belief, the relevant governing norm is typically taken to be truth. It is often claimed, however, that the connection of belief with truth mitigates against the possibility of direct control over belief determination.⁷⁹ In this section I will briefly examine the relationship between belief and truth, and conclude that contrary to this popular view the normative structure of belief determination, including the connection between belief and truth, has no deleterious effect upon control. Rather, when properly understood, the normative structure of belief enhances the argument for moderate direct control, by showing us further parallels between belief and action.

Since we are accepting that truth is at least one of the governing norms of belief, examining our standard evaluative practices for belief will provide a useful window into

⁷⁹ Such a claim is the foundation of Adler's argument against control (see Adler (2002)), but Audi (2001), Alston (1988) and Williams (1973) all hold to variants of this general view.

this topic. Making the case that the normative structure of belief strongly parallels that of action will require building upon the distinction and discussion of L1 and L2 rationality from Chapter 2. I will expand upon the taxonomy of evaluation that I made there, though it will of necessity remain quite incomplete and provisional.

In the previous chapter I established that there are at least two distinct and separate forms of evaluation for both belief and action, and there are significant parallels between the forms of practical and theoretical evaluation. I wish to build on that discussion here. I begin with belief and evaluation at L1. Recall that the evaluation of belief at L1 is of the relationship between the belief formed and the external world to which it refers. Beliefs are correct at this stage of analysis if they are true. If a subject has the belief “There is a fire engine going by my house”, and in fact there is a fire engine going by the house, this belief will evaluate positively relative to the norm of truth operative at L1. If, on the other hand, the subject has simply imagined this state of affairs and there is no fire engine going by, the belief is incorrect with respect to truth, whatever else may be said for it.

In the realm of action, by contrast, truth is not the governing norm at L1, but rather some conception of *goodness*. To be clear, I have no desire to saddle my argument with any claim that all action aims at the Good, conceived as a univocal and objective source or object of positive value. I am here using goodness as something of a catch-all, a symbol for a family of considerations including goodness, benefit (for the subject or for others), value, etc. I am explicitly using the term in a subjective, internal sense, referring

to the perception or evaluation of the actor herself, rather than to any objective standard existing or standing apart from the actor's evaluation.⁸⁰ Whether any such standards of goodness exist independently is a metaphysical question quite outside the boundaries of this inquiry. My claim is simply that action, and in particular the deliberative action that is the focus of our discussion, aims at one or more of the cluster of related concepts noted above.⁸¹ I take such a claim to be at most minimally controversial.

While the conception of goodness may be subjective, however, the evaluation of individual actions with respect to it is not. Agents can fail to live up to even their own standards of goodness. If Stephen attempts to give quinine to his mother because he believes it will cure her smallpox and thus be beneficial to her, but it instead has no effect on the disease (being instead a treatment for malaria) and furthermore makes her deaf, then his action was not "good" in any meaningful sense of the term. It failed to live up to his own conception of goodness or any reasonable notion of benefit for his mother. So whether an action is in fact good is something subject to objective analysis, even if one accepts that the standard of goodness at which the action aims is itself subjective.

As the above discussion makes apparent, the norms of belief and action at L1 clearly differ. But it is not clear how this difference alone can ground the claim that we

⁸⁰ I do this because I take this to be the weaker claim, and all that is required to establish my point. I of course do not wish to preclude the possibility of objective notions of goodness, I simply don't wish to distract from my main point by defending a stronger claim that seems unnecessary for my purposes.

⁸¹ For an influential contrary view, see Velleman (1992b). Watson (1978) articulates a related conception, though he speaks in terms of valuational systems, rather than specifically of a concept of the Good.

lack control over belief determination while retaining it over action. One might argue that the truth of a belief is simply not up to us, it is simply a fact about the world over which we (typically) have no control. However, as I've noted, both the truth of a belief and the goodness of an action are dependent, at least in part, on external facts over which we have no control.⁸² Our control over whether an action is in fact good is similarly limited by these facts. It is not clear how the bare difference in norms points to a meaningful difference in terms of control between the practical and theoretical realms.

Still, the different norms of belief and action *does* mean that different types of reasons will be relevant for belief determination than for action. An alternative claim might be that this difference in the general content of belief-relevant reasons and action-relevant reasons is the basis for the lack of control over belief determination. An argument of this sort would have to claim that something about the content of belief-relevant reasons make control over belief determination impossible. As we saw in Chapter 2, however, and as we will see in the next section, control is not a principally a matter of content but rather of the structural relations between beliefs or actions and other mental states internal to the believing subject. For the intellectualist, control is a matter of appropriate structural relations between beliefs or actions and judgments that are the outcome of deliberation. The voluntarist adds in the complication of the will, but the essential issue remains structural, not a matter of content. While the reasons that matter

⁸² Self-fulfilling beliefs may be an exception to this general rule. I would argue that there are corresponding exceptions in the case of action as well, and of course the existence of self-fulfilling beliefs does nothing for the critic of control.

for belief and those that matter for action are different, this difference cannot ground any disanalogy between control for belief determination vs. action.

Of course, there are structural difference between belief and action at L1, as we've identified above. The truth of belief is determined by the relationship between mind and world. The goodness of action, on the other hand, is a matter of the relationship between the action, which is at least in part a feature of the external world,⁸³ and other aspects of the world. Focusing a bit more on these structural issues, however, helps us to see their irrelevance to control. Both truth and goodness are at least in part external relations, and it is precisely because of the externality of truth and goodness that considerations at L1 are irrelevant to control. Taking the relevant belief and action as givens, I may not (except in very special circumstances) be able to control whether my belief is in fact true, or whether my action is in fact good. But this has little bearing on whether or not I can control the belief or the action itself. As we saw in the previous chapter, control is an internal matter of relations between mental phenomena of the subject. It is at L2 that these relations are evaluated, and where any differences relevant to control will ultimately be found.

While not sufficient to ground a difference in control, at L1 there are nevertheless clearly different normative concepts at work in the theoretical and practical realms. At L2, however, the differences between the norms of belief and action become less clear.

⁸³ As Peter Graham has pointed out in conversation, part of the individuation of an actual may include its associated intention, which is most plausibly thought of as a mental state.

What becomes more clear, on the other hand, are the structural similarities in the determination of action and belief. This can be clearly seen by looking at examples of our evaluative practices related to L2 rationality.

Belief is inextricably connected with truth, but truth is not the *only* standard of evaluation for belief. To the contrary, we regularly make evaluations of beliefs at L2 contrary to or in the absence of information regarding their truth status. We often offer either negative or positive evaluations of beliefs whose truth we are unable to evaluate, simply based on the reasoning used to arrive at the belief. When Sally tells me that she believes her son who ran away from home a week earlier is fine because this is what the tarot card reader at the mall told her, I will likely negatively evaluate her belief, even if I have no idea whether or not she is correct (and despite the fact that I sincerely hope her belief proves to be true). My criticism will be based not on the truthfulness of the belief, but rather on the conviction that her belief cannot be adequately justified by the pronouncements of a tarot card reader, that this is not a legitimate basis upon which to form a belief.

Even more strikingly, we will make the same types of negative evaluations about beliefs which we know to be true. If Paul tells you that he won a great deal of money based on his belief that the Dodgers would win on Saturday, you likely will congratulate him on his successful wager, and perhaps even form a favorable view of his baseball evaluation skills. But when he informs you that his belief in the Dodgers' impending victory was based on the number of times his dog Tommy asked to be let outside on

Friday night (“odd, rather than even, clearly foreshadowing a good pitching performance by Chad Billingsley”, he says), your evaluation of his belief in the Dodgers’ victory will clearly sour, despite the fact you know that it proved to be correct. This is because you do not see Paul’s belief (and, it should be noted, his resulting bet) as appropriately formed, despite the fact that you know it turned out to be true.

In a similar manner, we will sometimes make positive evaluations of beliefs that we know to be false. Any parent of small children is familiar with praising one of his or her child’s beliefs, which, while false, displays excellent reasoning from premises which the child doesn’t know to be faulty. To avoid any pejorative connotations from associating the practice with small children, however, consider an example involving adults. Mitch and Steve, good friends from opposite sides of the political spectrum, are discussing healthcare reform. Mitch, a conservative Republican, may vehemently disagree with Steve’s belief that rationing of healthcare services is both necessary and appropriate as a part of any comprehensive reform package. While committed to the falsity of this claim, Mitch may nevertheless feel that Steve’s belief is commendable for its honesty and its rational consistency with his other beliefs about healthcare and social policy, though Mitch holds many of these beliefs to be false as well.⁸⁴

⁸⁴ The extreme example of such cases in the epistemological literature concern Demon Worlds or Brain in a Vat (BIV) scenarios. In such cases the subject in question is universally deceived, and all or nearly all of his/her beliefs that refer to the external world are false, though many of them are nevertheless justified. Part of the point of such examples is typically the same as I am attempting to make above, that justification comes apart and is separate from considerations of truth.

Clearly something more than simply the truthfulness of a belief is at issue in these examples.⁸⁵ It seems apparent that at L2 belief has considerations against which it is evaluated which are not connected exclusively with truth. The same is true of action, with respect to goodness. We often offer forms of positive evaluation to actions even when we don't think they are good or beneficial. When we hear that Maria has called the police on her husband, mistakenly thinking him to be an intruder in their house, we may be inclined to say things like "Well, it makes sense that she would do that, given that she thought there was an intruder in the house", or even "That was a perfectly reasonable reaction, since no one was supposed to be home." We offer a favorable, or at least a non-negative evaluation of Maria's actions, explicitly referencing her other internal doxastic and practical states as a basis for the evaluation.

The unifying thread in all these examples, theoretical and practical, is *justification*. In each case, the belief that passes or fails the L1 test of truth or goodness has an opposite result with the principal test at L2, the test of justification. As we saw in

⁸⁵ Now, one possible response to the examples above is to say that I have misread the significance of the evaluations in these situations. While we would negatively evaluate the beliefs in the cases of Sally & Paul above, the evaluations would nevertheless be made on the basis of considerations of truth, just in a more roundabout fashion. In both examples, the objector might suggest, our negative evaluation is based upon our conviction that the mechanisms or methods utilized to form the belief are faulty, because they are not truth conducive or reliable. Such belief forming methods, though they may happen to provide true beliefs in these cases, do not tend to lead to true beliefs, and thus are worthy of condemnation rather than praise. The principal question is still one of truth.

I am happy to embrace such an analysis, at least in part. My purpose is not to attempt to completely separate the concept of justification from considerations of truth, an implausible project at best. Whether or not the ground of our analysis of justification is some notion of truth conduciveness is a question far outside the scope of my concerns, but even if we accept this claim it has no bearing on my claim that individual analyses of the justifiedness of beliefs and actions are separate and distinct from questions of their truth.

Chapter 2 and above, for both belief and action evaluation at L2 is in terms of the relationship between the relevant mental state (belief or intention) and the subject's other mental states. Beliefs and action are evaluated with respect to the strength of the support provided by their supporting mental states. The structural similarity of justification for action and for belief is clearly evident. The issue of justification in each realm is best understood in terms of a uniform question, "why do/believe x ?" For both belief and action, the question is answered by the giving of reasons. What counts as a reason of course varies according to the relevant mental state being supported. In both cases, however, the structure is the same. The justificatory norm functions as a barrier to be crossed, a threshold, below which both belief and action do not merit the honorific "justified". Beliefs and actions adopted without sufficient reasons to meet the relevant standard of justification invite our negative evaluation and censure at L2.

It is worth noting in passing that L1 and L2 do not exhaust the possible levels of evaluation for belief or action. And consideration of additional levels of evaluation may offer further support for the argument of the similarity in the normative structures of belief and action. As I note above, the test of justification is a measurement of the support for a belief or intention against some standard. We may reasonably inquire about the uniformity and basis of the standard utilized. For example, in the realm of belief, William James has famously argued that one's approach to acquiring or managing one's beliefs may have various different connections with truth. This is a topic worthy of its own detailed inquiry; here I simply wish recall James' widely acknowledged point that the

maximization of truth and the avoidance of error are quite different (and sometimes incompatible) doxastic goals, which are pursued in different ways. Both, James argues, are legitimate goals for the believing agent and the doxastic attitudes formed under either policy are properly called beliefs. Yet they result in markedly different justificatory standards, as evinced by the quite different sets of beliefs that may result from the different approaches. Think, on the one hand, of the skeptic who questions everything and embraces no belief without sufficient support, and on the other of the enthusiastic inquirer who treats all or most propositions as innocent until proven guilty. For the former, truth is the exclusive aim of the doxastic enterprise, for the latter, truth is simply one of the limiting conditions upon inquiry.

Paul Helm suggests that such differences can be explained by differing *belief policies*, meaning different approaches to evidence and differing justificatory standards.⁸⁶ The skeptic, Helm suggests, employs a strongly verificationist belief policy, requiring significant support for any belief before adopting it. The enthusiastic inquirer, on the other hand, employs a falsificationist belief policy, where beliefs presented for consideration enjoy a comparatively warm reception and instead require some measure of criticism before they are rejected.

The notions above are easily translated into the practical realm, where individuals clearly employ varying standards for the justification of action. If such ideas are correct (which I am here simply positing, rather than defending), then it follows that the

⁸⁶ Helm (1994)

standards of justification utilized by individual believers and actors themselves require justification. What precisely would this justification be in terms of? Most probably, it would be in terms of the highest-level goals, projects and commitments of the agent. At this level, which we might call L3, evaluation of belief and action policies achieve an even greater unity than at L2, hinting at a unity of practical and theoretical reason that supports the notion of control put forward here. Pursuing such a line of thought is of course impractical here, but the possibilities are nevertheless intriguing.

Returning to the previous discussion of L2, in the case of both belief and action, the justificatory/explanatory question is structurally identical. Whatever the specific vector along which it is being measured, justification is a threshold for both belief and action prior to which either belief or action are inappropriate. What it means for a belief or action to be justified is that it bears the proper relationship to an appropriate set of reasons. Given these structural similarities, whether there is one or are many standards against which either belief or action may be measured seems immaterial to the question of whether we possess control. Moreover, as we will examine in detail in the next section, the possession of this relationship to reasons is not typically taken to inhibit control over action. Why should it be thought to do so for belief?

Based on these considerations, I argue that the issue of the normative structure of belief and belief's connection with truth poses no specific challenge to direct control over belief determination. Rather, the normative structure of belief shows us another way in which belief and action are relevantly analogous, and thus offers one more reason to

think that if we accept the reality of direct control over actions, we ought also to embrace direct control over belief as well.

3.3. Conceptions of Control: Guidance & Regulative

In my first chapter I argued for a presumption in favor of direct control over belief, and in Chapter 2 I suggested how this presumption might be theoretically supported on an intellectualist account. In my discussion of the dominant voluntarist outlook, I have argued just above for a strong similarity in the normative structure of belief and action, in support of the claim that we ought not accept control over the latter while denying it over the former. Thus far, however, I have not discussed at length what the direct control over belief which I allege might consist in on a voluntarist framework. If, as I argue in the Introduction, control is the central issue for questions of responsibility for believing, then a plausible account or analysis of control is key to establishing the reasonableness of much of our normal talk about belief. I will attempt to provide such an account below, employing the useful distinctions made by John Martin Fischer that I noted in the opening of this chapter. As I will try to show, my view is not necessarily reliant on Fischer's account of control; as I have argued throughout, my arguments with respect to control are generalizable to any plausible account of control over action. I utilize Fischer's distinctions below because he provides one credible and highly influential way of dividing up the intellectual landscape on issues of control and responsibility.

What separates actions from mere activities, happenings, or events? One plausible answer is that actions are done by agents, and are performed for reasons. This means that with respect to actions a distinctive form of explanation is available that is not available for “mere” events. We can ask both “Why did the rock fall onto the car?” and “Why did the captain land the plane in the field?” The former question may be explained in terms of physical forces, geological events, chemical reactions, etc. But in most cases we would not expect an explanation in terms of reasons or motives, because the rock’s falling is not an action, it is “merely” an event (unless of course the rock was pushed, pulled or otherwise manipulated by an agent). The latter case is different. Insofar as the landing of the plane is an action, we expect an explanation that is in significant part comprised of appeals to reasons or motives. Any explanation that did not include an account of the reasons or motives of the captain would be an answer to a different question, perhaps about an event (“why did the plane come down in the field?”), but not about the action in question.

Clearly any adequate account of control must include this rational component, the distinctive role of reasons in actions that are under our control. However, many philosophers and non-philosophers alike would suggest that the rational component of actions, while necessary, is not sufficient to explain what it means for us to possess control. An extended consideration of the above example will help to clarify. Say that we were offered a seemingly complete explanation of the captain’s landing the plane in the field, with appeals to her reasoning, (“she judged that it was unlikely that she would

make it to the runway,") observations, ("she could tell that the field was clear of trees and other large obstructions", etc.), as well as other relevant factors. But suppose it was then added that the captain of the plane was in fact a complex robot, or a highly advanced experimental autopilot system. Many of us would at this point feel that the situation was different than we had originally surmised, and that in fact we no longer had an explanation of action of the sort in the original example. This can be seen from the fact that many of us would now find it inappropriate to praise the "captain" for the successful landing of the plane. Indeed, we would be much more likely to praise the human creators of the system, or to criticize them if it failed, than to attribute responsibility to the system itself.

Why would we find it difficult to attribute responsibility, and thus praise or blame, to the "captain" in this latter example? Of course, one might argue that some condition of rational control is not satisfied, that the robot could not possess the necessary mental states, etc. With enough reworking of the example I think we might assuage such worries, if they exist. But pursuing this, I suspect, takes us off track. The most straightforward explanation for why many would hesitate to praise or blame in such a circumstance is that they think that some condition of control quite apart from rationality has not been satisfied. In the example in question, many would argue, the robot or autopilot was not in control because, in the end, the outcome was not up to the robot.

But in what way is the outcome not "up to" the robot, not within its control? Many who hold this position would suggest that the robot lacked sufficient control over

the outcome because it was not in control of its decision to land the plane in the field. Certainly, the navigational and maneuvering equipment of the plane was subject to the robot's decision-making abilities, as it were. But the decisions made by the robot, it may be said, did not originate within the robot itself. They were simply the deterministic product of the robot's pre-programmed software and the inputs of the surrounding environment. Given the circumstances of the example, the robot could not do otherwise than it did. Precisely simulate or recreate the exact circumstances of the example, and the robot would do precisely as it did, every time, without fail. The robot lacks real access to alternative possibilities, it cannot do otherwise, and thus, many would argue, it lacks control over the landing of the plane, and is not the proper subject of praise or blame for the outcome, whether positive or negative.⁸⁷

The broad outlines hastily sketched above, and the two aspects of control contained therein, are given a much fuller exposition in the work of John Martin Fischer. Fischer dubs the notion of control focused on the agent's rational faculties *guidance control*, and he advances a vigorous defense of it across numerous well known articles

⁸⁷ In the above paragraph I paper over and otherwise ignore current debates in philosophy of action concerning the proper grounding of responsibility on incompatibilist accounts of action. Many incompatibilists would argue that the proper ground of control and thus responsibility is not mere access to alternative possibilities, but rather "sourcehood", the ability to be the ultimate originator or source of one's actions. McKenna (2001) defends a version of source incompatibilism, as does Pereboom (2003). Kane (2000) embraces the a sourcehood or ultimacy condition on control & responsibility, though he also requires alternative possibilities somewhere in the agent's causal history. Widerker (2002) rejects the idea of a sourcehood condition independent of the Principle of Alternative Possibilities (PAP). The proper account of sourcehood, and its relationship to PAP are contentious subjects in current action theory, and I will for my purposes (I think safely) ignore them here.

and books.⁸⁸ The notion of control that adds the additional requirement of access to alternative possibilities Fischer calls *regulative control*. These two conceptions of control will provide the framework for my discussion in the remainder of this chapter. I will first summarize the main features of each of these types of control. I will follow this with a discussion of the applicability of each conception of control to belief. My argument will follow a familiar pattern: if one accepts the plausibility of a given account of control for action, one ought to recognize its concomitant applicability to belief determination.

As I've noted, on Fischer's account guidance control is fundamentally rational control.⁸⁹ It consists of two elements, *moderate reasons-responsiveness*, and *mechanism ownership*. Moderate reasons-responsiveness is itself analyzable into two parts: *reasons-recognition* and *reasons-reactivity*. The former is the ability to recognize salient reasons when they are available. The second is the ability to respond to such reasons when recognized. All of these notions are explained by Fischer in terms of a *mechanism*. Much ink has been spilled over the precise characterization of a mechanism, but for our purposes, it can remain substantially vague. It is enough to say that a mechanism is a system or process by which an agent comes to perform an action. For example, the activity of my practical reason in the production of an action can be a mechanism in

⁸⁸ See Fischer (2006) and Fischer & Ravizza (1998) for extended discussions of guidance and regulative control.

⁸⁹ By "rational control" here I do not mean control operating successfully in accordance with accepted norms of rationality. Someone can be deeply irrational (at least L1 irrational) and nevertheless be in control of their actions or beliefs. I mean instead the broader notion of control effectuated through the characteristic action of a subject's rational faculties, control exercised through the use of reason (whether the faculties are employed well or poorly).

Fischer's sense. Say that in circumstances C I recognize reasons R as supporting action Φ , and thereby form an intention to Φ , and then Φ . We could say that my mechanism was moderately reasons responsive if in circumstances C*, where I was instead presented with reasons R*, which support action Ψ , I were to recognize these reasons, and react appropriately, instead forming an intention to Ψ .

This leaves mechanism ownership. Mechanism ownership has both a historical component, related to how a mechanism came to be present within an agent, as well as a psychological component, concerned with the agent's understanding of her mechanism, whether she takes ownership of it. In order for mechanism ownership to be present, the mechanism must have been acquired in a "normal" way, say through learning, development and growth, rather than through manipulation, such as hypnosis or direct interference with an agent's brain. There are three discrete psychological components to mechanism ownership. To possess ownership over a mechanism, I must see myself as an agent, as a source of decision and action with causal influence on the world. I must additionally see myself as properly subject to the reactive attitudes in my use of the mechanism. Finally, I must possess the first two attitudes or dispositions on the basis of evidence, meaning that there must be some appropriate connection between my attitudes and reality. Fischer is clear that these psychological aspects of mechanism ownership do not require conscious deliberation or reflection on the part of the agent. They are conditions or realities which an agent in principal has access to, but may never have deliberated about. One need not be a philosopher to have mechanism ownership. A

normal youth who has taken ownership of her practical reason, who operates as an agent in the world and responds within appropriate ranges to attributions of responsibility, may nevertheless never have reflected in an abstract sense upon her employment of her practical reason and the fairness of our social practices of attributing praise and blame. Nevertheless, she may still satisfy both of the conditions of guidance control in her actions, and be properly held responsible for them.

Let us turn now to regulative control. One of Fischer's standard characterizations of regulative control is as a kind of "guidance control plus."⁹⁰ Think back to my attempt at an intuitive characterization at the beginning of this section. One might be persuaded of the necessity of guidance control for proper attributions of responsibility; nevertheless, one might maintain that guidance control is not enough. The distinctive feature of regulative control is the requirement of access to genuine alternative metaphysical possibilities. I'm not in control, it's not up to me, unless I could have done otherwise than I did. I may do Φ , but in order to have regulative control (and thus, on such an account, to really be responsible) I must also have been able to do Ψ , holding the past prior to my action fixed. It is thus a requirement of advocates of regulative control that causal determinism –understood as the doctrine that at any point in the history of the universe the entirety of the future is entailed simply from a complete description of the state of the universe at that moment and the physical laws of the universe– is false.⁹¹ In addition to

⁹⁰ See, for example, Fischer & Ravizza (1998), pp. 31-34.

⁹¹ Causal determinism is characterized in this way by various interlocutors in the free will debate, here I take the characterization from Fischer & Ravizza (1998), p. 14. Fischer has noted in conversation that he is

the falsity of causal determinism, however, any plausible account of regulative control must provide a characterization of how agents possess access to the alternative possibilities in question. The will, characterized as the faculty of decision making and/or intention formation, is often understood to be the source of this access. Libertarians who accept something like a regulative account of control, often view the will as a power, which accords to the agent the ability in certain circumstances to decide on any of two or more possible actions, and form the intention to execute the chosen action. The precise conditions under which the agent is able to engage this power and under which alternative possibilities are accessible to an agent varies depending on the account.⁹² But all such accounts make a requirement of direct control that the agent could have done otherwise than she did.

We now have an exceedingly brief sketch of both guidance and regulative control as they pertain to action. I will now turn to considering the implications of each of these explanations of control for belief determination. Again, my purpose here is not to argue either for or against either of these conceptions of control. Clearly each is influential and has a measure of persuasive power. My goal is instead to argue that whichever of these two general accounts one accepts, one ought to accept an analogous account of control over belief determination.

officially agnostic as to whether causal determinism rules out alternative possibilities, though this is the view of the various libertarian incompatibilist theories that I discuss under the rubric of regulative control.

⁹² Plausible accounts of regulative control will also describe conditions of indirect control and tracing requirements for properly attributing responsibility for those many actions of which we do not enjoy direct regulative control. Again, these details can safely be left aside for our purposes.

3.4. Guidance Control Over Belief Determination

Consider first guidance control. I noted above that guidance control has two principal elements: moderate reasons-responsiveness and mechanism ownership, both of which can be broken down into discrete parts. For reasons-responsiveness, these parts are reasons-recognition and reasons-reactivity. It should be immediately clear, given our discussion in this chapter and the various examples in Chapter 1, that both of these requirements of guidance control are regularly satisfied for belief. Keeping in mind our focus on paradigmatic cases of deliberately formed beliefs and actions, these requirements amount to my ability to recognize reasons for a belief present in a given set of circumstances, and to respond to them. Say that in circumstances B I am presented with reasons R in favor of proposition P, and I form a belief in P. Then we may say that the mechanism of my theoretical reason is possibly moderately reasons responsive if, when I am in circumstances B* and I am presented with reasons R* in favor of proposition Q, I recognize the relevant reasons and form a belief in Q.

It seems reasonable to think that this characterization of moderate reasons-responsiveness for belief is applicable to us much of the time. Stacey, for example, is curious about the origin of specific species of plant found near her home, the reed canarygrass. She browses to Google on her computer and types in the name of the plant, and then the name of her hometown. Putting stock (as the vast majority of us do) in the relevance & wisdom of Google's ranking algorithms, Stacey selects the first link in the

result set presented to her. At the linked page, Stacey reads an article from a prominent botanist arguing that the presence of the reed canarygrass in her region is due to 19th century Dutch settlers of the region, who commonly kept a variety of the plant in their homes for its supposed medicinal qualities. After reading the article, Stacey considers the arguments made in favor of the theory. Finding them plausible, and having made a cursory search into the background of the botanist in question, she forms a belief in this theory as the proper explanation of the reed canarygrass' origin and presence in her hometown.

Now, imagine instead an alternative scenario (our B*), in which Stacey types into the Google webpage *first* the name of her town, *then* the name of the reed canarygrass. This subtle difference leads to a slightly different ranking of pages in the result set. Again taking the first link in her results, Stacey in this case ends up reading a well-documented paper by a top plant pathologist arguing that the reed canarygrass found in this region is principally the product of a variant introduced into the neighboring region by 16th century Spanish explorers, who brought the seeds of the canarygrass as part of the feed for their horses & livestock. The key link, the pathologist argues, is established by genetic testing of both the local plants and preserved seeds found in historical museums in southern Spain, near Sevilla. After reading this article, Stacey considers the arguments offered for this alternative theory (though, of course, to her in this case it isn't "alternative" to anything). Finding these arguments plausible, Stacey forms a belief in this theory as the proper explanation of the origin of the local reed canarygrass.

Clearly in the above scenarios Stacey's theoretical reason is moderately reasons responsive. Her mechanism both recognizes and responds to the reasons available to her in both scenarios. Moreover, I would argue that both of these scenarios are highly plausible analogues to common cases of deliberately determined actions normally thought to be under our direct control. We can easily imagine closely similar scenarios where Stacey goes through an almost identical process, not in order to determine what to believe about the origin of the reed canarygrass, but rather to decide on the most appropriate method of eliminating the grass from her backyard. There is nothing particularly special or unique about these examples, they are typical employments of our theoretical reason in the determination of belief.

If then our belief-forming mechanisms can be moderately reasons responsive, can we also possess appropriate ownership over them? Recall that Fischer's concept of mechanism ownership has both a historical component and psychological components. The historical development of our belief-forming mechanisms closely matches that of our action-producing mechanisms. Both are refined over time through the instruction of our parents and peers, as well as our observation of the beliefs and belief-forming practices of those around us. This formation process for belief includes the same kind of normative instruction, the application of blame ("you shouldn't believe everything a salesperson says, honey") and praise ("that's a good argument. I can see now why you think that.") that accompanies the development of our practical reason. The historical requirement for mechanism ownership seems no great challenge to direct control over belief.

Fischer's psychological requirements for mechanism ownership are perhaps a slightly different story. Again, the three requirements include seeing oneself as an agent whose decisions have causal influence in the world, seeing oneself as properly subject to the reactive attitudes for these decisions and actions, and possessing these two attitudes or dispositions on the basis of appropriate evidence. It is the first of these the three requirements that poses something of a challenge for the account of guidance control over belief. This is because it seems clear that our beliefs in fact do not have the right sort of causal influence on the world. This type of causal influence is the province of action, not belief. How then can we have the required attitude? Here, I would suggest, a mild revision of this condition for mechanism ownership is both required and appropriate, given the type of phenomenon under discussion. As I've noted throughout, there is no perfect analogy between belief and action. Our decisions concerning beliefs do not have this form of causal influence on the "world", understood as the totality of our experience outside of ourselves. However, they do have such influence on our internal lives, and particularly our mental states, which, as I have been arguing, they directly determine. So, the appropriately revised condition for guidance control over belief demands that for mechanism ownership, subjects see themselves as doxastic agents whose doxastic decisions have influence on their internal mental states, determining their beliefs and influencing other mental states.

Normal adult believers typically recognize this influence, just as they recognize the influence of their practical decisions on the world. When I consider my belief that a

certain restaurant has the best Thai food in town, I do not typically wonder in bewilderment where in the world the belief has come from, or see it as something alien or external, imposed on me from the outside. I recognize that it is rooted in an implicit or explicit reasoning process that takes as its inputs things like my experiences eating at the restaurant, my reading of reviews in the local paper, conversations with colleagues and friends, etc. In fact, our standard practice of deliberation in both theoretical and practical contexts clearly indicates our recognition of the efficaciousness of such an activity. “Thinking it over”, in both the theoretical and practical case, is a goal-oriented activity aimed at resolving a question (what to think, or what to do), and is predicated on our conviction that the outcome of the process, namely the relevant decision, will have the desired influence or impact of determining the appropriate mental state of belief or intention. Recall again that Fischer’s conditions are not requirements that a subject must have consciously reflected upon, but are rather simply attitudes which must in principle be available to the subject in question. Surely this is the case with respect to belief for most normal adult believers. We move through the world, deliberating and determining our beliefs with just such a self-conception.

Much less effort is required to establish the satisfaction of Fischer’s other two psychological requirements for mechanism ownership. The entire discussion in Chapter 1 and indeed, the main premise for this book are both predicated on our common social practice of invoking the reactive attitudes towards each other in regards to our beliefs. And it practically goes without saying that normal adult believers who engage in this

practice take themselves to be properly subject to these sorts of attitudes, to praise and blame in relation to a broad subset of their beliefs. If I know that you take my belief on some controversial issue (say, the question of whether abortion should be legal) to be poorly founded, I am not at all surprised if you voice your opinion by criticizing my belief and speaking contemptuously of my reasoning or other belief-forming practices. If I find merit in your critical stance, I may even revise my belief in line with the criticisms offered. However, even if I don't agree with you, I will not find it odd, inappropriate or unfair that you would criticize me in this way, given your convictions. I am in fact likely to offer the same form of criticisms, and to direct the same negative attitudes towards you for your belief that I see as shakily supported, or perhaps simply deeply morally flawed. Fischer's third condition is much the same. Again, if my discussion in Chapter 1 has merit, our default view of ourselves includes a conception of doxastic agency, and our standard social practices indicate the normalcy and fairness of praising and blaming. This is surely sufficient evidence to ground the self conception that I have argued for just above.

It is therefore plausible to claim that all of the conditions for guidance control are frequently satisfied with respect to belief. We possess both moderate reasons-responsiveness and mechanism ownership with regards to at least a large number of our deliberately formed beliefs. At least a large number of our belief determinations are responsive to reasons in much the same way as our actions are. Working from the assumption that guidance control is the correct account of the control necessary for

responsibility, we may therefore conclude both that we possess direct control over a large subsection of our beliefs, and thus that the responsibility and control statements that we regularly make regarding such beliefs are justified and appropriate.

Fischer's account of guidance control, which I have focused on, is one of the most influential compatibilist theories of action and responsibility in current philosophical literature. However, while it would be impossible here to do a comprehensive survey of compatibilist theories, I would like to at least gesture at how one might make my argument more general. The points that I have made here can, I believe, be generalized to all or nearly all compatibilist accounts of action. While voluntarist compatibilists make room in their accounts for the will, the focus of control is typically upon the role of our reason in our actions. Given their rejection of the regulative or alternative possibilities condition on control, one of the distinctive elements of the will is eliminated. Since the will is no longer the faculty by which I may access the alternative possibilities which the libertarian typically demands, it becomes instead principally a mechanism for effectuating my judgments. But given the kinds of similarities noted above in the function of reason in the production of action and belief, it seems plausible to think that compatibilist accounts of action will also be satisfied for belief. When the focus is on the function of reason in control, as we've seen above and in Chapter 2, it is difficult to find a meaningful distinction to ground the claim that we possess control over action, but not over belief. Of course, any such generalized claims I make are little more than informed conjecture at this point, and would require substantial argumentation and the examination

of the details of individual compatibilist theories to establish. Compatibilists will undoubtedly have a variety of objections to such a thesis. I will touch on one related objection in the final section of this chapter, and address a number of objections arising out of compatibilist theories of action in Chapters 4 & 5. Here, however, I will put aside the discussion of guidance control and compatibilist accounts and turn to the examination of regulative control.

3.5. Regulative Control Over Belief Determination

We have established the plausibility of direct control over belief determination given Fischer's compatibilism, with guidance control as the criterion of control sufficient for responsibility. I now turn to the case for regulative control over belief. I have characterized regulative control above as "guidance control plus". To the rational control required by guidance control is added access to genuine alternative possibilities; the ability (holding the past fixed) to choose alternatively to the option decided upon in the actual sequence of events.⁹³ This type of control implies that while I decided to go to the store, I could have (without changing the past or the laws of nature) decided to stay home and finish my book. On the theoretical side, I decided to believe John's story, but I could have chosen not to. I chose to embrace the theory of Leif Erickson's discovery of the Americas after reading a relevant text, but I could have withheld my assent and engaged

⁹³ For a similar general account of libertarianism from a sympathetic viewpoint, see Clarke (2003), pp, 15-17.

in further inquiry. Is it credible to think that we possess such control over belief determination?

I have already tried to make the intuitive case in Chapter 1, appealing to a variety of plausible examples and our common linguistic practices. Here again I must attempt to provide a plausible theory to explain these cases. In doing so, however, it is crucial to recognize the specific dialectical situation. As I have said throughout, my contention is that *if* one accepts a particular form of direct control over action, one ought to also acknowledge its efficacy for belief determination. The targets of the argument in this section, then, are those who embrace a libertarian incompatibilist account of control over action, and the relevant question is whether there is any reason to think that such control, if present for action, is not available for belief determination. I need not and will not burden myself with defending regulative control as the true account of the necessary conditions of moral responsibility, or defending libertarian incompatibilism more broadly.

While in the minority amongst contemporary philosophers of action, a variety of libertarian accounts exist. They can be broadly categorized into three groups, all of which qualify as libertarian in virtue of their acceptance of the alternative possibilities condition on control:⁹⁴

⁹⁴ This general division of libertarian theories can be found in Kane (2005), Clark (2003) and O'Connor (2000). Clark refers to the first set of theories as "noncausal", while Kane follows O'Connor in characterizing them as "simple indeterminism".

Noncausal libertarianism/simple indeterminism – Free actions are free in virtue of one or more necessary noncausal features they possess, connected with the agent’s reasons, decisions, or intentions.⁹⁵

Event-causal libertarianism – Free actions performed by agents are free in virtue of distinctive indeterminism in the event-causal sequence leading to the action.

Agent-causal libertarianism – Free actions are caused by agents who are substances via agent causation, a unique form of causation not reducible to any form of event causation, and itself not subject to causal determination.

Many philosophers, of course, regard the libertarian incompatibilist account of control as mysterious in any of its forms. These philosophers argue that none of the various arguments offered in favor of regulative control are successful in their aim. Such arguments, they say, are either irrelevant (such as event-causal libertarians’ insertion of indeterminism in the generation of action), or mysterious, ad hoc, and metaphysically implausible (such as agent-causal libertarians’ postulation of the agent-causal power). I will not attempt here to rehearse or refute these well known arguments within philosophy of action. This is a debate I will not enter into, since I need not undertake the task of attempting to convince such philosophers that these forms of regulative control are intelligible or plausible.

Instead, I will focus on the narrow question of whether the additional alternative possibilities condition of regulative control is plausibly satisfied for belief determination. If we possess access to genuine metaphysical alternative possibilities for some of our actions, can we also say that we possess similar access with respect to some of our

⁹⁵ Carl Ginet and Hugh McCann are commonly cited as principal examples of noncausal libertarian theory. See Kane (2005), pp. 53-57, and Clark (2003), pp. 17-24.

beliefs? The mechanism or vehicle for this access for the voluntarist is, broadly speaking, the will, the faculty of decision. But how precisely is this supposed to work?

As I noted above, the varieties of libertarianism are numerous, and to keep my argument as general as possible I will not focus on the details of a specific variant. As I hope to show, this is not necessary for my purposes. Instead, to clarify and focus the discussion I will offer a broad characterization of deliberative action (our paradigmatic type) that is consistent with (a) my previous discussions of action (b) Fischer's account of regulative control, and (c) with many (if not all) of the specific accounts of libertarianism offered by the defenders of the varieties noted above. On this account, an agent is in direct control of an action Φ when each of the following is true:⁹⁶

- (1) The agent in circumstances X deliberates about whether or not to Φ .
- (2) The agent in X terminates her deliberation with a judgment that she ought to Φ /that Φ -ing is best/the available reasons support Φ -ing.
- (3) The agent makes a decision to Φ , resulting in an intention to Φ .
- (4) The agent non-deviantly Φ 's under the governance of her intention.⁹⁷
- (5) The agent could have chosen not to Φ , or to perform some alternative action Ψ , holding the past prior to 1-4 fixed.⁹⁸

⁹⁶ Note that this is offered as a set of sufficient conditions for action subject to direct control, not a set of necessary conditions. One or more of these items may not be necessary for freedom and direct control (for example, all free actions do not require prior deliberation). This is simply meant to be a generalized description of a paradigmatic instance of direct control.

⁹⁷ The "non-deviantly" qualifier is meant to avoid the type of causal-deviance objections raised by Davidson (1973) and others. As I discuss in Chapter 2, section 3.6 and further in Chapter 4, determining the appropriate characterization of intention in the production of action & belief is complex, and I do not mean to endorse any particular account of intention here. However, as I make clear in section 3.6, I think any overly robust notion of intention is problematic for both belief and action.

This general account is distinctively libertarian because of (5), but it intentionally leaves open the question of how precisely (5) is satisfied, that is, by virtue of what the agent could in fact access the alternative possibilities in question. I will briefly note some of the possibilities explanations below, but as I hope will become clear, my argument regarding the satisfaction of libertarian or regulative control does not depend on the details of (5).

Examples will be useful to connect this account with our experience, as well as to unpack its implications. Take first a characteristic case of deliberative action. Say that I am trying to decide whether to stop at Starbucks on the way to work in the morning. As I drive, I deliberate about various considerations. I love to get a tea in the morning, but stopping will almost certainly make me late for work. Of course, being a few minutes late to my job is not a big deal (after all, I work at a university), and I'll likely be more functional in the morning hours if I am mildly caffeinated. But I've been to Starbucks almost every day this week, and I really can't afford to spend so much money on breakfast drinks on a monthly basis. I deliberate for a minute or two, and finally conclude my deliberation with a judgment. Perhaps my judgment is that I ought not stop for a drink. Or perhaps my judgment is that it's not clear to me what is best, what I ought to do. Either way, on this conception, the path to action is not complete. The activity of my reason culminating in judgment still leaves room for the will, the faculty of decision.

⁹⁸ Some versions of libertarianism would be comfortable with a variant of this claim that holds the past fixed up to 2 or 3 of the above sequence, but this version of the statement has the virtue of not excluding certain versions of event-causal libertarianism that attempt to place the relevant indeterminism relatively far back in the action determination process (see Dennett (1978), cited in Kane (2005), pp. 64-65).

Even if I have made a judgment about what I *ought* to do, I still must decide what I am *going* to do. In the circumstances described, where the considerations are relatively evenly balanced (and the decision perhaps somewhat trivial) and my judgment either only moderately inclines me to one option, or leaves me in relative equipoise, the advocate of regulative control will argue that I retain the ability to decide in favor of either course of action. Both options remain open to me; I can either decide to go to the coffee shop, or to go directly into work. I will of course eventually make a decision and follow some course of action, but this does not change the fact, according to the libertarian, that whatever I decide in this circumstance, I could have done otherwise. And it is this fact about my decision that satisfies point (5) of my above account and which makes the action truly under my direct control.

Now consider a similar case of belief. I am watching a History Channel special on the life of Mikhail Gorbachev.⁹⁹ Gorbachev is someone I deeply admire as a person, in part for his role in ending the Cold War and introducing democratic reforms into the former Soviet Union. The biographical feature highlights these notable achievements, but also spends a significant time discussing Gorbachev's personal life in less than glowing terms. Using interviews with one of Gorbachev's sons and a former close friend, it paints a fairly compelling picture of the former Soviet General Secretary as an indifferent or even abusive father and a closet alcoholic. Rather befuddled by the mixed portrayal, and

⁹⁹ While of course Gorbachev is a historical figure, please note that the details of this example are made up out of whole cloth, and have no connection with reality that I am aware of.

saddened by the apparent personal failings of someone I admire, I decide to do a bit more research on Gorbachev's life. In short order, I find websites that dispute the claims made by the documentary, arguing that Gorbachev's son is a prodigal greedy for his father's personal fortune, and that the alleged friend from the documentary is a disgruntled former Communist official deeply embittered at Gorbachev over his own downfall in the post-Communist era. On the other hand, I also find websites supporting the documentary's perspective. I reflect for a time on the various pieces of information I've encountered, as well as what I know of their sources. I am inclined to believe the compelling testimonies from the documentary, but have many questions about the quality of the son and "friend" as sources. I judge that the evidence as it stands perhaps narrowly favors the portrayal of the documentary, that Gorbachev's character is deeply flawed. Yet, here again, this judgment does not decide the issue of what I believe. While I see the evidence I have access to at present barely favoring the negative portrayal, I may decide to retain my present belief in Gorbachev's admirable character. This decision is an act of the will. Like the practical case, this decision could go either way, and, as in the practical case, for the adherent of regulative control, it is this essential liberty that satisfies (5) and makes the belief subject to my direct control.

Both the practical and theoretical example comply with my account of direct control above, as well as Fischer's requirements of regulative control. Moreover, there is nothing unique or uncommon above the examples. Similar cases abound in our daily experience, and it would be simple to produce a variety of both practical and theoretical

examples in the same vein. This being the case, we may say that belief determination is sometimes subject to our control in much the same way as our actions are, and, as the corollaries to the thesis of moderate direct control suggest, that we some are sometimes able to decide what to believe. Moderate direct control is plausible, given libertarian voluntarism.

It may be objected that I still have not provided an explanation of *how* it is that I could have decided in favor of *or* against my belief in Gorbachev, or, for that matter, how I could have decided for *or* against stopping to get tea, all while holding the past fixed. I have appealed to the will, but this in itself is not a complete explanation. Two things are relevant in response. First, the major forms of libertarianism all do attempt to provide an answer to the “how” question, and I discuss two of them below. But second and more important is recognition of the fact that in this context no further response is required to the objection.

The general claim of the event causal libertarian is that indeterminism introduced at some point in the deliberative process makes the variability in outcome possible.

Randolph Clark outlines a general version of this theory:

Let us start with a sketch of a rather simple view of this (event-causal) type. It employs an event-causal theory of action. And it imposes, for free action, the very same requirements as do many good compatibilists accounts (for compatibilists accounts, at least in recent times, do not typically require determinism). It differs from compatibilists views primarily just by also requiring, in order for action to be directly free, that certain agent-involving events (such as the agent having certain beliefs and desires and a certain intention) that cause the action must nondeterministically cause it. When the requirements of this account are

satisfied, when an agent acts...., it will have been open to her not to do what, in fact, she did just then.¹⁰⁰

Depending on the version of event-causal libertarianism, the indeterminism in question may be found at various points in the deliberative process, including judgment and decision, leading up to action.¹⁰¹ But whatever its precise location, the claim of the event-causal libertarian is that, because of this indeterminism, until the agent's decision to Φ is made there is a non-zero probability that she will instead choose to $\sim\Phi$, or perhaps to Ψ . Different event-causal libertarians fill out the description of how the requisite indeterminism functions through the agent's deliberative faculties to actually make other possibilities accessible, but all such theories share the commitment to (1) the presence of indeterminism in the action determination process as the basis of access to alternative possibilities, and (2) the sufficiency of event-causal descriptions to capture the activity of the agent in the production of the action.

The agent causal libertarian's claim, on the other hand, grounds the answer to the "how" question not principally in the causal indeterminism in the action-generating process, but the nature of the agent herself. On this type of account it is the agent causal power, operating through the agent's decision, which makes various possibilities accessible. The agent on this view is typically viewed as an enduring substance and her causal action as agent is not reducible to instances of event causation. Timothy

¹⁰⁰ Clarke (2003), p. 29. The initial parentheses are mine.

¹⁰¹ See Clarke (2003), pp. 57-92 for a variety of different accounts that locate the indeterminism in different points of the action-generation process.

O'Connor, a prominent agent-causal theorist, describes the agent-causal power in the following way:

The direct causing by agents of states of intention goes like this: parallel to event causes, the distinctive capacities of agent causes ('active powers') are grounded in a property or set of properties. So any agent having the relevant internal properties will *have it directly within his power* to cause any of a range of states of intention delimited by internal and external circumstances. However, these properties function differently in the associated causal process. Instead of being associated with 'functions from circumstances to effect', they (in conjunction with appropriate circumstances) make possible the agent's producing an effect. These choice-enabling properties ground a different type of causal power or capacity – one that in suitable circumstances is freely exercised by the agent himself.¹⁰²

An agent, therefore, while constrained by the circumstances in which she finds herself, is nevertheless sometimes able to directly bring about any one of a number of decisions, to access the multiple alternative possibilities required by (5) in the account described above. This capacity, as O'Connor notes, is grounded in properties of the agent, and allows the agent to sometimes freely (meaning without prior causal determination) bring about a particular outcome. This distinctive freedom of the agent makes it possible for her to choose either to Φ or to $\sim \Phi$, while holding the past fixed, as is demanded by libertarians as a necessary condition of direct control.

Potential answers to the "how" question of libertarianism are thus available, though much more would have to be (and has been) said to attempt to convince the compatibilist. But the more fundamental response in the current dialectical situation is to

¹⁰² O'Connor (2000), p. 72.

reject the demand for further information. Recall again that I am here addressing the libertarian, rather than the compatibilist, regarding the plausibility of direct control over belief. Postulating the indeterminism of the deliberative and decision-making process, and/or the agent cause as a potentially counter-normative power that gives us access to multiple alternative possibilities, is unsurprising and should be uncontroversial to the libertarian. In fact, it is important to the libertarian position that such possibilities exist. Notice further that the two principal explanations of the “how” of regulative control, mentioned briefly above, contain nothing that would limit their application to circumstances of practical deliberation and the determination of action as opposed to theoretical deliberation and the determination of belief. Nothing about the indeterminacy of the event flow that marks decision, or the properties with which the agent is endowed, prevents the relevant control over decision from extending beyond practical decisions to theoretical. The general account that I have described and examples I have offered is clearly compatible with either of the principal accounts, event or agent-causal. As the objection implies, either explanation, for belief or action, may be wholly dissatisfying (and wholly unnecessary) to the compatibilist advocate of guidance control, but she is not the one we are trying to convince. There is no need to further defend the details of regulative control to those who have already accepted it, in one form or another.

To be clear, to say that we possess such control for some of our beliefs is not to say that we possess it for all, or even most of our beliefs. It may be the case that many of my beliefs are not subject to this kind of control, either because they arise through a

different mechanism than deliberation, or because the evidence in favor or against the belief is so compelling as to remove even the possibility of a different conclusion. However, as I will discuss in subsequent chapters, to point out that I cannot at this moment believe something which I find wholly implausible is no more an argument against direct control over belief than the claim that I cannot now perform some action that I find utterly abhorrent is an argument against direct control over action. Direct control over belief determination, like direct control over action, is certainly mediated through our reason, and as such is subject to rational constraints. Certainly I will be at least partially excused from responsibility for such beliefs or actions, but this has no negative impact on the argument for direct control.¹⁰³

It is worth noting in passing the differing scope of beliefs that are subject to direct control on the regulative vs. guidance conceptions. Of course, much will depend on the details of a particular account, but it seems apparent that the conditions for regulative control will typically be satisfied less often than those for guidance control. This is because, strictly speaking, regulative control adds both a metaphysical and a psychological condition to control. As I've noted above, for control to be present on the libertarian account, we must have access to an alternative possible outcome, while holding the events preceding the relevant decision fixed. This of course requires the

¹⁰³ Considering these excusing conditions for belief and action actually serves to strengthen their parallels regarding control. The conditions under which I can/will be excused for a particular belief or action are exceedingly similar. Circumstances of rapid decision making under duress, cases of extreme inculcated bias or prejudice for which one is not responsible, instances of phobia or other forms of psychological impairment, all are potential responsibility-excusing conditions for both belief and action.

falsity of causal determinism. But it additionally requires that the force of reasons in favor of the actual outcome not be so compelling that it is psychologically impossible that we would have chosen otherwise. Of course, in many circumstances this latter condition is not fulfilled. In such cases, the libertarian will likely argue that we lack control over the belief in question. The compatibilist adherent to guidance control need argue no such thing. The conditions of guidance control described in this section can still be satisfied in such a circumstance, just as they may be satisfied even when I find the support for a particular action rationally compelling.¹⁰⁴ I will leave to the reader to decide which of the two accounts is most plausible in this and other regards, but it is worth noting the difference in scope, as well as the further similarities between our accounts of direct control over belief determination, and the respective accounts of control over action.

We have established, then, that the conditions for direct control over belief determination are satisfied, both for guidance and for regulative control. The rational control over action required by advocates of guidance control is something we regularly exercise in the determination of our beliefs as well. And, if one accepts the additional alternative possibilities requirement of regulative control for actions, one should recognize that this condition is also satisfied for a variety of beliefs. These considerations, along with those raised in the previous sections of the chapter, provide

¹⁰⁴ This remark is not intended to imply that advocates of guidance control, or compatibilists in general, have no method available to distinguish between compelled or manipulated belief determination and free belief determination. I take this to be a separate question. Assuming compatibilists are successful in making such a distinction, I am simply arguing that the scope of beliefs that would be considered as motivated in a way detrimental to responsibility would be much larger for the compatibilists than for the libertarian. Similar arguments have been advanced within the philosophy of action (see Van Inwagen (1989)).

considerable support to the thesis of moderate direct control, and offer a theoretical explanation buttressing our intuitive conception of belief as something for which we are directly responsible.

3.6. Intention: A Potential Objection

I will consider a number of arguments against direct control on a voluntarist conception in the next two chapters, but one additional objection deserves to be discussed here. This concern is not specific to either guidance or regulative control; if successful, it applies to both conceptions of voluntarist control. Recall my statement in the opening paragraph of this chapter that the voluntarist has a special challenge because of her need to show that we can believe at will. One way in which this challenge manifests itself concerns the connection between the will and intention formation. The objection takes the following form.¹⁰⁵ Our control over our actions is executed principally through the will. The primary mechanism of this control is through the formation of intentions. In parts 3 and 4 of my above account of deliberative action I judge that Φ is best for me to do, and decide to Φ . This decision results in the formation of an intention to Φ , which in turn leads to my Φ -ing. But this characteristic of action has no analog for belief determination, because the formation of beliefs does not involve the essential mediating role of intention. There is therefore a substantial disanalogy between the determination of

¹⁰⁵ I want to thank Dion Scott-Kakures both for raising this objection in conversation, as well as helping me in working through a response to it. A related objection to the one described here is offered by Nettleman (2006).

action vs. belief and between practical and theoretical examples. An essential component of direct control over action is not present in the determination of belief, and this calls into question whether we actually possess such control over our beliefs.

There much to be said in response to this objection.¹⁰⁶ Stated as above, the objector impales herself on one or the other horn of a dilemma. The objection is ambiguous concerning the nature of intention. What does it mean to say that voluntary action essentially involves the formation of an intention? What are such intentions supposed to be like? I suggest that there are two broad options available to the objector: a robust notion of intention, and a more minimal notion of intention. Embracing either conception, however, leads to the demise of the objection.

On the robust notion, intentions are conscious mental states of the form “I intend to Φ ” or “I am going/will to Φ ”. When I engage in an action that is under my direct control, it must involve some thought of this form. I judge that the holding my tongue in the argument is the best thing for me to do, I form the intention “I will not respond to those statements”, and I voluntarily do not respond. I judge that I ought to exercise more, and I form the intention “I will exercise more regularly”, which guides me in my future actions. Here we have a clear mediating role played by intention, one that seems lacking in at least most of our belief formation, dissolution, and retention.

¹⁰⁶ I additionally discuss the role of intention in belief determination in Chapter 4, sections 3 and 4.

However, the robust notion of intention also gives us a highly implausible conception of action itself. Certainly it is the case that with regard to some of my actions I engage in the process described above, involving the conscious or explicit formation of an intention. This is particularly true with respect to complex actions that take place over an extended period of time, where my intention governs my execution of a plan.¹⁰⁷ But clearly I do not engage in such a process for all voluntary actions, even those formed through a deliberative process. In many (perhaps the majority of) cases, I simply decide to Φ and then Φ , with no conscious intention formation. Thus, if the objector opts for the robust notion of intention, we may plausibly reject her claim as based on an implausible requirement for action. If the objection is true in this form, if this type of conscious intention is a requirement of voluntary action, then not only belief determination but also a wide swath of our actions thought to be clearly under our control are in fact not. Apart from being generally implausible, it is unlikely that any of our interlocutors, voluntarist advocates of responsibility for and control over action, would be willing to embrace such a result. But if actions under our direct control frequently do not satisfy the intention requirement so construed, then our belief determinations need not do so, either.¹⁰⁸

¹⁰⁷ Michael Bratman writes extensively about the role of intention in planning, practical rationality and agency extended over time, in Bratman (1987) and elsewhere.

¹⁰⁸ It is worth noting that even on this conception of intention, some belief formations, dissolutions or retentions may still qualify as intentionally undertaken. Some beliefs involve great internal struggle, and are resolved in a decision that seems to involve element of intention. The wife in an abusive situation, struggling with her husband's perennial promises to change his behavior, may reason to herself in precisely this way, thinking "I won't believe him. Not this time." Admittedly, whether or not this constitutes an intention under this description is somewhat ambiguous. Much more can be said on this topic, though I will not pursue this particular response to the objection here.

The alternative for the objector is a more minimal notion of intention. On this view intention is not essentially a particular conscious phenomenon. Instead, intention is identified simply as that which mediates between judgment and action when actions are under our control. On this view, to say that I acted voluntarily *is* to say that I acted intentionally, and the will is the faculty of intention because intention *is* what precedes voluntary action, of which the will is essentially a part. On this conception, intention is functionally described, rather than phenomenologically characterized. I intentionally dive out of the way of the oncoming car, even though I never form the thought, “I shall now dive out of the way”. The quarterback who successfully passes to his third receiver on a play does so intentionally (meaning under the direction of an intention), but not because he explicitly formulates the intention “I shall now pass to my third receiver” (he would have been sacked by the time he finished the thought!) before acting. And, to revive an example mentioned previously, I intentionally open the car door while trying put my screaming child in, even though no thought of the form, “I will open the door” passes through my mind. To return to a point made previously by Steup, I can tell by introspection after the fact that I performed a particular action intentionally, but not because I can recall a particular mental phenomenon of intention. Rather I can tell that I did the action intentionally because I can recognize introspectively that I caused the action, in a characteristic way (lacking external coercion, etc.) The minimal notion thus grants us more plausible results in a variety of cases of action where we clearly are in control of our actions, but where it is implausible to suppose that we possessed a conscious or explicit state of mind identifiable as an intention.

The minimal conception of intention is not burdened with a particular and implausible phenomenology or the requirement of specific conscious mental states that must directly precede action. However, what the objection gains in plausibility in this respect, it loses in effectiveness against direct control over belief determination. If intention is simply that which functionally precedes a voluntary action, and requires no particular conscious mental state, then why deny that beliefs are at least sometimes formed intentionally? While not vacuous, the requirement becomes vastly simpler to satisfy. As argued above, the characteristic psychological story of deliberative belief formation closely matches the account of deliberately determined action. If the standard story for action is sufficient for us to say that action is voluntary or intentional even without the presence of a conscious intention, then it is not clear why we should be unwilling to say the same about belief. In the absence of considerations to the contrary, it is plausible to think that this portion of the account of belief determination may match that of action as well. The concept of intention as a conscious mental phenomenon could perhaps have provided such a contrary consideration, but it was abandoned for the reasons noted above.

This objection to direct control over belief based on intention thus fails. Whether further objections raised by other authors fare better is the subject of the next two chapters. However, at this point it is worth taking stock of our progress. I have offered both an intuitive argument for direct control as well as theoretical support to direct control over belief determination on both an intellectualist and a voluntarist conception.

The intuitive argument (Chapter 1) arises from our standard ways of talking about belief, as well as our commonplace experiences of various beliefs. The theoretical argument in the present chapter is rooted in a plausible and influential account of account of action and the metaphysical and psychological conditions of control and responsibility. The obvious parallels between action and belief in terms of their normative structure and relation to reasons provide additional support. The stated goal of this chapter was to strengthen the argument for moderate direct control on a voluntarist conception, by providing a theoretical framework to support the presumption in favor of direct control established in Chapter 1. I suggest that in light of these considerations, this basic goal has been met and the burden of proof is squarely on the shoulders of the critic of direct control. Whether this burden can be met by the critic is the topic of our final two chapters.

CHAPTER 4 – VOLUNTARISM: CONCEPTUAL ARGUMENTS AGAINST DIRECT CONTROL

4.1. Conceptual & Psychological Objections

We have now articulated theoretical frameworks for direct control over belief on both an intellectualist and voluntarist conception. As I noted in Chapter 3, a much larger literature exists on control from a voluntarist perspective, including a number of prominent articles denying our ability to believe at will. Any comprehensive account of control over belief determination must respond to such arguments, and this will be my primary concern in this and the final chapter. Of course, it is impossible to survey all the arguments made against believing at will; I will concern myself with some of the most prominent arguments offered, examining what I take to be a representative sample of the various positions within the literature.

A useful distinction exists within this literature on believing at will¹⁰⁹ that will be helpful in categorizing my interlocutors. Arguments against direct control typically fall into one of two categories: conceptual or psychological. Those making conceptual arguments claim that it is something contained in the concept of belief or intentional action that makes it impossible to believe at will, to voluntarily decide to believe. In the present chapter I will concern myself with conceptual arguments to this effect offered by

¹⁰⁹ This is the term most frequently used by the critics of direct control, and I will use it interchangeably with “direct control over belief determination” throughout these chapters.

Bernard Williams,¹¹⁰ Dion Scott-Kakures,¹¹¹ Jonathan Adler,¹¹² and Robert Audi.¹¹³ The psychological arguments against believing at will contend that it is not something analytic to belief but rather some contingent feature of human psychology that makes believing at will impossible or unlikely. William Alston¹¹⁴ and Audi offer the two psychological arguments I will consider in Chapter 5.

4.2. Conceptual Arguments

Before beginning my response to the various conceptual arguments against believing at will, it is important to clarify the goal and identify what precisely will serve as a refutation. The proponents of these conceptual arguments are, as I noted, attempting to show that believing at will is necessarily impossible, that there is something amiss with the very concept of believing at will. Because of this ambitious goal, one need only show that believing at will remains possible in order to refute these arguments. If the concept of direct voluntarist control over belief is not incoherent, or does not produce the alleged incoherencies in the concepts of belief or intentional action, then the conceptual arguments in question fail.

¹¹⁰ Williams (1973)

¹¹¹ Scott-Kakures (1994)

¹¹² Adler (2003)

¹¹³ Audi (2001)

¹¹⁴ Alston (1988)

While this is the dialectic of the discussion, demonstrating the bare possibility of believing at will would be insufficient for the project of these chapters, which is to show that it is not merely possible, but also plausible, given voluntarist assumptions. My goal then, for both the conceptual and psychological arguments, will not be merely to offer a refutation, but to show that the idea of direct control over belief outlined in the previous chapter is not seriously undermined. Certain of the arguments I consider (Williams, Scott-Kakures) have been shown to fail, in the minimal sense noted above, by other philosophers whose responses I will discuss. However, because of my specific goal, I will in these cases still offer my own responses focused on their connection with my positive argument for direct control.

4.3. Williams

Bernard Williams offers perhaps the best known examples of conceptual arguments against believing at will in his seminal paper “Deciding to Believe”. Williams advances three criticisms of control over belief, all centered on the nature of belief itself.¹¹⁵ He begins by arguing that belief is aimed at the truth. As I’ve noted in previous chapters, there are controversial aspects to this claim, but it is not one which I wish to

¹¹⁵ It is worth noting that Williams limits the scope of his arguments in the essay to factual beliefs about the immediate external world. That is to say, his arguments are not directed at more abstract or theoretical beliefs. One upshot of this is that it is not immediately clear that Williams’ argument strictly contradicts my argument in Chapter 2. But it is plausible to think that a form of at least Williams’ first argument could be extended more broadly (and indeed has been by others, see Hieronymi (2006)) to cover other types of beliefs.

dispute. Williams' assertion that belief is aimed at the truth involves a cluster of related claims, which we can briefly state:

- i. Truth and falsehood are a dimension of the assessment of belief, unlike many other psychological states or dispositions.
- ii. To believe that p is to believe that p is true.
- iii. To say 'I believe that p ' carries, in general, the claim that p is true.¹¹⁶

All three of these claims are eminently plausible, though it is the first two that are particularly relevant for our discussion. Williams suggests that this truth-directedness of belief raises difficulties for the idea of believing at will. His argument is worth quoting at length:

If I could acquire a belief at will, I could acquire it whether it was true or not; moreover I would know that I could acquire it whether it was true or not. If in full consciousness I could will to acquire a 'belief' irrespective of its truth, it is unclear that before the event I could seriously think of it as a belief, i.e., as something purporting to represent reality. At the very least, there must be a restriction on what is the case after the event; since I could not then, in full consciousness, regard this as a belief of mine, i.e. something I take to be true, and also know that I acquired it at will. With regard to no belief could I know—or, if all this is to be done in full consciousness, even suspect—that I had acquired it at will. But if I can acquire beliefs at will, I must know I am capable of doing this; and could I know that I was capable of this feat, if with regard to every feat of this kind which I had performed I necessarily had to believe that it had not taken place?¹¹⁷

I think that Williams makes two distinct points in this passage, one about the acquisition or formation of beliefs, and one about their retention. The former claim is sometimes not

¹¹⁶ i, ii, and iii are all from Williams (1973), p. 137.

¹¹⁷ Ibid., p. 148.

identified in discussions of Williams' work, perhaps because Williams himself appears to abandon the line of thinking so quickly.¹¹⁸ What Williams seems to be suggesting is that, because believing at will involves the ability to believe things irrespective of their truth, I would be able to form an intention to believe some proposition p , all the while thinking that p is not true. There is, as Williams notes, some deep incoherence about this prospect, rooted in the truth-orientation of belief, that speaks against its possibility.

Williams' second claim concerns not belief formation, but belief retention. It is perhaps most clearly expressed in a structured form. I take the argument to be something like the following:

- (1) One cannot hold a belief and simultaneously know or suspect that one had acquired it at will (ie. without justification).
- (2) Therefore, one must believe that one has acquired none of one's beliefs at will.
- (3) In order to believe at will, one must know that one is capable of believing at will.
- (4) But it is impossible for one to know that one can believe at will while simultaneously believing that none of one's beliefs have been generated at will.
- (5) Therefore, it is impossible to believe at will.

So the conflict is between my ability to believe at will, and my need to disbelieve that any of my beliefs were formed at will. Stated in this way, the problems identified by many commentators, which I discuss below, may already be apparent.

¹¹⁸ Jonathan Bennett seems to suggest as much in Bennett (1990), p. 92. I'm not sure that this is the reason, or that this is an accurate divination of Williams' thinking, he makes all three of his points rather tersely at the end of his essay, without much further discussion.

Williams' third argument, made separately from the first two, is rooted in a functionalist account of belief.¹¹⁹ This argument is focused on perceptual or "empirical" beliefs in particular, though it could presumably be generalized to other kinds of belief. A dominant and broadly functionalist way of thinking about perceptual beliefs suggests that they are individuated and possess their meaning in virtue of their typical causes and relations with other mental states in one's cognitive economy. On this view, Williams suggests, a perceptual belief just is a mental state that has a certain kind of cause, namely that it is produced in the right kind of way by one's perceptual organs in contact with the external environment. But if this is the correct way of thinking about perceptual beliefs, then deciding to believe is conceptually impossible. One might produce something belief-like through the activity of the will. But it would not be a belief, because, as Williams says, "there would be no regular connection between the environment, the perceptions and what the man came out with, which is a necessary condition of a belief or even of a [belief-like]*B*-state."¹²⁰ Anything produced at will would lack the proper causal ancestry and connections to qualify as belief.

So we have three arguments from Williams against the idea of believing at will, all rooted in the concept of belief. Believing at will is impossible because: (1) we could not prospectively regard the mental state to be acquired as a belief, (2) we could not satisfy the requirements of believing at will and still know that we were capable of such

¹¹⁹ Williams (1973), pp. 148-49.

¹²⁰ Ibid., p. 149, my addition in brackets. The distinction between beliefs and *B*-states is not relevant for our discussion.

belief, and (3) any belief-like mental state generated at will would not be a belief.

Williams' arguments have been much-criticized; I offer here some of the most telling criticisms before turning to my own discussion of his claims.

Jonathan Bennett has famously responded to Williams' second claim, about belief retention, by inventing the fictional community of Credam, where exactly those circumstances that Williams denies in fact obtain:

Credam is a community each of whose members can be immediately induced to acquire beliefs. It doesn't happen often, because they don't often think: 'I don't believe that P, but it would be good if I did'. Still, such thoughts come to them occasionally, and on some of those occasions the person succumbs to temptation and wills himself to have the desired belief...When a Credamite gets a belief in this way, he forgets that this is how he came by it. The belief is always one that he has entertained and has thought to have some evidence in its favour; though in the past he has rated the counter-evidence more highly, he could sanely have inclined the other way. When he wills himself to believe, that is what happens; he wills himself to find the other side more probable. After succeeding, he forgets that he willed himself to do it...

...After successfully willing himself to have a certain belief, a Credamite may later get evidence that that is what he has done; e.g. someone may tell him. Then he either rejects the evidence (e.g. disbelieves the informant) or else accepts that he has willed himself to have the belief, and usually loses the belief because of his knowledge of how he got it. In the latter case, he still remembers having had the belief, and now remembers willing himself to acquire it...So each Credamite knows that he sometimes wills himself to believe something, even though it is never true that he *now* has a belief which he *now* remembers having willed himself to acquire.¹²¹

¹²¹ Bennett (1990), p. 93.

Whatever one thinks about the plausibility of Credam's existence, it is clearly possible, which is all Bennett needs to prove his point against Williams.¹²² Credamites satisfy the requirements of Williams' second argument, they know they can believe at will, but never think that any of their current beliefs were formed at will.

Should Bennett's fictional construct seem dubious, however, Dion Scott-Kakures offers an even more straightforward counterexample to Williams' general claim of tension between believing at will and knowing that one has done so: self-fulfilling beliefs.¹²³ Because of its nature, the genesis of a self-fulfilling belief is immaterial to the subject that continues to hold it. We may adapt William James' famous train robbery example to illustrate this point.¹²⁴ In our version, a group of passengers on a train are each faced with the question of whether or not he or she should arise in defiance of the thieves. One passenger decides that if she stands, other passengers will rise up with her in defense of the train. The other passengers have no such firm resolution, and the initial passenger possesses no great evidence in support of her belief. Nevertheless, as she rises, the other passengers see her, are inspired by the show of bravery, and rise from their seats, their numbers quickly overthrowing the thieves. By deciding that the other passenger will act in such a way, and acting on her initial belief, she thereby acquires all

¹²² Barbara Winters suggests an alternative criticism of Williams argument, see Winters (1979).

¹²³ This could I suppose be seen as a special instance of the type of case that Bennett (1990) mentions in a footnote on p. 93.

¹²⁴ James (1896), pp. 24-25.

the evidence she needs in support of her belief.¹²⁵ Realizing later that she simply decided to believe it and lacked evidence in its favor will not lead to the belief's dissolution.

Bennett and Scott-Kakures also comment on Williams' third argument, rooted in his functionalist analysis of perceptual belief. Though both are sympathetic (particularly Bennett) to the basics of Williams' functionalist account, both acknowledge that it fails as an argument against believing at will. The most plausible functionalist analysis of perceptual belief, Bennett notes, has such beliefs identified by their *typical* causes and connections within a subject's mental framework. But for his argument to go through, Williams needs the claim that perceptual beliefs are *always and essentially* caused by the interaction of the surrounding environment with the perceptual organs of a subject. And that claim, as Bennett says, we have no reason to accept.¹²⁶

Scott-Kakures' take on Williams argument is similar, and his indictment is no less severe:

As stated, this argument begs the question. The argument has it that beliefs must be caused in certain ways if they are to count as beliefs. Yet the alleged ways in which a belief must be caused are just the ways in which the apologist for believing at will must deny are essential to belief genesis. More important, if someone *does* manage to will herself directly into a psychological state which interacts with other psychological states

¹²⁵ James' cases differs from mine in that he is not explicit that the belief ("faith" is James' term) of the heroic passenger being specifically acquired through a voluntary act of direct control. James is engaged in a different debate, concerning evidentialism rather than doxastic voluntarism, and so the genesis of the belief is not so much his concern as is the evaluation of its rationality. I utilize the example because of its familiarity and don't think I have done any violence to it by stipulating this aspect within my version.

¹²⁶ Bennett (1990), p. 95.

and produces behavior in the typical ways, that state counts as belief even if produced in some aberrant fashion.¹²⁷

Scott-Kakures considers the possibility of strengthening Williams' argument by connecting it with current functionalist accounts that emphasize the connection between the typical causes of belief and the roles beliefs play in one's psychology and reasoning. But while the argument thus construed is an improvement over Williams' original, Scott-Kakures concludes that it nevertheless fails to rule out the possibility of beliefs being formed at will, and yet playing the appropriate part in one's reasoning.¹²⁸ The prospects for Williams' functionalist strategy seem dim.

What about Williams' first claim, that one could not prospectively view a belief to be willed as a belief? As noted, many commentators have ignored this portion of Williams' argument. Bennett, however, does briefly consider the issue. Bennett asks us to consider the case of someone seeking to directly induce a belief in herself for practical reasons. Could she not simply think to herself, he asks, "I shall get myself to believe that P; the belief will be false, and I shall be deluded, but this is what I want."¹²⁹ If it is possible for one to believe purely on practical grounds, then perhaps this form of reasoning would be open to our subject.

¹²⁷ Scott-Kakures (1994), p. 84.

¹²⁸ Ibid., p. 84-86.

¹²⁹ Bennett (1990), p. 92.

But is it possible to believe in this way *purely* on practical grounds? It is not clear that it is, or at least that it is directly. Recall Williams' requirement that one make the required move between unbelief and belief "in full consciousness" of the falsity of the proposition in question. If our subject were able to do this, she would find herself in a state of serious cognitive tension, with both a belief that the proposition in question is true and one that it is false. It seems dubious to think that she could get there directly at all. Bennett's phrasing of the reasoning above betrays this, our subject says "I shall get myself to believe that P", rather than the more straightforward and direct "I shall believe that P". If we were to substitute the more direction formulation, it would start to sound dangerously close to a version of Moore's Paradox, "I shall believe that P; the belief will be false." It seems implausible to attribute such a (truthful) statement to any minimally rational subject.

I am inclined at this point to give up on Bennett's line of reasoning, which does not seem particularly promising. I think the precise meaning of Williams' first argument is difficult to ascertain, as is the success of Bennett's response to it. I want to suggest that we need not respond in such a way to Williams' first argument; that we need not show that it is mistaken. Instead, I contend that Williams' first argument is simply irrelevant to the question of whether or not one can believe at will, because it rests on a false assumption about this phenomenon.

Williams opens his argument with the statements, "If I could acquire a belief at will, I could acquire it whether it was true or not; moreover I would know that I could

acquire it whether it was true or not. If in full consciousness I could will to acquire a ‘belief’ irrespective of its truth...” Of course, Williams cannot mean here what the straightforward reading of his sentences implies. As we all know, much to our chagrin, we have the ability to acquire beliefs that are not true. Williams is not speaking about the relationship between our beliefs and the world they represent, he is speaking about the relationship between our beliefs and our perceptions of the world. In the parlance of Chapters 2 and 3, he is speaking about L2 rather than L1 rationality. He is suggesting not that believing at will would allow us to acquire false beliefs, but rather that it would allow us to acquire beliefs while knowing or thinking that they are false. Direct control over our beliefs, Williams thinks, would allow us to embrace beliefs not simply irrespective of their truth, but irrespective of what we think about their truth. And it is understandable that he is disturbed by this possibility. As mentioned above, there is something deeply incoherent about this idea.

But, as the arguments of the previous chapter showed, this is a relatively implausible picture of direct control over belief. Williams’ willful believer is a straw man. There is no reason for the doxastic voluntarist to commit herself to such an inflated conception of control. Recall again that moderate direct control is the thesis that the determination of belief is sometimes subject to control in much the same way as actions are. Specifically:

- (1) An individual’s beliefs are responsive to reasons about what she ought to believe, in ways similar to how her actions are responsive to reasons about what she ought to do.

- (2) In certain circumstances, individuals have the latitude to decide what to believe.

Under no interpretation of this thesis was it suggested that individuals might be able to willfully believe disconnected from or in opposition to what they think is true. As I noted in the previous chapter, saying that direct control over one's beliefs requires the ability to believe irrespective of the evidence is analogous to saying that direct control over one's actions require the ability to act regardless of one's reasons, or one's conception of the goodness of a prospective action. No one claims the latter about action, why then should we accept the former about belief? But Williams requires precisely this for the success of his first argument. It is only in those circumstances where one knows or believes some proposition to be false that it seems hard, as Williams says, to seriously think of what one is acquiring as a belief. The advocate of moderate direct control concedes as much, and suggests that this immaterial to the question of whether or not one can believe at will. Williams' first argument simply assumes too much, packs too much into the meaning of deciding to believe, and thus fails to threaten the argument for direct control over belief. As his second and third arguments have been shown to fail, we may conclude that Williams poses little threat to the argument for moderate direct control.

4.4. Scott-Kakures

Though he is a critic of Williams' arguments, Scott-Kakures suggests that the spirit of Williams' project may be revived if we turn away from an analysis of belief and look instead to the concept of intentional action. Exercising direct control over belief,

Scott-Kakures suggests, must be an intentional action. It is simply to believe something intentionally or through a decision. But just as Williams argued that there was a tension between deciding to believe and the concept of belief, Scott-Kakures contends that there is an incompatibility between deciding to believe and the concept of intentional action.

As an intentional action, Scott-Kakures argues, believing *at will* involves the transition from one state, at $t1$, where I desire to believe that p , directly to another state, at $t2$, wherein I believe that p . But not just any transition will allow the process to count as believing *at will*. Specifically, the transition cannot occur via a deviant causal chain. This type of deviant causal process is widely discussed in the literature on intentional action, Davidson's case of the timorous mountaineer being perhaps the most famous example.¹³⁰ In this example, a mountaineer is climbing with a partner, whom he is supporting by his grip on a rope. The mountaineer forms an intention to let go of the rope, which would cause his partner to fall. This thought so unnerves the mountaineer that he loses his grip on the rope, and his partner in fact falls, though he never actually executes his intention. In this type of case, the transition between a desire and its realization is made, but through a deviant causal process that is irreconcilable with the action's being intentional. Believing intentionally is similarly incompatible with this type of causal deviancy.

So forming a belief *at will* must involve a transition from unbelief to belief consistent with a plausible account of intentional action. There is, Scott-Kakures confesses, substantial disagreement over the proper characterization of intentional action

¹³⁰ Davidson (1973).

in response to deviant cases. But there nevertheless is broad agreement surrounding the necessity of the intention playing a guiding, controlling or monitoring role in an intentional action.¹³¹ Part of this monitoring is accomplished through the content of the intention itself. So, when I intend to raise my arm, part of the content of that intention is self-referential, specifying “that I perform the action of raising my arm by way of carrying out this intention.”¹³² As Scott-Kakures says:

To say that an intention directs an activity is just to say that I monitor and guide my activity against the background of the intention. Thus when I intend to raise my arm and I succeed, I know when to stop trying. It is this guiding and monitoring of the agent’s by appeal to the content of the intention that serves to distinguish the brute Humean causes of the deviant cases from the rationalizing causes essential to intentional psychological explanations.¹³³

As an intentional action, then, believing at will must be able to satisfy the requirement of a monitoring intention. The problem, as some respondents to Williams have noted,¹³⁴ is that it seems to be essential for any plausible account of believing at will that there be a gap in the monitoring by the intention to believe at will in the transition between my unbelieving state at $t1$, and my later believing state at $t2$. This gap is required, Scott-Kakures suggests, by the necessity of my possessing at $t1$ the following belief:

¹³¹ Scott-Kakures (1994), p. 89.

¹³² Searle (1983), quoted from Scott-Kakures (1994), pp. 90.

¹³³ Scott-Kakures (1994), p. 91.

¹³⁴ See Scott-Kakures (1994), p. 83 specifically cites both Winters (1979) and Bennett (1990).

(B) A belief in p at $t1$ is unjustified.

This belief, Scott-Kakures holds, is necessary for three doxastic states that a subject must possess to move between not believing that p and believing that p directly by willing:

(S1) My believing that I do not believe that p ;

(S2) My not believing that p ;

(S3) My intending to believe that p .

A belief in B at $t1$ is most relevant to S3. Were I to not believe B, Scott-Kakures suggests, I would not need to form the intention to believe by directly willing it. If I thought at $t1$ that believing that p is currently justified, then I would simply believe on that basis.

So B is integral to the intention to believe at will. However, Scott-Kakures argues, this belief cannot be present at $t2$, where I believe that p . On the traditional analysis, to believe that p is to believe that p is justified, and so, were I to continue to hold B, I would simultaneously possess two clearly contradictory beliefs.¹³⁵ B must somehow drop out in the process. But if this belief is part of the content of the intention that is governing the belief formation, then this intention must lapse between $t1$ and $t2$, there must be some gap or “cognitive blind spot” that allows B to drop out. But this type of gap or lapse in the monitoring role of the intention is incompatible with intentional action. If this gap is a necessary feature of believing at will, then believing at will must be impossible. That is

¹³⁵ Of course, we all probably possess certain beliefs which are contradictory, but it is implausible that one would hold two so clearly contradictory beliefs in a manner where they would be both be so closely occurrent.

not to say that the transition between the states at $t1$ and $t2$ cannot occur. Scott-Kakures admits that such a movement might occur through a deviant causal path, but this type of deviant or anomalous transition is not worthy of the name “believing at will”.

- (1) To believe p at will at $t2$ I must believe B at $t1$.
- (2) However, believing p and believing B simultaneously is incoherent.
- (3) Therefore, I must drop my belief in B somewhere between $t1$ and $t2$.
- (4) But dropping my belief in B between $t1$ and $t2$ is inconsistent with my acquisition of a belief in p being an intentional action.
- (5) Believing at will must be an intentional action
- (6) Therefore, dropping my belief in B is inconsistent with believing at will.
- (7) Therefore, believing p at will is impossible.

Scott-Kakures’ argument is novel and intriguing, but it suffers from a central flaw, identified by Dana Radcliffe.¹³⁶ The problem lies in step 1, and concerns the alleged necessity of belief in B by the subject at $t1$. Scott-Kakures claims that B must be present at $t1$ as a necessary condition of forming the intention to believe at will. However, Radcliffe argues convincingly that B is not an essential component of a subject’s cognitive environment at $t1$. Of course, Scott-Kakures is correct to claim that a subject cannot occurrently believe that p is justified at $t1$. If she did, she would either already believe that p at $t1$, or have no need to believe that p directly by willing it, and thus would form no intention to do so. However, the same subject need not believe that p is *unjustified*; she instead could simply be agnostic as to the justificatory status of p at $t1$. In cases where the subject has not had time to survey the evidence she possesses for and against p , it is plausible that a subject would have no such belief on the matter. The

¹³⁶ Radcliffe (1997).

situation is similar in cases where the evidence relevant to a proposition is extensive and/or complex and nuanced, cases which I noted previously seem ripe candidates for deciding what to believe. Such subjects could still at $t1$ hold all of the doxastic attitudes (S1) through (S3) that Scott-Kakures argues are essential for forming the intention to believe that p directly by willing it, despite their failure to hold B. Since the presence of B at $t1$ is the lynchpin of Scott-Kakures' argument, it fails.

It should be clear that Williams' first argument and Scott-Kakures' argument are structurally similar. The latter argument is thus open to a structurally similar response to the one I made above to Williams' first argument. Radcliffe's argument is just such a response. Both Williams and Scott-Kakures paint a picture of belief and justification that is too stark, and use this to deny the possibility of believing at will. The more plausible picture, as discussed in the previous chapter, is of a spectrum of justification, and believers who may occupy any one of a number of points along it. At one end of the spectrum sit subjects who think some proposition p wholly unjustified, because they believe that it is in fact false, or they believe that substantial evidence contradicts it. As Williams notes, it is inconceivable that someone in this position might move directly and willfully to the belief that p is true. Somewhat further in on the spectrum are those subjects that don't necessarily believe that p is false, but nevertheless think p seriously unjustified. Here, too, it is hard to imagine how one might move directly and intentionally to the belief that p is true. However, at the risk of being repetitious, believing at will is not simply the improbable phenomenon of a subject propelling herself

from a state of viewing a belief as completely unjustified (or perhaps even as having evidence against it) to a state of wholehearted belief. Believing at will, or exercising this form of direct control over belief determination includes any instance where a subject bridges the justificatory gap between the state of non-belief and the state of belief by an act of decision, or a direct exercise of the will. As we saw, this movement may be much less dramatic than the cases envisioned by Williams and Scott-Kakures, it may even be the slight tipping of an evenly balanced scale. But this does not make the action any less an exercise of direct control over belief.

4.5. Adler

The first two authors surveyed criticized believing at will by focusing on the concepts of belief and intentional action, respectively. Our third author, Jonathan Adler, advances a series of arguments against direct control over belief in the course of the somewhat more oblique project of defending evidentialism in his *Belief's Own Ethics*.¹³⁷ Evidentialism, as noted in Chapter 2, is the normative thesis that beliefs ought to be apportioned to the evidence one possesses. Adler notes the natural affinity between traditional evidentialism and voluntarism with respect to belief. Traditional, or what Adler calls “extrinsic” approaches to evidentialism suggest that the relationship between belief and evidence is deontological, one of ethical or other forms of obligation

¹³⁷ Adler (2002). It is not possible in this chapter for me to summarize or respond to all of Adler’s argument, I am therefore focusing on the most relevant portion of his work, especially his second chapter, “Can One Will to Believe?”, pp. 55-72

impinging upon the believer to conform her beliefs to the evidence. It is of course natural (“ought implies can”) to think that this type of obligation only makes sense if one can in fact do this, if one can voluntarily conform one’s belief to the evidence. Thus the language of evidentialism seems to imply voluntarism.

Adler rejects this picture of evidentialism, opting instead for what he calls an “intrinsic” approach. By this he means to posit a conceptual connection between belief and evidence. The force of epistemic norms is conceptual rather than deontological. As he says, the “pivotal modals terms are ‘must,’ ‘cannot,’ or ‘can,’ not those of deontology: ‘ought,’ ‘obligatory,’ or ‘permissible.’”¹³⁸ Given this conceptual connection it is perhaps unsurprising that Adler rejects voluntarism as well. If beliefs are not suggested by the evidence, but rather bound or determined by it, then one might think that our control over our beliefs is dubious. On this picture, rather than our freely conforming our beliefs to the evidence, our beliefs are compelled by it.

Adler advances two related lines of argument to support his denial of direct control.¹³⁹ The first he calls, appropriately, the “conceptual argument”; the second he terms the “instability argument”. Adler, like Scott-Kakures, is self-consciously indebted to Williams’ famous arguments. In contrast to Scott-Kakures, however, Adler sees

¹³⁸ Ibid, p. 55.

¹³⁹ Adler considers a number of individual cases of belief in the course of arguing against direct control. I do not consider most of these cases here for two main reasons. First, Adler claims that his consideration of individual cases is intended only to clear away putative counterevidence to his broader conceptual thesis. Second, many of the cases are very similar to those considered by Audi and especially Alston, and thus shed little new light on the issue of direct control.

brighter prospects for Williams arguments as originally formulated. Both the conceptual and instability arguments, Adler says, are closely connected with Williams' claim that "it is not possible to regard oneself as holding a belief, while recognizing it as due to the exercise of the will."¹⁴⁰ I will begin by briefly outlining Adler's conceptual argument, and consider certain responses to it, and then turn to his instability argument.

Adler's conceptual argument starts with what he calls the *subjective principle of sufficient reason*. This principle requires that "when one attends to any of one's beliefs, one must regard it as believed for sufficient or adequate reasons."¹⁴¹ The support for this principle is found in what Adler calls the "incoherence test", which is based broadly in Moore's Paradox. Recall that Moore's claim is that it is impossible to truthfully assert, "*p*, but I don't believe it." The statement is not a contradiction, because both of the conjuncts of the statement could be simultaneously true. But nevertheless the assertion of such a statement is incoherent. The incoherence is caused by the fact that a truthful assertion that *p* by an individual carries with it a certain kind of commitment, namely that the individual believes that *p*. Thus, the person saying "*p*, but I don't believe it" appears to be asserting, or at the very least implying a contradiction of the form "I believe that *p*, but I don't believe that *p*." And it is incoherent to attribute to any person such an obviously contradictory set of occurrent mental states.

¹⁴⁰ Adler (2002), p. 58.

¹⁴¹ Ibid., p. 26.

Adler makes use of this close connection between assertion and belief. His incoherence test employs a related locution, “ p , but I don’t have sufficient/adequate evidence that p ”. Consider for example, “The number of stars is even, but I lack sufficient evidence that the number of stars is even.” This statement is incoherent, Adler suggests, in much the same way as Moore’s original. Plug in any p that you like and you will get the same result. And the phenomenon holds true for negative forms of the same statement as well. “ $\sim p$, but I have more than sufficient evidence that p ” sounds similarly incoherent.¹⁴² As Adler says, “The incoherence test exposes not only what cannot be believed, but also what must be believed.”¹⁴³ Adler claims that the incoherence of such statements is explained by the conceptual connection between belief and evidence. Belief and evidence are conceptually bound in the way described by the subjective principle of sufficient reason; when I consider any belief, I *must* regard it as supported by adequate reasons. This “must” is conceptual, not deontological. To not think this about some proposition just is to fail to believe it.

The incoherence test is relevant to the issue of direct control because instances of believing at will violate this conceptual connection between belief and evidence. When one believes at will, Adler suggests, one explicitly believes in spite of one’s acknowledged lack of evidence. As we saw in considering Scott-Kakures’ arguments,

¹⁴² One interesting issue which I will not explore in this chapter is difference between this negative form of the statement and the original positive form. The ring of incoherence in the second statement is much less severe, it seems to me. h

¹⁴³ Adler (2002), p. 31.

were one to see oneself as possessing evidence for p that one regarded as sufficient, there would be no need to form the intention to believe it at will. Only when I have failed to form such a judgment about the evidence can I generate the relevant intention. But in such a case I will end up, Adler says, with a mental state something like the general form of the incoherence test. My mental state will be something like, “ p , but I believe it because I choose to,” as in “The number of stars is even, but I believe it because I choose to.” But such a statement, Adler claims, is just as incoherent as those considered above. It is this kind of incoherence that Williams was referring to in his discussion of believing at will, when he said that I could not regard any such state as a belief. And it is this incoherence that makes it impossible for me to form beliefs on various subjects simply as I choose. Belief is conceptually constrained by evidence, and believing at will is thus impossible.

Adler’s argument initially seems appropriately categorical, but it in fact involves an important qualification or limitation on its scope. This qualification figures prominently in my response to Adler’s conceptual argument (which is his principal criticism of direct control). My response has several distinct parts. I will first consider how one might respond to Adler while accepting the legitimacy of his assertion test. I will then suggest some important reasons why the incoherence test is misleading as a measurement of the plausibility of believing at will. Finally, I will turn my attention to Adler’s consideration of an alternative response to his conceptual argument. This in turn will lead to a discussion of his instability argument.

First, let us consider how one might respond while accepting the legitimacy of the incoherence test. The chief limitation of Adler's incoherence test is the scope of belief to which it applies. As discussed just above, the statement that comprises Adler's incoherence test is of the form, "*p*, but I lack sufficient evidence that *p*." The statement "*p*" in Adler's version of the paradox is meant to imply what he calls "full belief." By this, Adler means belief with a high degree of certainty and commitment, belief "full stop". If you were to ask me, "*p*, or $\sim p$?", and I replied confidently, "*p*," this would be a paradigmatic example of full belief as Adler conceives of it. It is, Adler thinks, the standard form of assertion, and he holds full belief to be the paradigmatic form of belief, both in the sense of being the most common form, and in the sense of being most clearly belief.¹⁴⁴

I find these claims regarding assertion and belief dubious. While assertion with full belief may be the "standard" form of assertion, it is certain not the only form. And much more importantly, full belief is not the only (and perhaps not even the paradigmatic) form of belief. Belief is a gradated concept. This is clear from the fact that I can compare or rank my beliefs relative to my conception of their probability. If you ask me, "Are you more certain that your wife loves you or that Aaron Burr shot Alexander

¹⁴⁴ Strictly speaking, there is at least one further qualification to Adler's argument worth mentioning. Adler's initial subjective principle specifically concerns circumstances "where one attends" to the belief in question. Here Adler restricts the impossibility to cases where one is focused upon one's belief. This qualification is perhaps not particularly troubling, though, given Adler's characterization of believing at will. Adler explicitly requires that direct control involve an intention to believe, and the formation of the belief by or under the direction of that intention. It is at least hard to imagine cases where this requirement is satisfied, but the subject is not attending directly to the belief in question.

Hamilton?”, and I opt for the former proposition, I am expressing my judgment that it possesses superior justification, that it is more probably true than the latter. But I still clearly believe the latter proposition. And, since belief occurs along a spectrum of probability, we could determine beliefs that I hold which I think are less probable, or of which I am even less certain of than this latter proposition (say, my belief that Russia’s population will drop below 100 million persons by 2050). My belief in this third proposition is weaker than full belief, but it is still clearly a belief.

Supporting the claim that the attitude I’m describing is belief is the fact that I would be willing to assert my belief in the proposition. “Did Aaron Burr shoot Alexander Hamilton?” you ask me. “I believe he did,” or “I think he did,” I might respond. Note the specific grammar of my response. Belief that is weaker than full belief (let us call it “qualified belief”, for lack of a better term) is characterized by this type of self-referential or self-conscious assertion. The speaker is flagging, both for himself and his interlocutor, the fact that while he believes the proposition in question, he holds this belief to be something less than certain.

It is interesting to note that this form of assertion, incorporating qualified belief, seems far less troublesome when plugged into Adler’s incoherence test. “I think that God exists, but I believe it because I choose to” does not smack of incoherence in anything like the way that Moore’s original paradox did. It is not clear, barring some conceptual commitment to involuntarism or Adler’s form of evidentialism, why one ought not attribute such a confluence of mental states to a subject. It is not even clear that “but”,

which implies some incongruence or reversal, is the proper form of conjunction between the two statements. It could just as easily be “and”, as in “I believe that Jill and I will be happy together, and I believe it because I choose to.”¹⁴⁵

I claimed in Chapter 3 that on a libertarian conception direct control over belief is possible at least in cases where the evidence is inconclusive, is unclear or seems only somewhat tilted one way or the other. Interestingly, this correlates with the conclusion of this interchange with Adler as well. Circumstances of full or unqualified belief are determined by instances of compelling evidence. These are cases where I believe things with a high degree of certainty. I cannot achieve certainty (directly) through an act of the will, this is something that only comes as the result of very strong evidence. But we often lack such evidence. In many such cases, circumstances require us to make up our mind. Whether they do or not, however, in instances of inconclusive evidence we have the ability to decide to believe or to withhold belief. If we do form a belief intentionally, the resulting state will be one of qualified belief, typified by the tempered or self-referential forms of assertion noted above. But though my assertion of such a belief is qualified, it is nevertheless still clearly a belief.

Thus far I have offered a response to Adler while accepting (for the sake of argument) the legitimacy of the incoherence test as a gauge of the plausibility of direct control. I now wish to question that legitimacy. In reality, the incoherence test is

¹⁴⁵ This is admittedly an awkward formulation of such a sentence, but the lack of a contrastive conjunction is just as evident in the more natural form of the sentence, “I choose to believe that Jill and I will be happy together.”

fundamentally misleading. It fails to give an accurate characterization of direct control over belief, and any force that it possesses as a challenge to the plausibility of direct control over belief determination derives from this fact, rather from real problems with the possibility of believing at will.

How is the incoherence test misleading? The incoherence alleged by Adler in cases of believing at will seems to derive from a kind of stark contrast between the two parts of the statement under consideration. Consider the incoherence test applied to a case of testimony, as in, “John is telling the truth, but I believe it because I choose to.” Adler claims that there is some significant incongruity between “John is telling the truth,” on the one hand, and “but I believe it because I choose to” on the other. Now, if my discussion in the preceding paragraphs is correct, part of the incongruity derives from the implausible force of the assertion in the former clause of the sentence, representing full belief. We simply don’t acquire the kind of certainty associated with full belief in beliefs generated by intentional decision. So, the test case is unrealistic. But we can remove this variable from our discussion by adding in a self-referential indicator of qualified belief, making the former clause, “I believe that John is telling the truth.”¹⁴⁶

¹⁴⁶ It should be noted that statements of the form “I believe that *p*”, or “I think that *p*”, do not always function to indicate or flag the subject’s lower confidence in a belief. Such assertions can be made for a variety of purposes, for example, the emphasis in the assertion may be on “I”, rather than on “believe”, and serve principally to indicate that it is I who believe, not someone else. Much of this will of course depend on the conversational context of the statement. But I think it is fair to say that the indication of some degree of uncertainty is one of the most common usages of this type of statement.

Now our sentence is “I believe John is telling the truth, but I believe it because I choose to.” Does any incongruity remain? I suspect Adler would claim that it does. Where then does this incongruity derive from? I think it becomes apparent at this point that the circumstances described by the sentence are ambiguous. We can add clarity and further our discussion by fleshing out different possible descriptions of the subject’s reasoning in such a case. The legitimacy of such an approach is explicitly endorsed by Adler, who uses the same tactic in fleshing out the incoherence in the more general version of his incoherence test.¹⁴⁷ Consider the following three possible descriptions of the subject’s reasoning:

- A. I believe that John is telling the truth, but I believe it because I choose to. I have no prior acquaintance with John, and presently he seems evasive and dishonest. I have no reason to believe that he is telling the truth, other than that he has asserted “*p*”.
- B. I believe that John is telling the truth, but I believe it because I choose to. John is a friend and has in many cases told me the truth in the past. In circumstances such as these, John has sometimes been unreliable, but he seems sincere in this instance. It seems to me that I probably ought to believe John, though I find myself somewhat hesitant for reasons I can’t entirely put my finger on. So I choose in the end to believe he is telling the truth.
- C. I believe that John is telling the truth, but I believe it because I choose to. John is a good friend and my relationship with him has great value to me. The matter in question is not of great importance, though it has some emotional significance. I am ambivalent about the evidence of John’s truthfulness in this instance; the matter is simply not clear to me. But if I do not believe that John is telling the truth, it is likely that it will have negative consequences on our relationship. Because of the ambiguity of the evidence and the value of our friendship, I choose to believe John.

¹⁴⁷ See Adler (2002), pp. 30-31.

The first case, A, clearly differs from the latter two. There is some clear form of incoherence in A. It seems clear that I should not choose to believe John, and it seems psychologically implausible that I could choose to believe John in such a case. But, as should be obvious, A bears no resemblance to the type of case around which the previous chapter's argument for direct control was built. A is an instance of believing against the evidence, or at the very least while having no good reason to do so, and I have never suggested that this is possible. Moreover, the impossibility of such a voluntary belief is no threat to the general possibility of direct control. An analogous case could easily be constructed for action, and it would seem just as psychologically implausible that one could so act voluntarily.¹⁴⁸ But this would be no threat to the general possibility of action.

B and C, on the other hand, contain no obvious contradiction or incoherence in the subject's reasoning. Indeed, B and C look like reasoning patterns that we commonly engage in, and match with reports individuals regularly give of their reasoning. B and C differ with respect to the role that practical or non-evidential factors play in the adoption of the belief, and I include both for reasons that will be apparent below. B matches precisely the account of control offered in Chapter 3, where a subject in the face of inconclusive or unclear evidence may choose to adopt a belief or withhold belief.

Most importantly for our purposes, it is apparent that once we specify the reasoning involved in the cases, the incoherence required by Adler can only be produced

¹⁴⁸ Consider, for example, the following analogous reasoning for action: "I choose to burn my hand, though I have no reason to do so. I am under no coercion to choose this, and I am aware that the action will cause me great harm and pain, which I do not desire. But I choose to burn my hand."

by begging the question against the voluntarist. This incoherence is obvious in the case of Moore's Paradox, because the contradiction is straightforward, and immediately apparent. The contradiction in the subject's assertion is somewhat less obvious in the case of Adler's incoherence test, and requires one to fill in the reasoning that makes it clear, as Adler recognizes. And in cases of believing at will like B and C, the incoherence is not at all straightforward, even when one does specify the reasoning involved in acquiring the belief. These seem like plausible accounts of how one might acquire a belief. A contradiction is only apparent and the incoherence alleged in the believing subject is only forthcoming if one has an existing theoretical commitment to a conceptual connection between belief and evidence that precludes such a belief acquisition. But this is just to beg the question against the voluntarist.

Adler's incoherence test is misleading, particularly in cases of direct control, because it erroneously generalizes the incoherence found in case A to other cases. Certainly, as conceded in Chapter 3, it is incoherent for a subject to believe against the strong tide of the evidence, or in complete absence of evidence. But the terse form of Adler's incoherence test, with its contrastive conjunction, "*p*, but I believe *p* because I choose to" implies just this kind of case. This (plus the assertion of full belief) is the source of the intuitive feel of incoherence that Adler correctly identifies. But, as we see when we make the reasoning explicit, the kinds of cases where true incoherence exists are irrelevant to the question of believing at will. For these reasons, we ought to reject Adler's incoherence test as a guide to the possibility of direct control over belief.

I would suggest that we now possess sufficient reason to reject Adler's conceptual argument against direct control over belief. However, I mentioned previously that Adler makes an additional qualification to his conceptual argument that would lead us to a discussion of his instability argument. This qualification is connected with examples like case C above. Adler's concern has to do with the phenomenon of making up one's mind. He asks us to consider the case of Irene, who has borrowed a book from me and appears to have lost it. The evidence in the circumstances is not conclusive, but leans in this direction, though personal and other factors mitigate against my believing this. Moreover, for certain practical reasons, I need to make a decision on this issue, and cannot just put it aside.¹⁴⁹ In such an instance, Adler admits that I can make up my mind in favor of the belief, which "can be a voluntary act." "With qualifications to come," he writes, "I admit that to judge all-out that Irene lost my book is to satisfy the conditions for believing that she did."¹⁵⁰ But this seems to be an admission that one can, in certain circumstances, voluntarily decide to believe. How can this be consistent with Adler's earlier claims?

Adler recognizes the problematic nature of this admission. He responds by suggesting various reasons that we ought not to think that voluntarily judging that something is true is the same as believing at will. He suggests ways in which we might see this action as only indirectly coming to believe. Perhaps the judgment is really one to close inquiry. Or perhaps the decision is only one to induce belief, not a decision to

¹⁴⁹ As many readers will recognize, this scenario sounds much like those that William James focuses on, situations where an issue is "living", "momentous", and "forced." See James (1896), especially pp. 3-4

¹⁵⁰ Adler (2002), p. 61.

believe. The bulk of Adler's response to this concern, however, is his instability argument. Having accepted the possibility of making up one's mind in favor of a belief, and offering an account of this that sounds suspiciously like deciding to believe, Adler attempts to argue that such a belief would be essentially unstable. This instability argument, modeled after Williams' second argument, asserts that it would be impossible to *retain* a belief formed at will in this manner.

The precise character of the instability that Adler alleges is somewhat unclear. His initial account matches Williams', where the instability is caused by a subject's recognition of the general connection between belief and evidence, and simultaneous awareness of the lack of evidence supporting this particular belief. In support of this characterization, however, Adler cites Daniel Dennett's examples of "brain writing", where false beliefs are inserted directly into the mind of a subject.¹⁵¹ The example Adler mentions concerns Tom, in whom is inserted the false belief, "I have an older brother living in Cleveland." Dennett argues, and Adler agrees, that any such instance of brain writing would quickly fall prey to reality. As the various incongruities between the actual world and the inserted belief arose (such as the fact that I don't know anything else about my alleged sibling), the belief would quickly become unstable, and would shortly be rejected by a rational subject.

Hold apart for a moment the fact that this case has no bearing on whether or not it is possible to exercise direct control over belief. The example also suffers, as an analogy

¹⁵¹ Ibid., pp. 62-63.

for believing at will, from the fact that it deals with a belief stipulated to be false. But of course there's no reason to suppose that all beliefs formed at will would be false. The example then relies for the alleged instability not on incongruities concerning evidence specifically, but rather incongruities caused by the falsehood of the belief. But in cases where one chooses to believe something that is in fact true, such incongruities will likely not be present. If anything, it is likely that the belief will receive confirmation from reality, rather than confusing disconfirmation. And even if we do consider a willed belief that is false, the instability pointed to here is instability caused not by the source of the belief in the will, but rather by the belief's falsity. Dennett's example thus offers no support to Adler's claim that willed beliefs are essentially unstable.

Adler then returns to the example of Irene to support his allegations of instability. Perhaps, he suggests, the plausibility of deciding to believe that Irene lost the book was based on the fact that the case was not open to resolution by evidence, and yet required action on my part. If we remove the former constraint, and suppose that I will find out the book's whereabouts in 12 hours, we can see the instability resurface. This instability, Adler argues, is shown by considering again my assertion in such a case, "Irene lost my book, although in 12 hours I may learn otherwise." Such a belief formed at will is, Adler says, inherently unstable in the face of potentially decisive future evidence.

Again, it is unclear how Adler's claims in this case bear directly on the possibility of believing at will, other than perhaps to argue that such beliefs could play no significant, ongoing role on one's cognitive economy. More importantly, however, is the

fact that Adler has incorrectly diagnosed the source of the incoherence in the assertion “Irene lost my book, although in 12 hours I may learn otherwise.” Any incoherence in such a statement arises not from any problem deciding to believe in the face of potential future evidence. The problem is instead between forming a full belief, whether (*per impossible*) at will or through some other mechanism, and the possibility of decisive future evidence. That the problem is with full belief is shown by the naturalness of the assertion incorporating a qualified form of belief, such as “I *believe* Irene lost the book, although in 12 hours I may learn otherwise.” I believe all sorts of things while aware of the possibility that the belief very shortly may be disconfirmed. Such a belief, explicitly recognizing a certain level of uncertainty, has no difficulty with the explicit possibility of future disconfirmation. Yet it remains belief, and remains open to initiation through an act of decision, of making up one’s mind. There is no problem with believing at will in the face of future evidence. It is perhaps incongruous to form or hold a full belief while simultaneously focusing on the possibility of its being decisively undermined in the near future. Simply the awareness that such decisive evidence is forthcoming is inconsistent with such certainty. This is why Adler’s stronger assertion rings strange in our ears.

Discussing these cases has taken us somewhat afield of the kind of instability initially alluded to by Adler and posited by Williams, the instability allegedly caused by exercising direct control over belief.¹⁵² But we did address such incongruity earlier in our

¹⁵² The focus on individual cases may seem odd in another respect as well. Adler notes that, given he is making an explicitly conceptual argument, the reliance on cases may seem inappropriate, or unequal to the task of showing that believing at will is conceptually impossible. Adler suggests (pp. 61-62) that the role of

discussion of Williams and saw how, as a conceptual argument demonstrating the impossibility of believing at will, it leaves much to be desired. The above discussion indicates that Adler does not meaningfully advance beyond Williams' position. I conclude, therefore, that Adler's instability argument fails to show how direct control is impossible. And given that his initial conceptual argument also fell far short of this goal, leaving the argument of the previous chapter basically intact, we may conclude that Adler has not offered a convincing response to the thesis of moderate direct control.

4.6. Audi

I turn next to arguments made by Robert Audi against believing at will. As I read him, Audi makes both psychological and conceptual arguments against believing at will. I consider his psychological arguments in my next chapter, though I confess it is sometimes hard to keep the two apart in discussing Audi's claims. Audi distinguishes two conceptions of doxastic voluntarism that are his targets. *Behavioral voluntarism* is the view that belief itself is a type of action that is sometimes voluntary. *Genetic voluntarism* holds instead that only the formation or initiation of a belief is sometimes a voluntary action, as opposed to the belief itself. Both forms of doxastic voluntarism are ultimately mistaken, Audi argues. Audi appears to take this division as exhausting the possibilities

individual cases in his argument is not to show the impossibility of believing at will, but rather to clear away alleged counterevidence to his conceptual thesis. Whether this is all the work such cases are doing is questionable, but not particularly important given that, as we have seen, most of the individual cases Adler considers offer at best ambiguous support to his argument.

for doxastic voluntarism, and thus if neither position is viable, the idea of believing at will ought generally to be rejected.

Audi begins his conceptual arguments by noting certain differences between theoretical and practical reasons, reasons for believing and reasons for acting. He argues that reasons for believing are always propositions, and as such have truth values. Audi explains seemingly obvious counterexamples, such as phenomenological experiences at the base of perceptual beliefs (ex. “being appeared to greenly”), by arguing that such experiences are more properly understood as grounds for belief rather than reasons for belief. Reasons for action, on the other hand, are in general not most appropriately expressed in terms of propositions, and are not truth-valued.

These differences in reasons, Audi contends, point to larger and more important differences concerning practical and theoretical reasoning, action and belief. Audi makes three related claims about belief, all of which he thinks show the impossibility of doxastic voluntarism:

- (A1) Beliefs are formed in a causal process which flows from world to subject; actions are executed in a causal process that flows from subject to the world.
- (A2) Beliefs are not events, and thus cannot be actions.
- (A3) Beliefs purport to represent reality, actions seek to change reality.

I will consider the arguments supporting each of these claims in turn. The argument for A1, which concerns the causal genesis of beliefs and actions, proceeds roughly as follows. There is a certain causal flow to practical reasoning. In practical reasoning and intentional action the causal sequence flows out from the agent to the world, directed at

modifying the world to fit the desire or desires which helped formulate the intention. In theoretical reasoning, the causal flow is reversed. Beliefs and theoretical reasoning function to reflect the world, to fit the mind to what is actually there. Since they reverse the causal flow of actions, beliefs cannot be actions, and thus cannot be subject to direct control.

Audi's claim A2 is that the very nature of action is further evidence against doxastic voluntarism. Actions are events, beliefs are not. Actions, like my typing the letter "Q", are events that occur in determinate spatiotemporal locations. The most plausible conception of beliefs, on the other hand, is that they are states of the mind. But as a state of the mind, a belief is not an event, and thus is not even a candidate for being an action. This insight explicitly dooms behavioral voluntarism, which implausibly sees belief itself as an action.

Audi says little in support of A3, though it is intertwined with his discussion of A1. Of course, the claim that beliefs purport to represent reality is far from controversial, and one that I've previously noted my agreement with. The specific implication of the claim, as I will note below, is difficult to tease out, though the general point seems clear. On Audi's interpretation, actions do not represent the world, but instead modify it. Thus it appears to be that the representational nature of beliefs makes them poor candidates for characterization as actions.

In response, it seems clear that Audi's denial of behavioral voluntarism seems correct. Here he, like both Williams and Scott-Kakures, helps us to see more clearly what

the voluntarist should *not* hold. On either an occurrent or dispositional account, belief seems more like a state than an event that could plausibly be construed as an action. To use the example again, possessing a belief is more like the state having one's arm up in the air than it is like raising one's arm. So we ought to reject behavioral voluntarism. But behavioral voluntarism was never particularly plausible to begin with, and has little to do with the argument of Chapter 3 or the thesis of moderate direct control more generally, which clearly is a form of genetic voluntarism on Audi's categorization. Do Audi's claims cause any concern for this argument, or genetic voluntarism more broadly?

Looking again at Audi's three claims, it is immediately obvious that A2 has no purchase with genetic voluntarism, which holds only that belief formations are actions. Since A2 concerns only beliefs or possession of a belief, it has no bearing on this thesis. While A1 and A3, on the other hand, are more relevant to genetic voluntarism, they also have a familiar ring to them. Concerns about the causal origin or development of beliefs, as well as their representational character, were at the heart of Williams' three arguments against believing at will. But how precisely ought we to understand A1? The claim requires interpretation to make its significance clear. We could take it to simply mean that external influences beyond our control play an essential role in the causal etiology of acts of belief formation. But this innocuous claim is no threat to doxastic voluntarism; influences from the external world figure prominently in the production of other actions as well without destroying control. Interpreting the claim more strongly, we might read it as arguing that beliefs are typically produced by such influences in a causally mechanistic

fashion that belies the possibility of control. On this gloss, I would suggest that A1 is most plausibly interpreted as a form of Williams' third argument based on a functionalist conception of belief. As we saw above, such arguments pose little concern to moderate forms of voluntarist direct control. To eliminate the possibility of believing at will requires an improbably broad reading of the functionalist requirement.

A3 is also ambiguous, but it similarly echoes arguments made by Williams, and roundly critiqued above. Why would the representational character of beliefs be a worry for doxastic voluntarists? One reason, suggested by Williams, is that beliefs are closely connected with truth. As the discussion of Williams' arguments showed, however, only those who hold to a fairly radical conception of believing at will need be concerned with the fact that beliefs are connected with truth or attempt to represent reality. Only if one is attempting to will to believe some proposition that one clearly thinks is false will this be a problem. But the voluntarist ought not to commit herself to this extreme form of control. If she does not, then Audi's argument on this matter is of little concern. This leads us to the conclusion that none of the claims in Audi's first argument has any purchase against genetic voluntarism.

4.7. Conclusion

We have seen in our survey of conceptual arguments against direct control that they orbit around a cluster of ideas concerning the nature of belief, the nature of intentional action, and the truth-directedness of belief. I have tried to show in my critical

interaction with this literature not just that these arguments fail, but specifically that none of the conceptual arguments offered by our four authors seriously threatens the case for moderate direct control over belief on a voluntarist conception. The broader task of defending direct control is only half-finished at this point, however, as we have not yet interacted with the other strand of argument in the literature criticizing believing at will. I therefore turn in the next chapter to the psychological arguments made by Audi and Alston.

CHAPTER 5 – VOLUNTARISM: PSYCHOLOGICAL ARGUMENTS AGAINST DIRECT CONTROL

5.1. Incoherence vs. Implausibility

In the previous chapter I considered conceptual arguments against direct control on a voluntarist conception. These conceptual arguments attempt to show that believing at will is incoherent, that it is incompatible with the very concept of belief, or intentional action. The question under consideration was whether any of these arguments poses a threat to the thesis of moderate direct control, which again is the idea that the determination of beliefs is subject to direct control in much the same way that actions are. We found that many of the conceptual arguments offered attack an over-inflated notion of control that the doxastic voluntarist need not support. I argued that many of the arguments made against direct control suffer from variants of the same flaws found in Bernard Williams seminal arguments against believing at will. In the end, I argued that none of these arguments causes significant difficulty for the thesis of moderate direct control over belief.

In the present chapter I take up the other line of argument regularly offered by critics of believing at will. These arguments are principally concerned with the psychological plausibility of direct control, how likely it is that we can exercise direct control over our beliefs given the most plausible conceptions of human psychology. Again in this chapter I will attempt nothing like an exhaustive survey of the relevant

literature; I will rather concern myself with considering two prominent and representative examples of psychological arguments against control, made by Robert Audi and William Alston.¹⁵³ Also, as in the last chapter, my specific concern will be with establishing that the psychological arguments under consideration offer no serious threat to moderate direct control, and it will be to this concern that I direct the bulk of my comments.

5.2. Psychological Arguments

As was the case with the conceptual arguments, I begin with a brief discussion of what a psychological argument against direct control over belief attempts to establish, and what constitutes a refutation of such an argument. Psychological arguments are typically somewhat less ambitious than their conceptual counterparts. These arguments do not seek to establish that direct control over belief is impossible as such. Instead, psychological arguments seek to establish that direct control over belief is impossible or improbable for us, for human beings, given the constitution of human psychology. Such psychological arguments typically consist of two parts: a brief discussion or gloss of the relevant aspects of human psychology, and a discussion of cases of putative direct control. The discussion of our psychology serves as a theoretical anchor for the discussion of individual cases. The consideration of individual cases typically involves offering alternative interpretations of alleged instances of deciding to believe,

¹⁵³ Audi (2001) and Alston (1988), respectively.

interpretations consistent with an involuntarist analysis of belief and the previously articulated claims about human psychology.

A refutation of such an argument may consist of a defense of the idea of believing at will, or an attack on the argument against direct control, or preferably both. It is important that such a refutation include a critical discussion of the conception of human psychology offered by the critic, and I will do so with each of our two authors. Given that the claims of psychological arguments are not as categorical as those of the previous conceptual arguments, it will not typically serve to simply articulate a counterexample to the arguments. Broadly speaking, one needs to show that the support for believing at will is superior to the argument offered by the critic. I take myself to have articulated the voluntarist case for direct control in Chapter 3; here I will focus on critical examination of the arguments against direct control.

Given that interpretation of cases plays such a key role in our analysis, I must begin by drawing attention to a significant disagreement with both Audi and Alston as to the dialectic concerning psychological arguments against believing at will. The specific issue concerns where the burden of proof lies. Reading both Audi and Alston one gets the impression that putative instances of direct control over belief are guilty until proven innocent. This despite the fact, which we discussed at some length in Chapter 1 and which both Alston and Audi admit, that our widespread practice is to speak of beliefs as subject to direct control, as when someone says, "I decided she was lying to me," or, "I decided that Smith's theory would not withstand experimental testing." I would suggest

that Audi's and Alston's approach is incorrect and misconstrues the relevant dialectic. Part of my case in Chapter 1 was the claim that we have a persistent, pervasive tendency to act as if some beliefs are subject to our direct control. This fact seems to me well established. And it seems to me, as I remarked previously, that when we encounter a widespread phenomenon with a *prima facie* straightforward explanation, this explanation ought to be afforded a certain deference. If individuals commonly report that they embrace or assent to certain beliefs through acts of decision, then the burden of proof is on the critic to show why this is not possible, not upon the supporter of doxastic voluntarism.¹⁵⁴ This does not remove the need to answer the critics' arguments, but it does affect how we receive these arguments in the first place. Direct control over belief starts with intuitive support and Audi and Alston must make the case against direct control, not simply gesture at potential alternative interpretations of the facts. But much of what both of these authors engage in appears to consist simply of offering alternatives to alleged instances of direct control. I am suggesting that this alone is not sufficient.

Let us turn now to the specifics of the arguments. Alston's and Audi's psychological arguments against direct control over belief share much in common. Both examine the putative "best data" for doxastic voluntarism, and argue it fails to provide adequate support for the thesis that one can sometimes believe at will. Both suggest that

¹⁵⁴ The fact that Audi is correct to criticize behavioral voluntarism has no significance for my claim here. All of the examples that Audi criticizes are ambiguous as to whether the subject is expressing a behavioral or a genetic voluntarist position, though they are not ambiguous as to whether the subject is expressing support, at least *prima facie*, for voluntarism. And if one rejects (as one should) a behavioral voluntarist approach, genetic voluntarism is clearly a preferable interpretation of the statements under consideration than is an involuntarist interpretation, barring an independent argument against voluntarism.

alleged cases of direct control are more appropriately understood as instances of indirect control. And both articulate similar principles of control, principles they claim are satisfied with respect to action, but not with respect to belief. For the sake of continuity with the previous chapter, I begin with Audi. As we have conceded the falsity of behavioral voluntarism, and identified genetic voluntarism with the basic argument of Chapter 3, I will examine the impact of Audi's arguments only on moderate direct control.

5.3. Audi Redux

Audi's psychological argument consists of two parts. The first focuses on analyzing the phenomenon of belief formation. Genetic voluntarism, unlike behavioral, argues only for the formation of a belief as a potentially voluntary action. This initially seems more plausible, Audi concedes. But how precisely should we understand the process of belief formation? Audi argues that forming a belief could be interpreted in at least two different ways: simply coming to believe, or causing oneself to form a belief. The former includes cases such as when I see a red fire truck on the street, and come to believe that there is a red fire truck on the street. As noted in Chapter 1, such instances of perceptual belief formation seem comparatively involuntary, and thus not good candidates for believing at will.

Causing oneself to believe, on the other hand, may be voluntary, but is not direct in the way necessary for believing at will. Typically, Audi says, causing oneself to form a

belief is not an act of belief formation, but rather an act that *entails* belief formation. This type of indirect action, however, is disanalogous with standard examples of direct action, such as raising one's arm. Raising one's arm is just an act of raising one's arm, not an action that entails raising one's arm. Causing oneself to believe, on the other hand, is more roundabout than the picture of control typically proposed by doxastic voluntarists. It may represent a form of control over belief, but it is not a direct form of control. But if these two are the only possibilities for belief formation, then genetic voluntarism is hopeless as well. And assuming that Audi has correctly divided up the landscape, having already rejected behavioral voluntarism, we ought to reject the possibility of believing at will.

The other half of Audi's argument consists of an examination of examples of belief where voluntarism's case seems strongest, such as cases of deliberation over evidence leading to belief. He asks the reader to consider a hypothetical interaction between a group of persons discussing another's testimony. Apparently, there is disagreement over whether the individual in question is telling the truth. Here, Audi suggests, several plausible responses to the testimony seem to imply a voluntarist interpretation:

To the question 'What did she do that so upset you?', the reply 'She believed her lying husband's fabrication' is an admissible answer. In the same context, another person might say, 'For my part, I can't decide whether to believe him or not; I'm still deliberating about whether his story is true.' Still another might say, 'I rejected his testimony as fast as she accepted it.' And Descartes famously spoke of *assenting* to a

proposition as a kind of thing we do (presumably at will) that entails coming to believe the proposition in question.¹⁵⁵

While the responses may invite a voluntarist interpretation, Audi argues that no example of this type *demand*s the analysis that it is an instance of believing at will. All such allegedly voluntarist cases can be understood under an involuntarist interpretation consistent with his theoretical arguments, both conceptual and psychological.

Audi attempts to establish this claim by distinguishing between what he calls positive and negative control over a state of affairs, and claiming that we lack the essential positive form of control over belief. Positive control consists in the ability to bring about a state of affairs, and negative control is the ability to prevent a state of affairs from occurring. Direct basic control, like that exemplified by raising one's arm, includes both positive and negative control. Audi's meaning here is a bit opaque, but what he seems to be saying is that in many circumstances, an individual has the ability to either raise her arm, or to refrain from doing so, and can bring about either state of affairs directly.¹⁵⁶ Those instances which seem initially to offer support for doxastic voluntarism, however, are more plausibly understood as consisting of at most negative control, Audi says. This control by itself is insufficient for believing at will. With regard to believing, negative control would only allow us to prevent the formation of a belief, or to permit it to be formed, but not to form it directly at will. We can see that negative

¹⁵⁵ Audi (2001), p. 94.

¹⁵⁶ One way in which Audi's meaning is unclear concerns whether his conception of positive control should be interpreted in a compatibilist or libertarian incompatibilist. I consider the implications of both interpretations below.

control is insufficient, Audi suggests, by considering by analogy involuntary events like sneezing or twitching.¹⁵⁷ We have negative direct control over such events; we can (sometimes) suppress a twitch by an act of will. But the twitch itself is not a voluntary action; it is not an action at all. We do not hold each other responsible for such events, because we see them as largely out of our control. Audi suggests that belief fits in a similar category of events.

Consider again a case of deliberating and deciding to believe, like the fictional wife above, who is being held responsible for allegedly voluntarily believing her husband's lie. Initially this seems a counterexample to Audi's earlier conceptual arguments. But on the view put forward above by Audi, such a case is more plausibly interpreted as deliberating and ultimately yielding to the evidence, which itself exerts the primary causal force towards belief. This is not believing at will, Audi argues, but rather letting a belief come about by an act of will. But then it is not the positive control exemplified by the arm-raising example. It is rather at most a more limited form of control. And if this is all we possess over belief, Audi suggests, then doxastic voluntarism is mistaken.

Before attempting to rebut Audi's claim, I think it is worth noting that even were we to concede both that Audi's distinction between positive and negative control is clear and intelligible (something I question below), and that his claim that we possess only

¹⁵⁷ I'm inclined to use the phrase "involuntary actions" here, but Audi uses "events" in discussing these cases, and I don't wish to prejudice the case against him.

negative control is correct, this would not undermine the main points that I am seeking to establish. Negative control, as Audi conceives of it, is still a direct form of control, and seems clearly to be a type of control for which we are reasonably held responsible. While limiting our control over belief to negative control might introduce interesting further distinctions between belief and action, and might overturn certain conceptions of believing at will, it will nevertheless not undermine at least a broad reading of moderate direct control. Recall that moderate direct control simply says that the formation and dissolution of beliefs are sometimes subject to control in much the same way as actions are. Here we would see a new way in which the control over action is more expansive than that over belief. But since, as Audi himself argues, we do plausibly possess negative control over belief, the main thesis would remain intact. Even the specific corollaries 1 and 2 of the thesis would remain intact if we were to concede Audi's points. Claim 1 indicates that our beliefs are responsive to our judgments about what we ought to believe, which is certainly still true where there is negative control. And claim 2 says that in certain cases it is possible for one to decide what to believe. If we allow that deciding what to believe includes deciding what not to believe, this is true even when our control over belief is limited to negative control. So even if completely successful, Audi's psychological argument does not undermine the basic thesis of my project.

Returning to the main line of argument, however, the most straightforward reply to Audi's reasoning is simply to deny that he has exhausted the possible modes of belief formation. It is certainly true that we sometimes form beliefs involuntarily, and we

sometimes cause them to be formed voluntarily but indirectly. But we also sometimes form beliefs directly by an act of decision, when we embrace or assent to a proposition concerning some matter. Here the formation of a belief is by the direct act of believing intentionally, or at will. The arguments of Chapter 3 offer support to this claim, which falls within Audi's notion of genetic voluntarism. To be fair, Audi is aware that he cannot simply dismiss the possibility of forming a belief at will without argument, and he offers a numbers of claims in support of his rejection of this possibility. Having questioned Audi's claims about belief formation, we must now show what is wrong with his interpretation of the cases.

Here we need to examine Audi's distinction between positive and negative control somewhat more closely. Serious questions exist about the plausibility of this distinction. Audi suggests that we possess both positive and negative control over actions, but only negative over belief, against the claims of genetic voluntarism. Rather than discussing abstractions, however, I'll instead focus on concrete examples from both the practical and theoretical realms. In the practical realm, consider the state of affairs of my writing my name on a job application. In the theoretical realm, consider the state of affairs of my believing my friend John's explanation for failing to pick me up for work.

Audi will suggest that if my writing my name is voluntary, then I exercised positive direct control over it. But he will claim that in the latter case I cannot possess a similar form of control, despite what I might say about deciding to believe John's account. One way to broach my concern here is to ask for a further account of the nature

of the positive direct control alleged in the case of action.¹⁵⁸ Possible accounts, as discussed in Chapter 3, can be divided up along compatibilist and libertarian lines. Consider first the compatibilist position. For the compatibilist, positive control cannot consist in my causally initiating the decision which leads to the action. My decision will necessarily possess a causally sufficient explanation that includes my decision but does not initiate with it; the causal chain resulting in my action extends deterministically back outside of me. For the decision and action to be voluntary, this causally sufficient explanation must minimally line up appropriately with my putative reasons for acting. But on this account, it is difficult to see how the story of control over action is relevantly different from that of belief. For both phenomena, the resulting state is causally determined, not just influenced, by factors outside myself, over which I possess no control. And in both cases, the resulting state is generated by my responsiveness to certain kinds of reasons.

Herein lies an important related point. Audi wants to compare believing to twitching or sneezing, two clearly involuntary activities. But here is a clear difference between such involuntary activities and rational belief. Beliefs are formed for reasons, sneezes and twitches are not. If I yield to a belief, it is to the force of reasons which move me. Sneezing and involuntary twitching both occur in a very different way. When I suffer

¹⁵⁸ Though I came to same the conclusion independently, this way of formulating this argument is heavily indebted to Steup (2001), who makes the point quite articulately on pp. 11-16 of his work.

from an involuntary twitch, I am compelled by sub-rational urges, or automatic physical processes disconnected from rational considerations. The same is true with sneezes.

Now, one might suggest that we sometimes decide to give into a sneeze for a reason. Say, for example, that my son feels the urge to sneeze coming on, and begins to resist it (perhaps he's playing a video game and doesn't want to ruin his chances). Assume that his resistance would have been successful. But then he quickly recognizes that sneezing while his mother is walking by increases his chances of staying home from school. With this thought, he stops resisting, and the sneeze erupts. Assume for simplicity's sake that he would otherwise have been successful in his resistance. Here he seems clearly to have sneezed for a reason. But of course here we would not call his sneeze wholly involuntary; here he clearly chose to sneeze. And the difference is marked by the role of reasons in the sneeze's occurrence. Rational responsiveness or susceptibility to reasons marks a clear line in the procession from the involuntary to the voluntary, especially for the compatibilist. And belief is clearly more like action than involuntary occurrences with regard to this consideration.

The only other possible distinction between action and belief on the compatibilist account seems to be found in their respective phenomenology. Audi claims that the reports of deciding in favor of a belief are illusory; I find this claim itself dubious. But even if we were to grant him the claim, we still cannot draw a clear line between the phenomenology of belief and action with respect to decision. Would Audi want to claim that my experience of yielding to the tremendous desire to kiss my wife, in spite of my

anger towards her, indicates that this action is involuntary? What about my yielding to the desire for chocolate cake set before me? Many actions possess the phenomenology of yielding to powerful desires and/or reasons. But they are nevertheless often voluntary, involving one's choice to perform the action. Here choice is essential to the performance of the action, but like in the cases of belief cited by Audi, the primary motivation comes from the reasons that influence one's decision. In both cases, I am a channel through which reasons flow through to particular outcomes, either actions or beliefs, respectively.¹⁵⁹

Given the various considerations noted here, it is hard to see how the distinction between positive and negative control stands up for the compatibilist. Audi's argument, however, is predicated upon this distinction and the lack of positive control in the realm of belief. And it is thus difficult to locate a meaningful distinction in control between belief and action on the compatibilist view.

The libertarian, on the other hand, can perhaps lay claim to a more robust distinction between positive direct control over action and the derivative control over belief. Here I can initiate an action, in the sense that it is not determined by factors outside of me. I certainly may be influenced by external factors, but the causally sufficient explanations of at least some actions find their genesis in my free choice. For the libertarian, this control is typified by the fact that, for at least certain voluntary

¹⁵⁹ Paul Hoffman and others have suggested to me that I have not gone far enough here, and that it is plausible to think that even certain cases of compulsion, as long as they involve rational compulsion, are voluntary.

actions, I have both the ability to bring about the action and the ability to not bring it about under the same set of conditions, or holding the past fixed. I can either write my name on the application, or refrain. But again, it is hard to see how this two-way power is lacking for belief, given the libertarian account. Even if we accept Audi's insistence that we at best yield to the reasons for belief, this still seems like a decision, one I could make or refrain from making. The very term "yield" implies a decision to relent. There are assuredly some beliefs for which the evidence is overwhelming or truly decisive, and I recognize it as such. On the libertarian picture in these cases it seems reasonable to think that I do not yield to the evidence, here my assent is compelled. But not all beliefs are this way, for many the evidence is far short of compelling. Here my assent is subject to my choice.

Perhaps Audi would respond here that we still have not established positive control over belief, only two-way control. But what then is positive control to be distinguished by? If we turn to our phenomenology again, the same points made for the compatibilist case will hold. Some voluntary actions are experienced as yielding to one's reasons or one's desires. And we still lack a compelling reason to dismiss the widespread claims to belief initiation through decision discussed in Chapter 1. Recall that we were looking here for support for the claim that deciding to believe is not possible, because such support was lacking in Audi's conceptual arguments. I suggest that it is similarly lacking from his psychological arguments.

Audi wishes to differentiate between the psychological processes when one “yields” to a belief, versus those when one decides and acts. But any plausible account of how the decision is made blurs the lines between the allegedly different forms of control. Audi’s distinction between positive and negative control ultimately fails to locate a relevant difference between belief and action. Audi seems to recognize this worry for his psychological arguments, saying the following in the course of making them:

Indeed, I grant that despite all I have said there is still a tendency to think that in the context of our willfully concentrating on evidence, particularly when it seems to us equally divided, the cognitive upshot of the process, namely belief formation, looks much like an action we might take on the basis of practical reasoning. Why not say that for certain cases, especially those in which the reasons leave us free to act or not and to believe or not, it is by an exercise of will that we act or believe, and here believing is voluntary in much the same way as action?¹⁶⁰

Audi’s answer to his question is that his claims are reinforced by his conceptual arguments, which we considered in the previous chapter. But as we saw in the previous chapter, these arguments establish only that behavioral voluntarism is false, and have little impact on the genetic voluntarist position advanced in Chapter 3, in support of moderate direct control.

¹⁶⁰ Audi (2001), p. 98.

5.4. Alston

We turn now to our second interlocutor, William Alston, and our final set of arguments against direct control over belief. As noted above, Alston's and Audi's arguments share much in common, focusing on individual cases, and attempting to establish that belief fails to conform to a plausible criterion of direct control. Alston initially argues that it simply seems abundantly clear that we lack the requisite control over belief, though, he concedes, this is a contingent fact. He is characteristically direct. "Volitions, decisions, or choosings don't hook up with anything in the way of propositional attitude inauguration," he states, "just as they don't hook up with the secretion of gastric juices or cell metabolism."¹⁶¹

This is vivid prose, but it is not clear that the examples that Alston marshals in support of this claim are relevant to the case. Alston gives us an idea of the kind of cases he has in mind, saying, "My argument for this, if it can be called that, simply consists in asking you to consider whether you have any such powers. Can you, at this moment, start to believe that the U.S. is still a colony of Great Britain, just by deciding to do so(?)"¹⁶² But as we've discussed previously, why should we think that the fact that one cannot believe something for which one has no reasons, or worse yet, something which all available reasons contradict, gives us insight into control over belief? I have already made this point with respect to Audi's arguments, so I will not belabor it here, except

¹⁶¹ Alston (1988), p. 263.

¹⁶² Ibid.

simply to note that Alston's claims here have little significance for the thesis of moderate direct control, the claim that we possess not unlimited control over our belief formation, but only control similar to that which we possess over our actions. Given that we cannot act against overwhelming reasons, why should we suppose we can believe in this way?

Alston does offer arguments more directly opposed to moderate direct control, however. He articulates a criterion of direct control, which he claims actions satisfy but beliefs fail to. Alston's criterion seems more explicitly libertarian than Audi's. He argues that one does not have direct control over some state of affairs *A* unless one is capable of directly bringing about either *A* or some other incompatible state of affairs, $\sim A$.¹⁶³ In the paradigmatic case of action we've noted throughout, I am free to either raise my arm, free to not raise it, or perhaps free to twist it behind my back. For believing, he suggests, this principle indicates that we lack direct control. Most obviously in favor of Alston's claim are those beliefs, such as perceptual beliefs, that are both immediately formed and obviously true or obviously false. I look outside, see the fire truck, and the belief that there's a red fire truck outside forms. In these cases, one lacks the ability to refrain from believing, at least directly, and thus cannot be said to believe at will.

Of course, most doxastic voluntarists would happily concede these cases, and argue, as I did in Chapter 3, that one's ability to believe at will extends only to those cases where the weight of the evidence or reasons is not compelling (a constraint

¹⁶³ Ibid., p. 261. Alston's precise statement is, "To have control over believing that *p* is to have control over whether one believes that *p* or not, i.e., over whether one believes that *p* or engenders some incompatible alternative."

similarly borne by one's ability to act). Theoretical beliefs, beliefs concerning scientific, historical, interpersonal, religious matters and the like, cases where one engages in deliberation and weighing of complex evidence before coming to the judgment that one proposition has greater support, these seem paradigm cases of belief potentially subject to control. These cases seem far removed from garden variety perceptual beliefs, and much closer to the model of deliberative voluntary action.

Alston makes two points in response to such cases, both of them germane to this discussion. First, Alston argues that even if he concedes that these beliefs are subject to voluntary control, this would not be sufficient to support a deontological conception of justification for belief. Such deliberative beliefs make up only a fraction of our doxastic superstructure, the bulk of which are instead beliefs that are either obviously true or obviously false. These tight limits on the scope of beliefs potentially under our control are inadequate, Alston claims, to support a generalized deontological conception of justification.

Here I would agree, to a point. I am seeking to support the use of responsibility language with regards to belief. I have heretofore made no explicit commitments to the idea that a deontological conception is the best or only way to cash out this language, even with respect to justification, though it is certainly one of the most natural. But I am arguing that the use of this language is grounded in our control over our beliefs. Direct control is one form of control that we enjoy. As I have discussed throughout this dissertation, only some fraction of our beliefs (varying dependent on the account) are

such that we might possess direct control over them. I certainly agree with Alston that this cannot be the whole story with respect to doxastic responsibility. In a complete account of responsibility for and control over belief (something beyond the bounds of this project) an account of the various forms of indirect control would have to be advanced.

Alston's primary target is this deontological conception of justification, and he advances arguments against indirect control as well. I do not consider these arguments here as they lie outside the immediate scope of this project. Ultimately I disagree with Alston about how wide the scope of beliefs subject to direct control is; I think he significantly underestimates the possible range of such beliefs. But I certainly accept that this form of control cannot account for all the ways in which we attribute responsibility to each other for our beliefs.

Alston's second claim is directed at the argument that the cases of deliberative and theoretical belief above could be instances of believing at will. There are two possible alternatives in these types of cases, he suggests. The evidence may be supportive (though only slightly) of one proposition over another incompatible alternative, or the evidence may be equally persuasive or unpersuasive for both alternatives. In the first case, Alston argues, we can see that the control condition is not satisfied. A philosopher who adopts theism because she believes the evidence generally favors it, he argues, is no

more capable of adopting atheism than one who finds the evidence for theism utterly compelling.¹⁶⁴

On the other hand, in cases where there is true evidential equipoise, the voluntarist finds no more support for believing at will. Beliefs, Alston argues, are held because they are seen to be more probable than not, or more probable than their alternatives. A decision to agree to one proposition over another when there exists no superior evidential support for the former is not a belief. Instead, Alston claims, it is more plausibly understood as a form of acceptance, for one of various practical reasons. One could be accepting the proposition for the sake of argument, or in order to ultimately engender belief within oneself through a more roundabout method. But one is definitely not believing at will.

Concerning the case of equipoise, I confess I am unclear precisely what to say. One thing that seems certain is that such cases, where one's evidence is truly indifferent to two incompatible propositions p and $\sim p$, are quite rare. Typically, even when the evidence is vexed, we find ourselves leaning one way or another. But in such a case as Alston describes, perhaps we are like Buridan's ass, torn between two equally attractive or unattractive options, with no evidential basis for deciding on one or the other. Of course, this seems to indicate similarities between belief and action, rather than dissimilarities. But even if it turns out that one could choose to act in such cases of true

¹⁶⁴ For this example, see Alston (1988), p. 266.

equipose, but not to believe (recall my discussion of other such potential differences in Chapter 2), this does not seriously undermine the case for moderate direct control.

In the other case, where the subject is inclined one way or the other, I suggest we need to carefully consider Alston's criterion of direct control. Recall that Alston requires bi-directional control over beliefs in order for believing to count as direct basic action. When initially outlining his control condition, Alston correctly notes that the incompatible alternatives to believing p include at least believing $\sim p$ or withholding belief concerning p . However, when he discusses deliberately acquired beliefs, Alston's scope of incompatibility changes. Discussing the case of religious belief mentioned above he claims:

Thus when our philosopher or religious seeker "decides" to embrace theism or the identity theory, what has happened is that at that moment this possibility seems more likely to be true, seems to have weightier considerations in its favor, than any envisaged alternative. Hence S is, *at that moment*, no more able to accept atheism or epiphenomenalism instead, than he would be if theism or the identity theory seemed obviously and indubitably true.¹⁶⁵

But of course atheism is not the only incompatible alternative with theism. Agnosticism and the possibility of withholding belief on the matter is another alternative which Alston here overlooks. Alston's example draws much of its plausibility from considering only the two polar opposites that he mentions. It is reasonable to think that someone who finds the evidence for theism slightly to moderately more plausible may not at that moment be

¹⁶⁵ Ibid., p. 266

able to believe atheism is true, by direct willing or otherwise.¹⁶⁶ But even if believing *against* the evidence is psychologically impossible, this says little about the possibility of willing to believe overall. It seems much more plausible to think that a deliberator in the circumstances Alston describes has the power to either embrace a belief in God or withhold belief on the matter, to remain agnostic, a possibility which Alston ignores. This possibility would constitute sufficient control for believing at will according to Alston's own standard, and would certainly fit within the framework of moderate direction control. I surely retain control over any action which I am at this moment either able to perform or refrain from performing, irrespective of any other options. Thus, while I have conceded the uncertainty of the possibility of believing at will for cases of true evidential equipoise, the remainder of our interaction with Alston's arguments have shown us no reason to doubt the probability of moderate direct control over belief.

5.5. Conclusion

I have argued in this and the preceding chapter that the contemporary arguments made against direct control over belief on a voluntarist conception fail, and specifically that they pose no threat to the thesis of moderate direct control articulated in Chapter 3. I have examined five representative examples from this literature, representing the best

¹⁶⁶ I should note here that while this claim may be reasonable, it is not obvious that this type of belief formation is impossible. As I have noted throughout, belief formation, retention and dissolution are psychologically complex phenomena, with various contributing factors. Since a response is not essential to my case, however, I leave this issue aside for discussion elsewhere.

case against doxastic voluntarism. Concerning the conceptual arguments against direct control, I have shown how William's central claims are mistaken, and then illustrated the explicit and implicit dependence relations between Williams and those, like Scott-Kakures, Adler and Audi, who have attempted to improve upon his arguments. With regards to the psychological arguments against believing at will, I have tried to show that they are often directed at an implausible conception of control, and that they fail to make any meaningful distinction between action and belief that supports direct control for the former, but not the latter. The upshot of this discussion is that the case for moderate direct control on a voluntarist conception remains intact. Individuals committed to a voluntarist picture of direct control over action should recognize themselves as also committed to a similar form of control over belief.

CONCLUSION

As I stated at the outset of my dissertation, my principal concern has been the topic of responsibility over belief. Our speech and experience are peppered with elements that imply or directly claim responsibility for our beliefs. My purpose has been to inquire as to whether there is any legitimate basis for these statements. As I discussed, the most direct route to establishing responsibility is via control, and so the topic of control has been my focus.

In the preceding chapters I have established the plausibility of moderate direct control. This thesis tells us that a significant percentage of our beliefs, or more specifically their formation, retention and elimination, are subject to our direct control. I made this argument first by appealing to what I called “common sense”, including our standard responsibility-attributing practices and experiences. I suggested that an intuitive and defensible interpretation of these factors provides a basic presumption in favor of direct control which the skeptic must overcome. I then argued that such a refutation is not available, first by offering theoretical frameworks in support of direct control, and then by examining and criticizing various responses offered in the relevant literature on doxastic control and responsibility.

In this latter task of providing theoretical support, I have tried to proceed in as neutral a manner as possible. It is impossible to operate without a theoretical framework of some kind, so I have instead operated by identifying the major distinctions in the

relevant literature (intellectualist, voluntarist, libertarian and compatibilist), and then in turn adopting as assumptions the various positions available within the debate over doxastic control. In each case, the broad structure of my argument has been to examine relevant criteria of control over action offered within the given theoretical space, and then to show its applicability to the formation, retention and dissolution of belief. I have thus not attempted to defend a particular position with regard to control, but rather have argued that those persons who embrace the reality of control and responsibility over action within one of these conceptions (which together comprise the majority of viewpoints on control) are thereby committed to a similar view regarding belief.

In arguing this way, I have departed from a significant portion of the literature on this topic, which focuses on justification, epistemic norms, and related epistemological concepts. It is certainly not my intent to suggest that such considerations are either irrelevant or uninteresting to considerations of doxastic responsibility. In adopting such an approach I was driven first by the conviction that the fundamental issue to be resolved when examining questions of responsibility is the issue of control, and second by the belief that the most perspicuous writing on control has come principally from the philosophy of action. As a conclusion to this effort I wish to briefly allude to some of the potential directions of further inquiry resulting from this work, as well as to mention some of the implications of my principal thesis.

While I have offered a somewhat detailed analysis of the nature and character of direct control over belief, I have not even begun to consider the role that indirect control

plays in making us responsible for our beliefs. Such an examination of indirect control offers one possibility for the extension of the effort that I have made in these chapters. Indirect control over belief, as I've noted previously, is much less controversial than direct control, but this is not to say that no arguments against it exist (Alston is again a main interlocutor here). A full account of the different mechanisms conditions of indirect control would provide valuable insight into the nature of both belief and control, and, in conjunction with the account provided in the preceding chapters, would form the basis of a comprehensive examination of doxastic responsibility and control.

More work also clearly remains to be done on issues of direct control. I briefly raised the issue of normative or pragmatic control in Chapter 2. This topic invites further exploration; the provocative work of Paul Helm on "belief policies" offers a starting point for the development of an account of this additional mechanism of control. I would even suggest that, if my project is successful, discussions of normative control and related issues should be the locus of further debate between advocates and critics of direct control. An exploration of normative control will allow for an inquiry into the role of practical or pragmatic factors in belief generation and modification, a topic of significant interest in literature on the ethics of belief. Normative control and the notion of belief policies also possess significant implications for religious epistemology and analyses of faith, religious, moral, and otherwise.

The epistemologically oriented literature on doxastic control that I've mentioned several times above offers a final avenue of extension for my work. If the general account

offered here is correct, this clearly has major implications for accounts of the epistemic norms governing belief, as well as discussions of the nature of justification. I can of course only gesture here, but it seems unlikely that any purely externalist account of justification is correct if what I have argued about doxastic control is true. This is not to say that the correct account of justification contains no externalist elements (as I've noted throughout, my account has focused on just a subset of our beliefs), but if our beliefs are subject to our rational control in any of the ways that I've described, then surely some elements of internalism must be included in any complete account of justification. This issue, as well as the ones I've mentioned just above, deserves the examination that I cannot offer here.

What precisely is the upshot of my argument? As I noted at the outset of my work, rather than requiring any change to our common practices of claiming and attributing both control over and responsibility for beliefs, my argument offers theoretical support to them. I take this to be a virtue of this account, as it shows that our commonsense self-conception regarding belief and control has significant contact with reality. This is of course not to say that all attributions of responsibility for belief are apt, such a claim is surely false. But evaluating such a claim would require much more analysis than I've offered thus far.

However, while I have provided nothing like a complete account and much more work remains to be done along the fronts I've mentioned, some conclusions can nevertheless be drawn. Though my arguments support our common practices regarding

belief, they do, if correct, undermine a significant body of philosophical literature. While I haven't exhaustively examined the logical space available with respect to control, I have surveyed the positions held by the vast majority of philosophers and shown that the commitments held concerning control over actions lead to similar commitments with respect to belief.

We thus have much more significant and direct control over our mental lives than most current philosophers have allowed. While, as I noted in Chapter 1, most of the critics of direct control have found ways to support our undeniably regular ascriptions of doxastic responsibility, the path to this conclusion has here been made much shorter and more direct. We intuitively take ourselves to be responsible for many of our beliefs. My arguments help to make the case that these intuitions need not be explained away. Nor does the straightforward normative character of these intuitions need to be attenuated to something less or different than what it seems to be, namely responsibility for something over which I exercise significant control. Our beliefs, so central to who we are, are properly seen as central to and products of, rather than simply constraints on, our agency.

BIBLIOGRAPHY

- Adams, Robert M. (1995) "Moral Faith," *Journal of Philosophy* 92, pp. 75-95
- Adler, Jonathan. (2002) *Belief's Own Ethics* (Cambridge: Bradford Book/MIT Press)
- Alston, William. (1988) "The Deontological Conception of Justification," *Philosophical Perspectives* 2, pp. 257-299
- Audi, Robert. (2001) "Doxastic Voluntarism and the Ethics of Belief," in Steup (2001), pp. 93-114
- Bach, Kent. (1984) "Default Reasoning: Jumping to Conclusions and Knowing When to Think Twice," *Pacific Philosophical Quarterly* 65, pp. 37-58.
- Bainton, Roland H. (1995) *Here I Stand: A Life of Martin Luther* (New York: Meridian)
- Bennett, Jonathan. (1990) "Why is Belief Involuntary?," *Analysis* 50, pp. 87-107
- Bennett, Jonathan. (1974) "The Conscience of Huckleberry Finn," *Philosophy* 49, pp. 123-134
- Bratman, Michael. (1987) *Intentions, Plans, and Practical Reason* (Cambridge: Harvard University Press)
- Bratman, Michael. (1985) "Davidson's Theory of Intention," in Lepore, Ernest. and McLaughlin, Brian. (eds.) *Actions and Events: Perspectives on the Philosophy of Donald Davidson* (Oxford: Basil Blackwell), pp. 14-28
- Broome, John. (1999) "Normative Requirements," *Ratio* 12, pp. 398-419
- Carroll, Lewis. (1897) *Through the Looking Glass: And What Alice Found There* (Philadelphia: Henry Altemus Company)
- Chisholm, Roderick M. (1968) "Lewis' Ethics of Belief," in Schilp, P.A. (ed.) *The Philosophy of C.I. Lewis* (LaSalle, IL: Open Court), pp. 223-242.
- Chisholm, Roderick M. (1957) *Perceiving: A Philosophical Study* (Ithica, NY: Cornell University Press)

- Clarke, Randolph. (2003) *Libertarian Accounts of Free Will* (New York: Oxford University Press)
- Cohen, Gerald A. (2000) "Paradoxes of Conviction," in *If You're an Egalitarian, How Come You're So Rich?* (Cambridge: Harvard University Press), pp. 7-19
- Cottingham, John. (2002) "Descartes and the Voluntariness of Belief," *The Monist* 85, pp. 343-360
- Davidson, Donald. (1973) "Freedom to Act," in Honderich, Ted (ed.) *Essays on Freedom of Action* (London: Routledge & Kegan Paul), pp. 137-156.
- Dennett, Daniel C. (1978) *Brainstorms: Philosophical Essays on Mind & Psychology* (Cambridge, MA: MIT Press)
- Feldman, Richard. (2002) "Epistemological Duties," in Moser, Paul K. (ed) *The Oxford Handbook of Epistemology* (Oxford: Oxford University Press), pp. 362-384
- Firth, Roderick. (1959) "Chisholm and the Ethics of Belief," *Philosophical Review* 68, pp. 493-506
- Fischer, John Martin. (2006) *My Way: Essays on Moral Responsibility* (Oxford: Oxford University Press)
- Fischer, John Martin. (1998) "Moral Responsibility and the Metaphysics of Free Will: A Reply to van Inwagen," *The Philosophical Quarterly* 48, pp. 215-220
- Fischer, John Martin and Ravizza, Mark. (1998) *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge: Cambridge University Press)
- Foot, Philippa. (2003) *Natural Goodness* (Oxford: Oxford University Press)
- Frankfurt, Harry. (1969) "Alternative Possibilities and Moral Responsibility," *Journal of Philosophy* 66, pp. 829-839.
- Gale, Richard. (2005) *The Philosophy of William James: An Introduction* (Cambridge: Cambridge University Press)
- Haack, Susan. (1998) "The Ethics of Belief Reconsidered," in Hahn, Lewis E. (ed.) *The Philosophy of Roderick M. Chisholm* (Chicago: Open Court), pp. 129-144
- Helm, Paul. (1994) *Belief Policies* (Cambridge: Cambridge University Press)

- Hieronymi, Pamela. (2006), "Controlling Attitudes," *Pacific Philosophical Quarterly* 87, pp. 45-74
- James, William. (1896) *The Will to Believe: And Other Essays in Popular Philosophy* (New York: Longman, Greens and Co.)
- Kane, Robert. (2005) *A Contemporary Introduction to Free Will*. (Oxford: Oxford University Press)
- Kane, Robert. (2000) "The Dual Regress of Free Will and the Role of Alternative Possibilities," *Philosophical Perspectives* 14, pp. 57-79
- Kelly, Thomas. (2003) "Epistemic Rationality as Instrumental Rationality," *Philosophy and Phenomenological Research* 66, pp. 612-640
- Kent, Bonnie. (1995) *Virtues of the Will: The Transformation of Ethics in the Late 13th Century* (Washington, DC: The Catholic University of America Press)
- Korsgaard, Christine. (1996) *The Sources of Normativity* (Cambridge: Cambridge University Press)
- Marty, Martin. (2004) *Martin Luther: A Life* (New York: Viking)
- McKenna, Michael. (2001) "Source Incompatibilism, Ultimacy, and the Transfer of Non-Responsibility," *American Philosophical Quarterly* 38, pp. 37-51
- Meiland, Jack W. (1980) "What Ought We to Believe? or The Ethics of Belief Revisited," *American Philosophical Quarterly* 17, pp. 15-24
- Nettleman, Nikolaj (2006) "The Analogy Argument for Doxastic Voluntarism", *Philosophical Studies* 131, pp. 559-582
- O'Connor, Timothy. (2000) *Persons and Causes: The Metaphysics of Free Will* (New York: Oxford University Press)
- Owens, David. (2000) *Reason Without Freedom: The Problem of Epistemic Normativity* (London: Routledge)
- Parfit, Derek. (forthcoming) *On What Matters* (Oxford: Oxford University Press)
- Pereboom, Dirk. (2003) "Source Incompatibilism and Alternative Possibilities," in Widerker, David and McKenna, Michael (eds.) *Moral Responsibility and*

Alternative Possibilities: Essays on the Importance of Alternative Possibilities (Burlington, VT: Ashgate Publishers pp. 185-99.

Pereboom, Dirk (2001) *Living Without Free Will* (Cambridge: Cambridge University Press)

Radcliffe, Dana. (1997) "Scott-Kakures on Believing at Will," *Philosophy and Phenomenological Research* 57, pp. 145-151

Raz, Joseph. (2008) "On the Guise of the Good," *Oxford Legal Studies Research Paper* 43 (Available at Social Sciences Research Network: <http://ssrn.com/abstract=1099838>)

Ryan, Sharon (2003) "Doxastic Compatibilism and The Ethics of Belief," *Philosophical Studies* 114, pp. 47-79

Schwitzgebel, Eric. (2002) "A Phenomenal, Dispositional Account of Belief," *Nous* 36, pp. 249-275.

Scott-Kakures, Dion. (1994) "On Belief and the Captivity of the Will," *Philosophy and Phenomenological Research* 54, pp. 77-103

Searle, John. (1983) *Intentionality* (Cambridge: Cambridge University Press)

Smith, Angela. (2005) "Responsibility for Attitudes: Activity and Passivity in Mental Life," *Ethics* 115, pp. 236-271

Steup, Matthias. (ed.) (2001) *Knowledge, Truth and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*. (Oxford: Oxford University Press)

Steup, Matthias. (2000) "Doxastic Voluntarism and Epistemic Deontology," *Acta Analytica* 15, pp. 25-56

Van Inwagen, Peter. (1989) "When is the Will Free?," *Philosophical Perspectives* 3, pp. 399-422.

Velleman, J. David. (2000) *The Possibility of Practical Reason* (Oxford: Clarendon Press)

Velleman, J. David. (1992a) "What Happens When When Someone Acts?," *Mind* 101, pp. 461-481

- Velleman, J. David. (1992b) "The Guise of the Good," *Nous* 26, pp. 3-26
- Watson, Gary. (2004) *Agency and Answerability: Selected Essays*. (Oxford: Oxford University Press), pp. 123-157
- Watson, Gary. (1978) "Free Agency," *The Journal of Philosophy* 72, pp. 205-220
- Widerker, David. (2002) "Farewell to the Direct Argument," *The Journal of Philosophy* 99, pp. 316-324
- Williams, Bernard. (1973) *Problems of the Self: Philosophical Papers 1956-1972* (Cambridge: Cambridge University Press)
- Winters, Barbara. (1979) "Believing at Will," *Journal of Philosophy*, v. 76, pp. 254-255.