2015

Moral Responsibility and Mental Health: An Interdisciplinary Project in Philosophy, Psychology, and Applied Behavior Analysis

Megan Marie McGrew
MORAL RESPONSIBILITY AND MENTAL HEALTH: AN INTERDISCIPLINARY PROJECT IN PHILOSOPHY, PSYCHOLOGY, AND APPLIED BEHAVIOR ANALYSIS

By

MEGAN MCGREW

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

Degree Awarded:
Spring Semester, 2015
Megan McGrew defended this dissertation on March 30, 2015.

The members of the supervisory committee were:

Randolph Clarke
Professor Directing Dissertation

Martin Kavka
University Representative

Michael Bishop
Committee Member

David McNaughton
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the dissertation has been approved in accordance with university requirements.
For Heidi. I wish I had known then, all that I know now.
ACKNOWLEDGMENTS

I cannot say enough good things about my dissertation director, Randy Clarke. If he ever
lost either hope or patience with me or this project, he was far too kind to ever show it. He has
also been incredibly generous with his time and attention offering excellent feedback from the
big ideas all the way down to tiny details.

I would also like to thank Kathryn Norlock and Michael McKenna who each did more
than their fair share in shaping my academic career. They each pushed me to be a better thinker,
but they also offered the support, encouragement, and affirmation that helped me to believe in
myself as a scholar.

Thanks to Michael Ruse, for being the first of my graduate professors to grade exams and
papers blindly. In my first year of graduate school, that did more than you can know to make me
feel welcome.

Special thanks to Rich Cordero, Sarah Whylly, Warren Harasz, Claire Millarker, and
Zachary Martin who all, in some way, helped me to sort through the earliest stages of this
project. Also, to my family – Randy McGrew, Robin Delarosa, Ryan McGrew, and Sarah
McGrew – for a lifetime of every kind of support imaginable.

Infinite thanks to Dave Donegan whose love, support, and tireless work ethic inspired me
to finish.

This project has been going on for so long and experienced so many permutations that it
is impossible to thank all of the people that have influenced my thinking, acted as a sounding
board, or supported and encouraged me along the way. I am grateful to each and every one
of them.
TABLE OF CONTENTS

List of Figures ................................................................................................................................ vi
Abstract .......................................................................................................................................... vii
1. INTRODUCTION .......................................................................................................................1
2. THE INTEGRATED MENTAL STATUS CONDITION .................................................................20
3. ACTING FOR REASONS ........................................................................................................38
4. UNDERSTANDING BASIC SOCIAL COMPETENCE ..............................................................55
5. REALITY AND PERCEPTION .................................................................................................73
6. FINDING RESPONSIBLE, HOLDING RESPONSIBLE, AND APPLYING IMS .......................86

REFERENCES ..................................................................................................................................98
BIOGRAPHICAL SKETCH .............................................................................................................103
LIST OF FIGURES

Figure 5.1 – Original Photo .................................................................74
Figure 5.2 – Boosted Original.................................................................74
Figure 5.3 – Focal Zoom Figure 5.2 .......................................................74
Figure 5.4 – Dual Tone Original...............................................................74
Figure 5.5 – Inverted Colors Original.......................................................74
Figure 5.6 – Posterized Original..............................................................74
Figure 5.7 – Heat-Mapped Original.........................................................75
Figure 5.8 – Neon Figure 5.7 .................................................................75
Figure 5.9 – Saturated Figure 5.8 ............................................................75
Figure 5.10 – Focal Black and White of Figure 5.9 .................................75
Figure 5.11 – Posterized Figure 5.10 .......................................................75
Figure 5.12 – Pencil Sketched Figure 5.11 .............................................75
ABSTRACT

In this dissertation, I draw from the fields of philosophy, psychology, and applied behavior analysis in order to present what I take to be a necessary condition for morally responsible agency, the Integrated Mental Status condition (IMS). I offer a view that privileges the mental status of the agent above other conditions when considering moral responsibility. I do this within a conceptual framework that takes moral responsibility to exist on a continuum and assessments of moral responsibility to be deeply contextual and localized to a time. Additionally, I make an argument for how and why finding and holding responsible can and should come apart in an analysis of moral responsibility.
CHAPTER ONE

INTRODUCTION

The Project

This project has changed dramatically since its inception. It was originally conceived of as an interdisciplinary conceptual analysis of a mental condition for morally responsible agency, drawing from the fields of psychology and analytic philosophy (particularly the moral responsibility literature), which would hopefully have some relevance in application. Over the past 2 years, however, I have had the opportunity to work very closely providing behavioral therapy to a number of children on the Autism Spectrum. This experience has dramatically influenced my approach to this project. I have broadened my theory to include insights from the field of Applied Behavior Analysis which, I believe, has sharpened my analysis in important ways.

An Interdisciplinary Approach

A discussion of moral responsibility, in application, need also be a discussion about human behavior. The benefit of bringing together these distinct fields is that the resources exist for asking very different questions and integrating the knowledge gained from such queries. Psychology can tell us a great deal about how people behave. Much of psychology deals with statistical analyses of how people behave and works to understand the underlying mental states that motivate behavior. Applied Behavior Analysis (ABA) is rooted in B.F. Skinner’s Radical Behaviorism. Though it grew out of psychology, it is a field entirely devoted to behavior; however, unlike psychological explanations, explanations of behavior in ABA are environmental and observable; understanding behavior means understanding observable functional relations. There is no field off-limits to philosophy; the methodological approach of analytic philosophy not only compliments both psychology and ABA, but allows for complex conceptual analyses that both serve as a check on the science itself and, at times, go beyond the scope of what the science is equipped to address.
Put Another Way

**Psychology** allows us to ask questions like: “How do people behave, on average (statistically)?” “Why do people, in general, behave in certain ways (psychological trends)?” “What is abnormal human behavior (statistically)?” “Why do individuals behave in particular ways (mentalistic explanation)?”

**Applied Behavior Analysis (ABA)** allows for different questions: “Why do people behave in particular ways (functional/environmental explanation)?” “What variables, in the environment, can be manipulated to alter human behavior (functional explanation)?” “Why is a particular individual behaving in a particular way (functional explanation)?” “How can we shape the behavior of a particular individual (functional/environmental explanation)?”

**Philosophy**, particularly Analytic Philosophy, allows us to ask virtually any question about human behavior (or just about anything else) remaining to us. Its methodological approach plays well with the data provided by research in the fields of Psychology and ABA while giving us the conceptual tools to assess the value and relevance of that data and to question methods that might, otherwise, be considered foundational (read: unassailable) within the field itself. Philosophy is also the only field that allows us to ask questions like “How should people behave, in general?” “How should particular individuals behave?” “What should a particular person do in a particular context?” “How should we think about human behavior? About morality? About moral responsibility?”

I have chosen to bring these three fields together for an analysis that aims both to clarify some conceptually murky waters and to have relevance in the actual world.

**My Thesis**

In this dissertation, I will defend the thesis that my Integrated Mental Status (IMS) condition is a necessary condition for moral responsibility on a *merit-based view*. It may also turn out that my Integrated Mental Status condition is sufficient for moral responsibility, but I

---

1 Many scholars have already seen the value of interdisciplinary study between psychology and philosophy; there are entire sub-disciplines devoted to it (Moral Psychology, Experimental Philosophy) and I believe incorporating lessons from a behavioral approach will allow for an even more robust analysis.

2 It is not readily apparent to me what else would be required for sufficiency. One thought is that an additional condition may be required in order to find agents not responsible in cases of coercion and manipulation cases and this may be the case. IMS may have the conceptual resources to deal with many of these cases (given the effects of
will not be arguing for sufficiency here. I will argue, further, that an interdisciplinary approach, when combined with a current time-slice view of responsible agency, an agential view of responsibility (as opposed to a mechanistic one), and a reinterpretation of old divisions regarding the concept of moral responsibility, can offer important insights into longstanding philosophical paradoxes, puzzles, and debates. These would include, but are not necessarily limited to questions about: instant agents, agents under the voluntary influence of a substance, the status of mentally ill agents, and how to account for psychopaths.

Project Overview

In what remains of chapter 1, I will explain why a comprehensive understanding of mental health requires going beyond the study of psychology and briefly address what I take to be the strengths and weaknesses of Psychology, Applied Behavior Analysis, and Analytic Philosophy. That discussion will, of necessity, be broad, sweeping, and over-generalized, but will explain how these approaches balance and complement each other in an analysis of moral responsibility while offering some basic background on each respective field. I will then address what I take to be at stake in a discussion of moral responsibility and introduce the condition that lies at the heart of my analysis: my proposed Integrated Mental Status condition (IMS) for moral responsibility. I will also highlight some important distinctions for the analysis to follow.

Chapter 2 will offer a breakdown of the Integrated Mental Status condition for morally responsible agency (IMS). Chapter 3 will address the strengths and weaknesses of what I take to be one of the best current theories of moral responsibility, the theory put forward by philosophers John Martin Fischer and Mark Ravizza in Responsibility and Control (1998). I will then explain how IMS builds on aspects of their reasons-responsiveness theory while avoiding some of the pitfalls. In Chapter 4, I will contrast the Social Competence aspect of IMS with Walter Glannon’s conception of normative competence as presented in The Mental Basis of Responsibility (2002). Chapter 5 will discuss relevant aspects of reality, perception, and subjectivity for an analysis of moral responsibility that, put one way, uses an objective assessment of subjective states to make objective claims about moral responsibility. In my final chapter, Chapter 6, I will bring together and summarize the key aspects of my view and discuss coercion/manipulation on the mental status of the agent) but I leave it open that it may not and consider it a question for further consideration at another time.
the relevance of IMS in assessments of moral responsibility, especially in the analysis of tough moral cases.

**Background**

**Experimental/Cognitive/Abnormal Psychology**

The American Psychological Association defines psychology as “the study of the mind and behavior,” and claims that the discipline “embraces all aspects of the human experience – from the functions of the brain to the actions of nations, from child development to care for the aged. In every conceivable setting from scientific research centers to mental healthcare services, ‘the understanding of human behavior’ is the enterprise of psychologists” (American Psychological Association, 2014). Psychology, thus, manages to cut out a large swath of territory, territory that it invariably shares with other disciplines and that has spawned numerous sub-disciplines within the field of psychology. For the purposes of this project, when referring to psychology, I am primarily referencing the areas of cognitive and abnormal psychology and I take psychology, on the whole, to be a discipline with related experimental and applied clinical treatment components.

The study of psychology (or even a series of psychological studies), while incredibly useful, can never be sufficient for understanding mental health. For one thing, that is not necessarily the objective; psychologists are often more occupied with the study of mental disorder or a particular aspect or human behavior than coming to an understanding of mental health. A suitably robust conceptual analysis of mental disorder could tell us a great deal about mental health, which, in psychology, is sometimes characterized merely as an absence of disorder and other times as a sort of mental flourishing (Oltmanns & Emery, 2012, 7). However, in psychological practice, mental disorder is generally taken to be a particular sort of deviation from a statistical norm or a person’s subjective experience of dysfunction (which may or may not be coupled with an objective assessment of a patient’s ability to function within society), and

---

3 ABA could be considered a sub-discipline given this definition, but I prefer to keep it distinct for the purposes of this analysis.

4 Which is to say, committed to the scientific method.

5 This, of course, being one area where psychology could benefit greatly from philosophical methods.
these interpretations don’t necessarily give us the sort of understanding of mental health needed for this kind of project.

While the field of psychology provides a wealth of empirical information, information critical to a truly informed, applied/applicable theory, I wish to discuss briefly some concerns about two common ways of considering mental health from a psychological perspective.

**A Statistical Model of Normalcy**

The diagnostic categories and criteria for mental disorder have changed (and continue to do so), often quite significantly, over time. The concern is that they change not merely on the basis of emerging science, but on the basis of prevailing cultural attitudes and norms. The Diagnostic and Statistical Manual of Mental Disorders (DSM), the standard criteria for the classification of mental disorders put out by the American Psychiatric Association, alters the scope of recognized mental disorders with each new edition.

Homosexuality was classified as a mental disorder by the American Psychiatric Association until 1974 and persisted into the 1980s via the DSM – III as the newly-minted, and widely criticized, *ego-dystonic homosexuality*. The most recent incarnation, the DSM – V, eliminates narcissistic personality disorder and adds a behavioral addictions category so that an addiction to gambling can be classified as a mental disorder. The DSM – V has also done away with Asperger Syndrome as a distinct diagnosis, reverting to an undifferentiated diagnoses of Autism Spectrum Disorder. While such revisions can be incredibly useful for treating certain individuals, they call into question what precisely is being tracked by the psychiatric category of mental dysfunction/disorder. The body of disorders (as put forward in the DSM) as well as the leading psychological theories attempting to define mental dysfunction⁶ are insufficient to the task of explaining what mental conditions are required for morally responsible agency.

The sort of statistical approach that psychology takes to mental health may be seen as analogous to approaching ethics from a standpoint of cultural relativism and subject to the same sorts of criticism. Simply put, a cultural relativist believes that what qualifies as morally right or morally wrong is a function of the expectations and norms of a given culture. The moral status of an action can only be judged from within the context of that culture. One cannot judge the actions of Nazis in Nazi Germany by any standards other than those of Nazi Germany, for

---

⁶ As with Wakefield’s evolutionary theory.
example, and would be similarly unable to judge a culture that prized feeding their third-born children to wolves (as a means of population control and as a sacrifice to their wolf-gods). Moral assessments based on cultural relativism can only ever speak to whether something would be right or wrong in a given culture; they can never say whether something is right or wrong *simpliciter* or whether one culture’s standard is better than another. Similarly, psychology can only speak to how people currently are and measure dysfunction with respect to current societal norms. The field of psychology on its own, and spanning domains of both diagnosis and treatment, is not well-equipped to pass reasoned judgment about how people *should* be or whether current societal norms are a worthwhile standard upon which to base analysis. This is not to say that individuals within the field cannot be appropriately critical, merely that psychology as a field lacks the conceptual resources to take on certain questions of value. Allen Frances, the psychiatrist who chaired the team responsible for the DSM-IV, after roughly a decade somewhat removed from the field, claims that even the conservative and well-researched, revisions in the DSM-IV probably did more harm than good and that the recently released DSM-V threatens to be more damaging and expand the reach of psychology/psychiatry well beyond its professional competence.

The writing is on the wall. “Normal,” badly needs saving; sick people desperately require treatment. But DSM-5 seemed to be moving in just the wrong direction, adding new diagnoses that would turn everyday anxiety, eccentricity, forgetting, and bad eating habits into mental disorders. Meanwhile the truly ill would be even more ignored as psychiatry expanded its boundaries to include many who are better considered normal. (Frances, 2013)

Even the most well-intentioned and broadly accepted attempts to classify mental status from within the field of psychology are problematic in application and can pose serious and very real consequences for those affected by a diagnosis.

**The Harmful Dysfunction View**

Many psychologists appeal to a harmful dysfunction view of mental disorder like that proposed by Jerome Wakefield. Wakefield claims that a condition is a mental disorder *iff* it meets two conditions: 1) it results from the inability of some internal mechanism to perform its natural function and 2) it causes some harm to the person, as determined by cultural standards.
Now, in order to assess the value of such a definition, we must first define some critical terms; much hinges on what is meant by inability, internal mechanism, natural function, and harm.

Oltmanns & Emery offer a basic explanation of harmful dysfunction in *Abnormal Psychology* (2010). The definition of inability is taken for granted; inability in common parlance may be used to connote a skill deficit, a performance error, or a lack of necessary motivation. As much psychological practice is devoted to the treatment of mental disorder, I think that the most charitable reading of inability in this case is one of “does not” rather than one of “cannot,” though perhaps it varies from case to case. The notion of internal mechanism (the malfunction of which satisfies the dysfunction portion of the harmful dysfunction condition) is somewhat vague, though the “…dysfunctions in mental disorders are assumed to be the product of disruptions of thought, feeling, communication, perception, and motivation” (Oltmanns & Emery, 2010, 7). As to natural function, Oltmanns and Emery explain that the “…natural function of cognitive and perceptual processes is to allow the person to perceive the world in ways that are shared with other people and to engage in rational thought and problem solving” (Oltmanns & Emery, 2010, 8). This view is subject to criticism but does offer some frame of reference from which to work. Accessing a collective vision of reality, rational thought, and problem solving are established as the natural functions of internal mechanisms.

Something like the harmful dysfunction view of mental disorder is at the heart of the definition put forth in the DSM – IV – TR which emphasizes the present distress (painful symptoms), disability (impairment in an important area of functioning), and a significantly increased risk of suffering, disability, pain, death or loss of freedom (Oltmanns & Emery, 2010, 8). The DSM excludes from consideration “an expectable and culturally sanctioned response to a particular event…deviant behavior…and conflicts that are between the individual and society” (Oltmanns & Emery, 2010, 8).

A harmful dysfunction approach to mental disorder holds a great deal of intuitive appeal, but remains problematic and falls prey to some of the same criticisms as a statistical approach. For example, a definition like that in the DSM is relativistic appealing to culturally relative definitions of harm and acceptability. I would argue that the best definition of mental disorder would have more universal application.

---

7 These are critical distinctions in Applied Behavior Analysis, though the concepts are not unique to ABA.
Mental Health as Flourishing

Thus far, I have discussed some common psychological approaches to mental disorder, but an analysis of mental health requires more than a definition of disorder. I would like to briefly address the psychological concept of mental flourishing.

They are typically people who experience many positive emotions, are interested in life, and tend to be calm and peaceful. Flourishing people also hold positive attitudes about themselves and other people. They find meaning and direction in their lives and develop trusting relationships with other people. Complete mental health implies the presence of these adaptive characteristics (Oltmanns & Emery, 2010, 9).

Mental flourishing, thus elucidated, has its merits. It is a plausible definition of happiness, well-being, or perhaps optimal mental health. I do have some concerns, however, about the net that this sort of definition casts. The first is that what are described as adaptive characteristics might be equally well described as side-effects of a privileged lifestyle endowed with plentiful financial resources, time, freedom from responsibility, freedom from worry, and personal freedom to pursue one’s own interests. Perhaps a privileged lifestyle does facilitate mental health. Certainly wealth, autonomy, and leisure can alleviate stresses that might negatively impact mental health. Even so, there are plenty of cases of highly privileged individuals that still seem to fall short of flourishing mentally. The second concern is that many people with severe cognitive limitations may exemplify the above definition. If a person is perpetually interested, in even the most mundane aspects of life, disposed to be happy, even in the face of extreme hardship, calm and peaceful, even in the face of injustice, self-satisfied, despite gross problems with her own character, or blindly trusting, of even the most obviously untrustworthy people, I would argue that she is neither flourishing mentally nor endowed with optimal mental health.

The criticisms leveled above should not be interpreted as dismissive of psychology at large. The field has much to offer and I will draw on that where it furthers my analysis. The point, rather, is that psychology has not yet arrived at an account of mental health that is sufficient for the analysis that I have in mind. Further, it is not clear to me that it ever could without availing itself of some of the conceptual resources utilized by other disciplines.

Applied Behavior Analysis

Though Applied Behavior Analysis is commonly considered a branch of psychology, I am treating it here as its own, distinct science. As it is a science that many people are unfamiliar
with, I will focus more on an overview in this section, and only briefly address criticisms. Applied Behavior Analysis (ABA) is rooted in B.F. Skinner’s Radical Behaviorism and “…is a scientific approach for discovering environmental variables that reliably influence socially significant behavior and for developing a technology of behavior change that takes practical advantage of those discoveries” (Cooper, Heron, & Heward, 2007, 3).

Though B.F. Skinner’s Radical Behaviorism establishes the philosophic roots for Applied Behavior Analysis, Applied Behavior Analysis is a discipline within the field of Behavior analysis which has two distinct branches: the experimental analysis of behavior and applied behavior analysis. The experimental analysis of behavior is practiced, primarily, in laboratory settings and yields most of the scientific literature underwriting the field. “This literature provides the scientific foundation for applied behavior analysis, which is both an applied science that develops methods of changing behavior and a profession that provides services to meet diverse behavioral needs” (The Behavior Analyst Certification Board, 2015).

The Behavior Analyst Certification Board (BACB), the chief credentialing organization for behavior analysts, points to a 1968 article by Baer, Wolf, and Risley, “Some Current Dimensions of Applied Behavior Analysis,” as a definitive expression of how applied behavior analysis should be understood.

Baer, Wolf, and Risley claim that there are seven dimensions of applied behavior analysis. Applied behavior analysis must be applied, behavioral, analytic, technological, conceptually systematic, effective, and have generality. I will briefly outline the dimensions of applied behavior analysis put forward in the article.

**Applied**

Applied behavior analysis is meant to be applied to real individuals, in the real world, and so “…applied research is constrained to examining behaviors which are socially important, rather than convenient for study” (Baer, Wolf, & R., 1968, 92). Here we see traces of Skinner’s worldview; many behavior analysts take it as their imperative to use behavioral interventions to improve the lives of individuals as well as society in general; “[t]he label applied is not determined by the research procedures used… [i]n behavioral application, the behavior, stimuli, and/or organism under study are chosen because of their importance to man and society” (Baer, Wolf, & R., 1968, 92).
Behavioral

Applied behavior analysis takes behavior – observable, measurable, behavior – as its subject. Applied behavior analysis eschews mentalistic explanations, not out of a denial that mental states and qualia exist, but because such things, to the extent that they cannot be measured\(^8\), are not useful in a functional analysis of behavior.

Applied research is eminently pragmatic; it asks how it is possible to get an individual to do something effectively … a subject's verbal description of his own non-verbal behavior usually would not be accepted as a measure of his actual behavior unless it were independently substantiated…there is little applied value in the demonstration that an impotent man can be made to say that he no longer is impotent. The relevant question is not what he can say, but what he can do. Application has not been achieved until this question has been answered satisfactorily. (Baer, Wolf, & R., 1968, 93).

Analytic

“The analysis of a behavior, as the term is used here, requires a believable demonstration of the events that can be responsible for the occurrence or non-occurrence of that behavior. An experimenter has achieved an analysis of a behavior when he can exercise control over it” (Baer, Wolf, & R., 1968, 93-94). While applied behavior analysis, constrained by real-world context and important ethical considerations, may not be able to demonstrate this control as decisively as it could be demonstrated in a laboratory setting, applied behavior analysts, nevertheless isolate and manipulate variables and take copious amounts of data on the effects. Data is graphed and analyzed and comparisons between baseline data and intervention data will yield varying degrees of certainty about whether control was established based on experimental design\(^9\) and other factors.

---

\(^8\) This leaves it open that with advances in technology that could make various brain states observable/measurable such things could be worked into an analysis. Skinner would argue against this. He did not deny brain states and acknowledged that they could potentially “throw light” (Skinner, 1953, 35) on functional relationships, but argued that they could never alter them. Many contemporary behavior analysts, however, hold a less rigid stance on this issue and practicing behavior analysts tend to work collaboratively with professionals from other fields (psychology, psychiatry, medicine, speech pathology, occupational therapy, etc.) to ensure the best treatment outcomes for their patients.

\(^9\) Behavior Analysts utilize a number of different experimental designs depending on the individual or group being studied and the type of behavior being analyzed. A design with reversals (return to pre-intervention conditions) can lend greater credibility to claims about control, in some cases, for example.
Technological

As with other sciences, applied behavior analysis requires that results be replicable, and to that end, it must be technological which “means simply that the techniques making up a particular behavioral application are completely identified and described...[t]he best rule of thumb for evaluating a procedure description as technological is probably to ask whether a typically trained reader could replicate that procedure well enough to produce the same results, given only a reading of the description” (Baer, Wolf, & R., 1968, p.95).

Conceptually Systematic

In addition to being technological, in the manner described above, applied behavior analysis, in its approach, needs to cohere with the principles of behavior and the best practice knowledge that can be gained through keeping up with the scientific literature in the field. Any old approach to changing behavior, cannot be considered behavior analysis; behavior analysis requires the maintenance of conceptual systems and “[t]his can have the effect of making a body of technology into a discipline rather than a collection of tricks” (Baer, Wolf, R., 1968, 96).

Effective

One of the hallmarks of applied behavior analysis is that it is effective in at least two different senses. One sense is that it effects change, change in behavior is tracked and graphed and if a change is not discernable, over time, interventions are adjusted until the targeted change is achieved. The other sense, is that the effects of applied behavior analysis are socially significant. A commonly accepted guideline amongst behavior analysts is that their behavioral objectives should be socially valid10 which is summed up thusly, “[in] evaluating whether a given application has produced enough of a behavioral change to deserve the label, a pertinent question can be, how much did that behavior need to be changed?” (Baer, Wolf, R., 1968, 96)

---

10 This is a term used in ABA that basically means that a behavior targeted for change should be socially significant; changing that behavior should improve the quality of life of the individual, the people who interact with that individual, and society at large. Habits that are merely annoying (tapping a pencil on a desk while thinking, for example) are not likely candidates for a behavior change program.
Displays Generality

Lastly, applied behavior analysis focuses on behavior change that will improve the life of an individual in multiple ways and over the long term; “[a] behavioral change may be said to have generality if it proves durable over time, if it appears in a wide variety of possible environments, or if it spreads to a wide variety of related environments” (Baer, Wolf, & R., 1968, 96). To this end, behavior analyst often target so-called behavioral cusps, which can be understood as behaviors that, once acquired, can bring the agent into contact with a range of new opportunities with far-reaching consequences (consider, for example, all that is possible once an individual learns to read). Generality is not achieved in every case, but it is something to work toward and should be considered when designing behavior programs; “[that] generality is not automatically accomplished whenever behavior is changed also needs occasional emphasis, especially in the evaluation of applied behavior analysis” (Baer, Wolf, and R., 1968, 96).

Did Skinner Have It All Figured Out?

The short answer is “No.” The entire field of Applied Behavior Analysis leans proudly on the Radical Behaviorism espoused by B.F. Skinner and most theoretical objections to Applied Behavior Analysis have to do with the fact that so many of ABAs underlying premises and guiding principles originated with Skinner’s Radical Behaviorism and that that particular brand of behaviorism holds that all of human behavior can be explained without reference to internal states.

I believe that some consideration does need to be given to postulated mental states in order to answer certain questions about human behavior. ABA frequently describes behavior as a product of a history of reinforcement, and there is more than a little truth to this if you are considering behavior, product, history, and reinforcement in the technical sense that behavior analysts do. What ABA cannot explain, for example, without appealing to internal states is why the same consequence can function as reinforcement for one person, punishment for another, and a neutral stimulus for a third. Even so, there are a number of important insights into human behavior that can be drawn from applied behavior analysis.

---

See George Graham’s 2010 SEP article, “Behaviorism” for an easy to follow discussion of the chief objections to Skinner’s view, namely that it prohibits hypotheses about representational storage and processing, failure to address/account for attendant phenomenal qualia associated with mental processing, and Noam Chomsky’s criticisms of Skinner’s analysis of verbal behavior.
Analytic Philosophy

Philosophy, like psychology, is a broad-sweeping discipline, but even more so. I will not consider all of philosophy in this project; that would go beyond the scope of any single conceivable work. What I am concerned with is philosophical methodology, borrowed from the analytic tradition, and a narrow subset of work in the realms of ethics (broadly construed), moral responsibility (more narrowly construed), and moral psychology.

It is worth spending a little bit of time here laying down some background on moral philosophy, especially where it is relevant to critical distinctions that I will be drawing in my analysis. In the Western tradition, Aristotle is generally taken to have been the first to formulate a proper theory of moral responsibility, in the 4th century BCE, laying the foundation for Virtue Ethics. In truth, he laid the groundwork for a lot more than Virtue Ethics. Central to Aristotle’s theory is the idea that “…one is an apt candidate for praise or blame if and only if the action and/or disposition is voluntary” (Eshleman, 2014, 4). Furthermore, a voluntary action only counts as voluntary if it possesses two distinct features: (1) it must have its origin in the agent (control condition) and (2) the agent must be aware of what she is doing/causing/being (epistemic condition) (Eshleman, 2014, 4-5).

There is a certain lack of clarity in Aristotle’s account. Aristotle’s overall proposal is very difficult to argue with, but the inherent ambiguity, especially with regard to the notion of appropriateness (Eshleman, 2014, 5) has led to a philosophical divide that persists to this day. In considering whether it is appropriate to praise or blame an agent, the notion of appropriateness can be thought of in (at least) two different ways: (1) such that praise/blame is appropriate in the sense that the agent somehow deserves praise/blame given the circumstances or (2) such that praise/blame is appropriate insofar as a particular type of reaction will result in some favorable consequence (e.g. change in behavior on the part of the agent); from (1) we get a merit-based view of moral responsibility and from (2) we get a consequentialist view (Eshleman, 2014, 5). A great deal of debate within the area of moral responsibility is still shaped by this critical divide, “[f]or example, those who accept the merit-based conception of moral responsibility have tended

---

12 There were, of course, Ancient Chinese philosophers of moral philosophy long before Aristotle.

13 Virtue Ethics is one of the “big three” ethical theories alongside consequentialism and deontology.

14 After reading his *Nichomachean Ethics* for the first time, as an undergraduate, I recall (sarcastically) summarizing his entire view as “Ethics is easy! Just do the right thing, at the right time, in right way, for the right reasons!”
to be incompatibilists” (Eshleman, 2014, 7). This distinction will play a critical role in my own analysis, when I parse out the benefits of keeping finding responsible and holding responsible conceptually distinct and offer an argument that finding a person responsible is not a necessary prerequisite to holding that person responsible.

While the resources, both with respect to conceptual analysis and with respect to the history of human ideas that fall under its purview, of analytic philosophy (as a field) are vast, it is not a discipline without shortcomings. In analytic philosophy there is a not-so-secret bias toward the elegant, deductive, clearly defined, and parsimonious, and there is perhaps too much emphasis on paring and shaping our theories to make them fit this ideal. Sometimes this means that we discount the messy cases altogether by narrowing the scope of our projects or concerning ourselves with mythical creatures like ideally rational agents. While I can appreciate the clarity that this can bring to an analysis, I have always found the messy cases far more instructive and an understanding of such cases to be far more vital if lessons from philosophy are to be applied in the real world.

The Benefit of Interdisciplinarity

My focus, in this project, is to offer a means of analyzing the sorts of cases that are often glossed over in the moral responsibility literature. I will draw primarily from the disciplines of philosophy, psychology, and applied behavior analysis. Their different approaches to the same or similar class of problems, I believe, can be used to great advantage in building a comprehensive and useful understanding of morally responsible agency.15

Philosophy is, of course, essential for this sort of project. It allows for great latitude in conceptual analysis, and that is certainly what is called for here. Where psychology and applied behavior analysis are sciences (social and natural respectively) with established standards, checks, and procedures, they take for granted that science is a self-correcting enterprise and cannot offer a self-reflective critique of their own methods. Further, while we can learn a great deal about how people actually are, psychology and applied behavior analysis can never directly take on the question of how people ought to be. Nevertheless, psychology and applied behavior analysis remain vital as a conceptual analysis of how people ought to be has very little relevance

15 Moreover, there is good precedent for doing so. The contributions of Experimental Philosophy as a sub-discipline demonstrate the value of interdisciplinary work between philosophy and psychology.
if we fail to take into account how people actually are. The strength and weakness of psychology is its reliance on statistics. Statistics can convey a great deal of important information, but defining “normal” as a statistical range can be deeply problematic. The strength and weakness of applied behavior analysis is that it dismisses mentalistic interpretations of human action altogether and focuses strictly on the functions of human behavior. This sort of approach can serve as a helpful check on a philosophical or psychological analysis so entrenched in the postulated inner workings of the mind (the interplay of first-order and second-order, conscious and unconscious desires, etc.) that the practical implications and applications become lost. It is important, however, that we don’t allow our analysis to become simplistic, in the process. It is my hope that these approaches will temper each other and offer a view of moral responsibility that considers the mental status of the agent in context and is informative, useful, and broadly applicable.

The conceptual analysis of philosophers in this area is often pre-occupied with questions about control and free will, most frequently with the mental stability, or more often rationality, of the agent taken for granted. Psychologists do not take mental stability for granted, but they are generally concerned with studying deviations from the statistical norm and helping the abnormals to act more normally when their deviation from the norm results in a subjective experience of dysfunction. Even if we were to fully flesh-out the statistical norm, it would be something variable and subject to cultural trends, and not something we could take, unreflectively, to be a necessary condition of morally responsible agency. I have drawn from what I see as the most relevant aspects of the philosophical, psychological, and applied behavioral literature in order to provide some substance and clarity to a condition that, while common-sense, is also quite slippery. In what follows, I offer a basic sketch of the integrated mental status condition that I have in mind. The condition will be discussed in further detail in later chapters.

**Commitments Underlying My View**

I will say at the outset that I do not think of moral responsibility as an all-or-nothing affair. The same agent may be morally responsible in one case and not in another, or more morally responsible in one case than in another. I will argue that this can be the case purely as a
function of the, often variable, mental status of agents. Agents can be more or less responsible and their level of responsibility can, and generally does, vary over time and between cases.

My objective is to propose a necessary condition for minimal moral responsibility. Satisfying all parts of the condition will be necessary for morally responsible agency, but agents may still be more or less morally responsible, once this threshold is reached, depending on how well they satisfy each part of the condition. Put differently, the minimal satisfaction of my proposed condition will show that the agent may be to some degree morally responsible (indexed to a particular case at a particular time) and the more robust the satisfaction of the condition, the greater the degree to which the agent may be responsible.

I take the presence of a mental status condition for morally responsible agency to be a fairly common-sense idea (and it is not without precedent in the literature). Certainly our day-to-day judgments regarding moral responsibility take mental status into account. We often excuse or, hold less responsible, people who are highly emotional, or mentally handicapped, or even under the influence, and there are provisions in the law that take such things into account as well. While I do think that a lot of our commonly held and deeply felt intuitions about moral responsibility rely on the assumption of a mental status condition, nowhere are the details of such a condition satisfactorily spelled out. Further, it seems that leaving such important criteria undefined is likely to lead to inconsistent, hasty, and rationalized ascriptions of moral responsibility.

I will attempt to lay the groundwork with a necessary multi-part condition for morally responsible agency that, when applied within a certain rational framework, offers a means of analyzing and assessing ascriptions of moral responsibility. I call this condition the Integrated Mental Status condition (IMS), and it has three major components that will be presented in-depth in Chapter 2.

---

16 I will, especially in Chapter 6, address drugs and alcohol specifically. At a certain level of intoxication, an agent won’t satisfy the condition; I will explain how and to what degree moral responsibility can be attributed in such cases.

17 I am not committed to saying that this condition is sufficient for moral responsibility, though it may be.
Important Distinctions

Before concluding this chapter and moving into the meat of my analysis, I want to lay out some distinctions that should be kept in mind moving forward.

Finding Responsible Versus Holding Responsible

That finding someone responsible is related but conceptually distinct from holding someone responsible is not a particularly novel idea. After all, in courts of law, we determine whether someone is guilty or not guilty (finding responsible\textsuperscript{18}) prior to sentencing, where it is determined what the consequences for that individual will be (holding responsible). It is important to understand that the Integrated Mental Status condition that I will propose applies only to finding responsible and not to holding responsible. While we generally take holding responsible to follow quite naturally from finding responsible, in the course of my analysis, I will argue something more controversial: that sometimes it is appropriate to hold responsible even without finding responsible. To be clear, I am not suggesting that we start charging people with crimes that they did not commit. My claim is that even when an agent fails to satisfy the Integrated Mental Status condition (and thus, cannot be found morally responsible) it may be appropriate to hold an agent responsible (which is to say, impose consequences as if she had been found responsible).

Hanna Pickard, an analytic philosopher, with experience in a clinical mental health setting, has reached a similar conclusion about finding and holding responsible for which she uses a slightly different vocabulary; my conception of holding responsible is similar to her conception of “detached blame” (Pickard, 2011). This will be more fully explained later on, but I will say that my reasons for arriving at this distinction are drawn largely from Applied Behavior Analysis and that it is a claim that, though it has implications for the case of the psychopath, is meant to apply quite broadly.

Responsibility as Spectral

It is important to think of moral responsibility as a spectral concept, not in the sense of ghostly or ephemeral, but in the sense that it is best understood as existing on a spectrum. With IMS, I am trying to identify the part of the spectrum that serves as the threshold before which we

\textsuperscript{18} Legally responsible, in this case, rather than morally responsible.
must find an agent not responsible and beyond which we must find an agent responsible to some degree. Importantly, the further beyond this threshold the agent is, the more morally responsible she should be found.

**Indexing**

I don’t think that I need to say much here, but I do want to make it clear at the outset that, on my view, assessments of moral responsibility must be indexed to a particular time. Whether or not Tina, for example, is a morally responsible agent (full stop) is a nonsensical question. It only makes sense to ask whether Tina is morally responsible at a particular time in a particular context. I have already introduced responsibility as a spectral concept, here I want to make it clear that the responsibility of the same agent can be located at different points along the spectrum, at different times and under different circumstances. I mean this in a more thorough-going way than, say, a commitment to the age of reason (an age at which children become responsible), or an allowance that senility or mental illness can, and do affect moral responsibility. On my view, the same agent can satisfy IMS at some time, fail to satisfy it at another, and satisfy it again later, regardless of the temporal proximity of the distinct times in question.

**Agential Versus Mechanistic Models**

My Integrated Mental Status condition is a condition that an agent must satisfy if that agent is to be found morally responsible. Many theorists have speculated about conditions for moral responsibility. These conditions tend to focus on certain types of control, certain capacities for reason, history of an agent, and autonomy. These are highly relevant concerns for assessing moral responsibility, but in-depth analysis of these concepts can lead to somewhat suspect ontological assumptions.

It is not uncommon, in the moral responsibility literature (see Fischer & Ravizza, 1998; Haji, 1998) to reference the mechanism that issues in some particular action or, especially in discussions of autonomy and control, an agent’s own mechanism. Fischer and Ravizza depend on the idea of an agent’s own reasons-responsive mechanism to do a great deal of work for their theory. Unfortunately, it is less than clear what, precisely, constitutes a mechanism. It is even less clear what conditions would make a mechanism an agent’s own mechanism, in the relevant
sense, what other sort of mechanisms (if not an agent’s own) we should consider, and what power imbues these *mechanisms* with the power to do so much theoretical heavy lifting.

I’ve yet to see a mechanistic account that satisfactorily answers these questions. As it stands, I see no reason to further obfuscate complex concepts with an additional layer of abstraction. This is why I prefer and have adopted a model of moral responsibility that speaks to the mental status of the agent herself (and does not posit anything beyond the agent) in the analysis to follow.
CHAPTER TWO
THE INTEGRATED MENTAL STATUS CONDITION

In this chapter, I will offer a brief sketch of the Integrated Mental Status Condition that I have in mind. Much the rest of the chapter will be devoted to a more in-depth explanation and assessment of each of its parts.

There are three major components to the Integrated Mental Status condition, hereafter IMS. The agent must have: 1) the ability to act for reasons 2) basic social competence and 3) access to reality. Each of these components has distinct sub-parts that are discussed below.

The Components

The Ability to Act for Reasons

There is a critical connection between reasons-responsiveness and moral responsibility. In outlining this component of IMS, I borrow heavily from ideas espoused in the reasons-responsiveness literature and especially the theory of reasons-responsiveness put forward by John Martin Fischer and Mark Ravizza in *Responsibility and Control*, which I have found very helpful in considering the relevant features of moral responsibility. That said, Fischer and Ravizza’s view has certain problematic consequences that I aim to avoid here. The ability to act for reasons, as I take it, requires 3 things: recognition, processing (of a particular sort), and motivation.

Recognizing reasons as reasons is a relatively straightforward idea, and my understanding of the degree to which this is necessary for morally responsible agency is pretty well in line with Fischer and Ravizza’s regular receptivity to reasons. Processing reasons involves weighing and ordering reasons in a sensible way. I realize that I am expecting sensible to do a great deal of work here and there isn’t really any way to explain that word without appealing to some normative standard. What it should mean in this case is that the ordering of

---

19 I say recognize reasons as reasons though some take recognizing reasons as reasons to be implicit in recognizing reasons. Given that it is possible to have an agent recognize various things about her environment, things that would generally be thought of as reasons to act, but fails to understand the relevance of those things, it is important to be clear that this is not the sort of recognition I am requiring. If Carol notices a small child wandering toward traffic and recognizes that the child could be injured or killed and, yet, she can only see these things as facts about the world, then she is missing something crucial.
reasons should be rational enough that it can be understood by a rational third party; the order need not be rigorously rational but neither should it be so idiosyncratic as to be uninterpretable.\(^\text{20}\)

The motivational element required has two parts. The first is very similar to Fischer and Ravizza’s weak reasons-reactivity, but does not take it for granted that “reactivity is all of a piece” (Fischer and Ravizza, 1998, 73). The agent should be able to summon sufficient motivation for action on the basis of some sufficiently strong reason; the agent may lack a sufficiently strong reason and so fail to act, but would satisfy the condition if they could have summoned the motivation to act in the face of a sufficiently strong reason.

Fischer and Ravizza merely require that there be some reason, in some possible world, that would have motivated the agent to act, had the agent been presented with that reason. The second part of my motivational element aims to avoid the bizarre consequence of Fischer and Ravizza’s view that agents who minimally meet the weak reasons-reactivity condition appear to be of questionably sound mind. If a highly idiosyncratic reason is the only reason that could have motivated the agent to act other than she did, it calls into question whether it is appropriate to say that she is reactive to reasons (even weakly). If Wendy plans to break up with Randall, and the only scenario in which she decides not to break up with Randall is the one in which he presents her with a jar of her favorite green candies, and, not in any scenarios where a similar quantity of green candies is presented in a box, or coupled with concert tickets to see her favorite band, it is unclear what precisely, Wendy is reacting to. Reactivity needs to require some weighing and ordering of reasons and should a reason exist of sufficient weight to motivate action, the agent should also be motivated to act for other reasons, should the agent judge them to be of similar or greater weight.

**Basic Social Competence**

The basic social competence that I have in mind does not require anything like mastery of social graces, but it is of crucial importance as I take the human capacity for morally responsible agency to stem from our status as social animals. A basic social competence is needed for an agent to recognize and attend to features of the world that may be morally salient.

\(^{20}\) Where interpretable means something like understandable rather than conceivable. Being able to conceive of someone weighing the fact that Mary’s nail polish is pink as the most important reason to help Mary out of her wrecked car does not make it understandable in an important sense.
This competence includes a basic understanding of beings deserving of moral consideration – that they are sentient, feel pain, can die, etc. The agent also needs a basic understanding of social expectations and norms, and at least some limited capacity to acquire this understanding when faced with new circumstances. This, of course, will require a moderate understanding of human communication and the capacity to exert some level of control over emotional impulses.

Access to Reality

This component is at once the most obvious and the most difficult to outline. I consider it well beyond the scope of this project to settle disputes over the ultimate nature of reality. This aspect of the component merely requires that the agent’s perceptual and interpretative apparatuses be relatively unimpaired. There is no getting around the fact that subjects superimpose their own interpretive filters on reality, but the primary objective here is to rule out active psychoses, hallucinations or deep-seated and unwarranted paranoia, for example, that prevent the agent from experiencing the world largely as it is. Below, I have included a table to serve as a basic overview of the condition.

Table 2.1: Integrated Mental Status Condition

<table>
<thead>
<tr>
<th>Primary Components</th>
<th>Sub-parts</th>
</tr>
</thead>
</table>
| Ability to Act for Reasons| • reasons-recognition  
                          • sensible weighing and ordering of reasons  
                          • motivation to act for some reason  
                          • motivation to act for (some) reasons of greater or similar weight |
| Basic Social Competence   | • a basic understanding of beings worthy of moral consideration (vulnerabilities, etc.)  
                          • basic understanding of social parameters, expectations, and norms  
                          • capacity to acquire a basic understanding of social expectations (in new situations)  
                          • capacity to exert some control over emotional impulses |
| Access to Reality         | • free of active psychoses (hallucinations, paranoia, etc.) that prevent the agent from experiencing the world (largely) as it is  
                          • perceptual and interpretive apparatuses are functional/uncompromised |

21 This is not to say that the agent will necessarily believe that such creatures are worthy of moral consideration. I am stipulating that such creatures are deserving of moral consideration and merely require that the agent understand pretty basic things about such creatures and have the general ability to distinguish between sentient beings and other things.
In what follows, I will address each component of the condition in a bit more depth, keeping in mind that each of the 3 major components will be taken on individually in Chapters 3, 4, and 5, respectively.

**The Ability to Act for Reasons**

As the first part of IMS relies heavily on the work of John Martin Fischer and Mark Ravizza in *Responsibility and Control*, I will briefly summarize their view of moral responsibility as Moderate Reasons-Responsiveness. The discussion here will be intentionally brief, as much of chapter 3 will be devoted to a more thorough analysis of their view.

**Fischer and Ravizza’s Moderate Reasons-Responsiveness**

Fischer and Ravizza offer a semi-compatibilist view of moral responsibility, arguing that what is required for an agent to be morally responsible is a certain type of control, namely, guidance control. Guidance control of an action involves an agent’s freely performing that action, via the agent’s own reasons-responsive mechanism, but is not the sort of control that requires alternative possibilities (i.e. that the agent could have done otherwise).

Fischer and Ravizza believe that moral responsibility requires that an agent be acting from a mechanism that is responsive to reasons. In considering how responsive the mechanism needs to be in order to qualify, Fischer and Ravizza consider a minimal responsiveness to reasons (Weak Reasons-Responsiveness) as well as a maximal responsiveness (Strong Reasons-Responsiveness) and dismiss them as too weak and too strong respectively. Weak Reasons-Responsiveness would merely require that there is some counterfactual scenario in which the agent, with the same mechanism operating, when presented with a reason for acting counter to the way the agent acts in the actual sequence, recognizes that reason as sufficient for acting differently, decides to act differently because of that reason, and so acts. Strong Reasons-Responsiveness would require that in any counterfactual scenario wherein the agent, with the same mechanism operating, is presented with a sufficient reason to act differently, she recognizes that reason as sufficient, decides to act differently because of that reason, and so acts. Weak Reasons-Responsiveness could be satisfied by insane persons who are not appropriate candidates for ascriptions of moral responsibility, and Strong Reasons-Responsiveness sets the bar far too high, and would likely exclude all human beings from candidacy as morally responsible agents.
Fischer and Ravizza develop their account of Moderate Reasons-Responsiveness by considering the components of reasons-responsiveness individually. Reasons-responsiveness requires both the ability to recognize reasons (receptivity to reasons) and the ability to act based on those reasons (reactivity to reasons). Fischer and Ravizza claim that moral responsibility requires a mechanism that is, in addition to being the agent’s own, regularly receptive to reasons, at least some of which are moral reasons. A mechanism that is regularly receptive to reasons can recognize and understand the weight of reasons in a way that demonstrates a rational pattern that minimally conforms to reality. In contrast, the mechanism need only be weakly reactive to reasons, and need not be reactive at all to moral reasons; there need only be one counterfactual scenario in which the agent, with the same mechanism operating, is moved to act by a reason to act otherwise that she deems sufficient. According to Fischer and Ravizza, this single instance of reactivity is enough to confirm that the mechanism has the “executive power” to react to a reason and so the general capacity to react to reasons that is required for moral responsibility (Fischer & Ravizza, 1998, 73).

The reasoning behind this asymmetry is that regular receptivity to reasons is necessary to ensure that the mechanism is appropriately rational and in tune with reality; without a general understanding of reasons or the ability to recognize a pattern of similar reasons, it would not be appropriate to hold agents responsible for a failure to recognize and act on sufficient reasons when they are present. Fischer and Ravizza don’t want to claim that an agent’s mechanism is required to react to all of the sufficient reasons that there are (Strong Reasons-Responsiveness) as they believe this would wrongfully exclude an agent whose mechanism is not reactive to some sufficient reason (on a whim, say) from being morally responsible. This is the reasoning that motivates Fischer and Ravizza’s asymmetry thesis, that an agent need be regularly receptive to reasons but only weakly reactive to them. Importantly, Fischer and Ravizza set no requirement that an agent’s mechanism ever be reactive to any moral reasons, even though that same mechanism must be regularly receptive to at least some moral reasons. It should become clear in what follows that the weak reactivity put forward by Fischer and Ravizza is too weak to serve as a basis for determining moral responsibility or to do the explanatory work that Fischer and Ravizza want it to do.

---

22 They are explicit on this point. See Responsibility and Control page 79.
A Crazy Problem

Alfred Mele, Michael McKenna, and others have offered critiques of Fischer and Ravizza that point to a serious flaw in the criteria for Moderate Reasons Responsiveness. We can think of cases where agents satisfy Fischer and Ravizza’s criteria for Moderate Reasons Responsiveness but do not seem to be morally responsible.

Mele famously considers the case of an agoraphobic (Mele, 1989; Mele, 1996). Imagine an agoraphobic who cannot bring herself to leave the house despite having very good reason to do so, say, attending her daughter’s wedding at the chapel next door. The agoraphobic is operating from a mechanism that recognizes a wide variety of sufficient reasons for leaving her apartment and so is regularly receptive to reasons. It is also the case that, should her apartment go up in flames, that same mechanism would cause the agoraphobic to flee for her life (perhaps to the chapel next door); the agoraphobic, thus, is at least weakly reactive to reasons as well. Fischer and Ravizza’s criteria for Moderate Reasons Responsiveness are satisfied by the agoraphobic in this case. However, this puts us in the position of having to say that the agoraphobic was morally responsible for staying home on her daughter’s wedding day and has the further implication that all manner of phobics, addicts, and compulsives are morally responsible for actions stemming from those phobias, addictions, and compulsions. This result seems deeply problematic.

The sort of criticism leveled by McKenna, that I will take up in chapter 3, seems to show that agents whose mechanisms only minimally satisfy Fischer and Ravizza’s conditions, may be insane and, in some cases, seem insane by virtue of that minimal satisfaction and not in spite of it. An agent whose mechanism is regularly receptive to reasons but only very weakly reactive, and reactive only on the basis of an extremely bizarre or trivial reason may not merit an ascription of moral responsibility.

Acting for Reasons and ABA

Before moving on to the next part of my Integrated Mental Status condition, I wish to briefly outline some core aspects of Applied Behavior Analysis that pertain to an ability to act for reasons and will become increasingly important to our analysis (especially the Chapter 4 discussion of psychopathology). Where psychology and philosophy can often take a mentalistic approach when it comes to explaining human behavior, applied behavior analysis looks to
functional relationships between behavior and the environment in order to explain behavior (and, therefore, actions). Behavior analysts do not deny mental representations, but they do not appeal to them, directly\(^23\), in their analysis of behavior. This analysis can become quite complex, accounting even for verbal behavior. What I will detail here are the so-called ABCs – antecedents, behaviors, and consequences - and the standard 4-part contingency that underwrites much of the functional analysis in ABA.

There is a contingent relationship between the components: antecedents influence the behaviors that follow them, the behavior affects the consequences that follow, and the consequences affect the likelihood of that same/similar behavior to occur in the future\(^24\). Antecedents consist of 2 distinct components: MOs (motivating operations) and SDs (discriminative stimuli). The behavior (generally referred to as a response) follows the antecedent and the consequence is some stimulus/stimuli that follows the behavior. Motivating Operations temporarily affect the value of a particular consequence. There are two types of motivating operations: establishing operations (EOs) and abolishing operations (AOs). Establishing operations will temporarily increase the value of a consequence, while abolishing operations will decrease the value of a consequence. A discriminative stimulus (SD) is an antecedent stimulus that signals the availability (or unavailability) of a consequence. Importantly, a stimulus becomes a discriminative stimulus as the result of learning,\(^25\) which is to say, after the agent has experienced a contingency where that antecedent stimulus became linked to a particular consequence; it is called a discriminative stimulus, because an agent must discriminate between all of the available stimuli in order to identify which stimulus signals the availability of a particular consequence.

---
\(^23\) It is often possible to address mental representations indirectly utilizing a behavioral approach and behavioral terms.

\(^24\) Reinforcement increases the rate of some behavior over time, while punishment decreases the rate of some behavior over time. What functions as a reinforcer/punisher can only be determined by observing the rate of that behavior, over time, after that consequence has been delivered. Reinforcement and punishment can be either positive or negative, but this does not refer to their function, positive reinforcement and positive punishment add something to the environment, while negative reinforcement and negative punishment take something away from the environment. Consequences are also important for turning neutral antecedent stimuli into the discriminative stimuli of the future.

\(^25\) I am setting aside rule-governed behavior here, for the sake of simplicity, and discussing cases that involve direct contact with contingencies. It is, however, worth noting that Applied Behavior Analysis has the conceptual resources to accommodate and explain rule-governed behavior.
The standard 4-part contingency looks like this:

\[(\text{MO} + \text{SD}) \rightarrow \text{R} \rightarrow \text{C}\]

An example should help to clarify somewhat. Sarah is working out hard at the gym. She realizes that she is quite thirsty. She sees the water fountain across the room, goes to get a drink (or several) from the fountain, and finds that she feels better. In this simple example, Sarah’s thirst functions as the motivating operation, more specifically, her thirst is the establishing operation temporarily increasing the value of something to drink. The water fountain functions as the SD; it is the stimulus in the environment signaling the availability of water (Sarah has obtained water from water fountains in the past). Sarah, motivated by the combination of her thirst and the availability of water, drinks from the fountain (this is the response). Sarah obtains water from the fountain and feels better; she is no longer thirsty and feels somewhat cooler and more comfortable; this is the consequence. It should be pretty easy to see how Sarah’s thirst and the presence of the water fountain motivated Sarah to drink from the water-fountain. The consequent quenching of thirst should function as reinforcement, and increase the likelihood that Sarah will seek water from fountains, when thirsty, in the future. Now, if, instead of cool water, warm, rusty, sludge had emerged from the water fountain, Sarah would likely discontinue her use of that particular water fountain (if not water fountains in general) in the near future.

What I most want to point out, at this stage, is that the functional components used to analyze an instance of behavior in ABA track the concepts that I will be discussing in important ways. I have said, already, that I take the ability to act for reasons to require 3 things: recognition, processing (of a particular sort), and motivation. These requirements are embedded in the 4-part contingency as well. It’s easy to see how the MO (motivating operation) relates to motivation, so I will focus for just a moment on the SD. The discrimination necessary for a stimulus to function as an SD requires recognition as well as processing. The SD must be recognized, and implicit in the recognition of an SD, as such, is an understanding of the likely consequences of a given response in a particular context. This will be discussed more in later chapters as there are implications, not only for the ability to act for reasons, but for the basic social competence and access to reality conditions of IMS as well.
Basic Social Competence

The basic social competence component of IMS has several different parts. In order to satisfy these parts, the agent must be able to recognize and have some understanding of appropriate targets for moral consideration, have the skills necessary to comprehend and navigate social contexts, and have the ability to control socially inappropriate impulses.

Value

It is perhaps controversial, but I consider sentient life to have intrinsic value\(^{26}\). I also take morality to be largely concerned with how we interact with the intrinsically valuable. Being morally responsible, on my model, requires that the agent have some sense of this sort of value and an understanding of the characteristics and vulnerabilities of the intrinsically valuable.

IMS requires that the agent understands the difference between a toddler and a rock and why kicking the toddler is wrong when, \textit{ceterus paribus}, kicking the rock is not. Some examples should help clarify what this point is trying to capture.

It is very common for very young children to be incredibly egocentric, and this egocentrism stems, not from a character flaw, but from an inability to appropriately distinguish between the self and the rest of the world. They do not understand that other people are unique and complex, and they cannot adequately consider another’s point of view. It is easy to see how this sort of egocentrism can act as a barrier to moral responsibility; indeed, a large part of the moral education of young children is helping them to understand other points of view.

We do not find\(^{27}\) very young children morally responsible when their capacity for understanding appropriate targets for moral consideration is undeveloped. Similarly, we shouldn’t find responsible those older agents who are similarly limited, whether it is due to temporary or permanent diminished capacity or a lack of the capacity altogether. We don’t, for example, want to find Lennie (from Steinbeck’s \textit{Of Mice and Men}) morally responsible for the death of his puppy. Though Lennie killed the puppy, he lacked an appropriate understanding of

\(^{26}\) This is consistent with a hierarchy of value. To say that all sentient life is intrinsically valuable is not to say that it is all equally intrinsically valuable. Neither does this rule out the possibility of other things having intrinsic value including rights or obligations that may not reduce to goods.

\(^{27}\) Keep in mind the distinction that I have drawn between finding and holding responsible. To say that we should not find responsible in these cases, is not to say that we should not hold responsible (i.e. implement consequences \textit{as if} responsible) in any of these cases.
the vulnerability of the puppy relative to his own strength and was motivated by the simple
desire to pet soft things.

To be clear, this is not to say that we should never hold such agents responsible; indeed,
doing so can serve as a crucial component of moral education and foster the development of
capacities that would allow an agent to be morally responsible at some point in the future.
Central to the practice of Applied Behavior Analysis is the understanding that consequences are
the most important variable in shaping long-term behavior change. This is why the distinction
between finding responsible and holding responsible is so important and why neither a merit-
based nor a consequentialist view of an applied theory for moral responsibility will suffice. A
child may not be morally responsible, but a child who does not contact consequences for
problematic behavior may never become morally responsible. I will speak more to this point in
later chapters and at length in Chapter 6.

Navigating social context

As a rule, human beings are taken to be social creatures, and so our ability to interact
socially – to read attitudes and emotions in others, to anticipate likely reactions to a statement or
action, to follow certain rules of etiquette, to express contextually appropriate emotions, to
understand a joke – is generally taken for granted. While social interaction comes easily to
some, perhaps even (statistically) most, people, in others it can require painstaking deliberation
and attention to detail or fall outside of one’s capacities altogether. Those people who are
unable, for whatever reason, to competently navigate their social context will be less morally
responsible in those contexts, as a presumption of such competence underwrites many of our
moral attitudes and expectations.

Jack

Jack has Tourette Syndrome, and in addition to simple and complex motor tics, Jack also
experiences complex vocal tics, uttering words or phrases that he cannot always suppress. Like
many with Tourette Syndrome, Jack’s tics become noticeably worse when suffering greater
anxiety or stress than usual. We should not, and hopefully do not, find Jack morally responsible
for his tics even when those tics amount to verbal utterances that would otherwise be considered
inappropriate and potentially hurtful. Thus, when Jack utters “fat bitch,” to a close friend that he
knows to be deeply sensitive about her weight, he is not responsible when the utterance is a tic.
Additionally, and importantly, assuming Jack satisfies all conditions of IMS and the utterance is *not* a tic, then he is morally responsible for saying something hurtful to a friend.

**Mona**

Mona has Asperger’s Syndrome, a type of Autism characterized by severe social impairment including, but not limited to: lack of desire or ability to interact with peers, a lack of appreciation for social cues, socially or emotionally inappropriate behavior, all-absorbing narrow interest in a particular subject, the imposition of interests and routines on the self and others, speech and language problems (developmental delay, perfect but superficial language structure, difficulty distinguishing between literal and implied meanings), and non-verbal communication problems (limited or inappropriate expression, use of gestures, etc.). To put it mildly, Mona’s condition presents a host of additional challenges when it comes to negotiating social situations. If, having little previous experience with children, Mona is gifted with a painting by her neighbor’s five year old daughter, she may not be morally responsible for hurting the little girl’s feelings when she replies, “This looks nothing at all like an elephant. It is a green and purple blob connected to a brown circle. Elephants are not green and purple.” Mona may, however, be responsible in other social contexts of a sort that she is more experienced with.

**Nick**

Nick was dramatically under-socialized as a child; he spent most of his time locked alone in a small room with little to no social interaction. He comes quite close to being what is often described as a “feral child,” or “wild child.” Due to his early isolation he failed to develop critical communication skills, and his lack of early development has greatly reduced his capacity for developing such skills in the future. Nick may never acquire the level of social development required by IMS, and to the extent that he doesn’t, in relevant contexts, he cannot be found morally responsible.

---

28 Though the category of Asperger’s Syndrome has been removed from the DSM-V, it is still widely used in clinical practice and within the Autism community.
Jack, Mona, and Nick

Each of the three cases presented I take to be relatively uncontroversial. I should think that most people would not ascribe moral responsibility to Jack, Mona, or Nick in the cases above. However, it is important that we are clear about why and to what extent we are not finding them morally responsible. As problematic as a blanket assumption of social competence can be for those who find social situations challenging or impossible, a blanket assumption of lack of responsibility is at least as harmful. This is why it is insufficient to narrow attributions of moral responsibility to an agent; such attributions must be narrowed to a particular agent, in a particular context, at a particular time. Otherwise, the claim that Jack, Mona, or Nick is not responsible in the cases above can have the insidious consequence of considering Jack, Mona, and Nick to be non-agents of a particular sort, always and forever inappropriate targets of our morally reactive attitudes. While this assessment may hold (in many to most cases) for Nick, it almost certainly does not hold (in most cases) for Jack and Mona.

I wish to use these cases as a further introduction to what Applied Behavior Analysis can bring to this discussion. I have said that I take it to be fairly uncontroversial that we should not find Jack, Mona, or Nick responsible in the above cases. The important lesson from ABA, however, is that it is often in everyone’s best interest that we still hold Jack, Mona, and even Nick responsible. To hold responsible, in this sense, is to put the agent in contact with consequences that are the same or relevantly similar to those they would contact if they were found to be responsible.

Jack may not be responsible for his vocal tic in the above case. He can, however, be expected to apologize, just as any agent might apologize for inadvertently stepping on someone’s foot. The harm was not intentional, but occurred nonetheless with the agent in question as the source. Additionally, should Jack notice an increase in his tics when anxious or stressed, he may look into techniques to help manage his stress and anxiety in the future. To blithely excuse every objectionable utterance, to act as if it did not happen at all, is to project onto Jack a level of helplessness that erodes his agency. There is great potential for this helplessness to become internalized and self-fulfilling.29 Further, consider how Jack is to be treated on the day that he feels a person or situation warrants an offensive utterance, and so utters – a history of failing to

29 See Ann Cudd’s discussion of stereotype threat in Analyzing Oppression.
hold responsible can result in the experience of a pernicious sort of invisibility for the agent in question.

Similarly, while the circumstances might be such that we do not find Mona responsible for hurting a child’s feelings, there are some very good reasons for allowing Mona to contact some of the natural consequences of doing so. Holding responsible need not be accompanied by any sort of venomous or mean-spirited affect (as a general rule, and not just in the case of Mona), and an opportunity to witness the child’s hurt and understand her own involvement in what preceded that child’s hurt will help Mona to better negotiate similar social contexts in the future.

**Impulse Control**

Impulse control, emotional and otherwise, is another important part of basic social competence and is closely related to appropriately navigating social context. Actual control, in this case, is far less important than a capacity for control. Compulsions, as irresistible impulses, are beyond an agent’s control in a way that strong urges are not; an agent will fail to satisfy IMS in cases of compulsive action but may still satisfy IMS when acting on impulse if the agent was sufficiently capable of curbing said impulse. Consider once again the cases of Jack and Mona.

Jack suffers from Tourette Syndrome and suffers from tics, both muscular and vocal, that are involuntary and largely out of his control. His tics fall in the same category as compulsions, and so Jack fails to satisfy IMS for actions that are the direct result of his tics. This, however, does not excuse Nick from actions stemming from emotional impulses that he was capable of controlling. Jack fails to satisfy IMS at time t when he says something derogatory and hurtful as a vocal tic but satisfies IMS and is morally responsible when he angrily punches Ben in the face for insulting his favorite sports team at t2.

Mona has Asperger Syndrome but has worked successfully with therapists and learned various techniques for communicating more effectively with others and modifying her behavior to better conform to social expectations and norms. As an adult in everyday situations, Mona will have the capacity to control her impulses; however, just as non-routine situations may result in problems with navigating her social context, they may also diminish her capacity for impulse

---

30 Jack, like many with Tourette Syndrome, can sometimes suppress his tics but only for short periods and with great difficulty that can cause increased stress and more severe tics.
control as the stress of an unfamiliar situation causes her to simply react. In such cases, Mona would fail to satisfy the impulse control component requirement of IMS.

Addiction to various substances can also diminish an agent’s capacity for impulse control. Should the addiction result in impulses that are irresistible\(^{31}\), or very nearly so, an addicted agent will not satisfy IMS with respect to a very narrow range of impulses, namely, obtaining and taking a certain drug. The addict may still satisfy IMS in other contexts and, if so, she is morally responsible regardless of her addiction.

**Access to Reality**

That an agent needs to be able to access reality in an important way is probably the most obvious condition of IMS, but it is often problematically taken for granted. Agents are assumed to have the appropriate access to reality and, as a result, very little thought is put into articulating precisely what such an understanding entails. There are a lot of places where an agent’s relationship with reality can become compromised, and understanding when and how this mitigates or obviates moral responsibility is essential to a full account of morally responsible agency.

**Reality**

Reality is a slippery concept, to say the least and, to be clear; I have no intention of offering a definitive account of reality here. I would prefer to set long-standing philosophical debates aside and, instead, appeal to our common-sense notion of reality. Everyone experiences reality subjectively; this is unavoidable.

Consider an analogy to photography\(^{32}\). In taking a picture, the goal is often to objectively and accurately capture some *real* moment in time. However, even the best cameras cannot fully capture an objective reality. For one thing, the scope of reality is narrowed by what fits into the viewfinder. Already, reality is being filtered. If you zoom in or use a telescopic lens, you are further filtering reality and capturing it in a way that you take to be better, clearer, but is several steps removed from your actual experience. Start using UV, polarized, or color filters and you

---

\(^{31}\) Whether this is ever actually the case is a separate concern that will be addressed, especially in Chapter 6. Certainly the impact of certain treatments calls this into question. See also (Pickard, 2013).

\(^{32}\) This analogy will be revisited at length in Chapter 4.
become that much further removed. Experiment with shutter speed and aperture settings and your pictures may become unrecognizable. Put on the lens cap and you are completely cut off.

Everyone has a camera and most people use zoom and autofocus and the occasional UV or color filter. This shades our reality but not in a way that prevents us from experiencing the world largely as it is. However, when we start making radical adjustments to shutter speed or taking pictures with the lens cap on we are no longer tracking something that is recognizable to others; our picture of reality no longer coheres with shared experience of reality. This is the point at which we are no longer talking about perceptual and interpretive filters but rather perceptual and interpretive barriers.

**Psychoses**

An agent with active psychoses is, perhaps, the clearest illustration of an agent who lacks requisite access to reality. Psychosis is a blanket term meant to cover agents who have lost touch with reality either through hallucinations, delusions, or severely disordered thinking.

There are many different types of hallucination and whether and to what extent they mitigate or obviate moral responsibility must be assessed on a case by case basis. In addition to visual (sight), tactile (touch), auditory (sound), olfactory (smell) and gustatory (taste) hallucinations, there are also hypnagogic (dream-like hallucination upon going to sleep), hypnopompic (dream-like hallucination upon waking), kinesthetic (sense of bodily movement), Lilliputian (perception of objects as smaller than usual), and somatic hallucinations (perception of an experience happening with the body), among others. A simple gustatory hallucination (ex. the taste of metal in your mouth) is unlikely to create the sort of perceptual barrier that would undermine the requirements of IMS. On the other hand, a hypnopompic hallucination or complex visual and auditory hallucinations could easily undermine IMS. The intuition that Jeff should not be found morally responsible for hitting Jane in the face when he is hallucinating that

---

33 (Medicine Net Inc, 2004)  
34 This is not to rule out a simple gustatory hallucination as ever challenging IMS. We can well imagine a highly trained CIA agent who tastes poison in her drink due to a gustatory hallucination. This hallucination, localized to a specific case, could make it such that the agent fails, momentarily, to satisfy IMS. Acting as if she had been poisoned (i.e. taking steps to subdue a perceived threat) might be excused; however, actions unrelated to the hallucination (i.e. robbing a bank) would not. The mental status of the agent was not compromised by the hallucination with respect to the constellation of perceptions, desires, interpretive reasoning, actions and intentions that lead to the bank robbery; not so for the agent’s perceptions and interpretive reasoning in the case of subduing a perceived threat.
she is a threat to his well-being (i.e. attacking bear, home invader, giant spider) is tracking something important, and I would argue that it is Jeff’s failure to satisfy IMS in this circumstance that is underwriting this intuition.

Delusions can be categorized as irrational beliefs that an agent holds with a high level of conviction and persists in believing in the face of overwhelming evidence to the contrary. Delusions, disordered thinking, and other associated disorders can provide both perceptual and interpretive barriers depending on their severity, persistence, duration, content, and context. This will be discussed at greater length in Chapter 5.

**Other Perceptual and Interpretive Barriers**

Broadly construed, there are three types of perceptual and interpretive barriers with the potential to undermine IMS: native barriers, learned barriers, and induced barriers. Though it is important to distinguish between the origins of these barriers in order to understand their impacts on IMS, it is also important to recognize that they are multidimensional and often go hand in hand.

Native barriers to perception and interpretation have a physiological basis. Examples of barriers of this type include Autism, brain tumors, and schizophrenia. There is a range with respect to severity of impairment caused by native barriers, meaning that some may undermine IMS while others do not or that some may only hinder IMS in certain situations and not in others. That said, as far as it is fair to generalize about such potentially variable and disparate impediments, as a class, native barriers will more often offer a substantive challenge to IMS than either learned or induced barriers. This is because native barriers have a greater potential to act as total barriers to adequate perception and interpretation. Learned and induced barriers are almost certainly more common, but are much more likely to act as partial barriers.

Learned perceptual and interpretive barriers can be very difficult to tease out. What I have in mind are biased and even irrational systems of information processing that have become more or less ingrained in our perception and interpretation of the world. These barriers are not inevitable and can be challenged and even dismantled. However, their internalized second-nature

---

35 I have chosen to distinguish between these barriers based on their origin. I find this a convenient and useful way to talk about them, but recognize that it is by no means the only way to parse the landscape. Severity of impairment is, of course, more important for determining their impact on IMS, but would have to be assessed on a case by case basis and would preclude, to some extent, a discussion about types of barrier/impairment.
status can render them largely or completely invisible to the agent (especially when surrounded by like-minded people) and render them somewhat immune to the powers of critical self-reflection. Patriarchy, bigotry, and dogmatic religious beliefs can all function as this sort of barrier. Most often learned barriers will function as partial barriers to perception, interpretation, and reasoning, rather than total ones.

Induced barriers tend to be of limited duration and are the result of some external force acting on the agent. These barriers may be self-induced, as with the voluntary consumption of alcohol and other types of drugs, induced by others (ex. blow to the head), or imposed by some other circumstance or set of circumstances (severe dehydration or extreme emotional trauma). In extreme cases, induced barriers may also be total barriers (ex. when they trigger elaborate hallucinations or delusions), but more frequently, induced barriers are partial barriers. Of all the barriers discussed, self-induced barriers inspire, perhaps, the most difficult questions with respect to assessments of moral responsibility.

In Chapter 5, I will discuss what constitutes appropriate and adequate access to reality in greater depth, building on my photography analogy, and offering an analysis of specific cases. This should shed some light on what sorts of barriers would preclude us from finding an agent morally responsible as well as whether and to what extent it might be appropriate to hold such agents responsible.

**In Conclusion**

In this chapter, I have laid the foundation for my Integrated Mental Status condition for morally responsible agency and the analysis to follow. I have discussed each of the 3 components of IMS: the ability to act for reasons, basic social competence, and access to reality and introduced related concepts from Applied Behavior Analysis that will be important in later chapters.

The next 3 chapters will take each of the 3 major components of IMS in turn. Chapter 3 will address the ability to act for reasons, focusing particularly on how it compares and contrasts to the theory of reasons-responsiveness put forth by Fischer and Ravizza. Chapter 4 is devoted to explaining basic social competence, comparing it with Walter Glannon’s “normative competence,” and answering relevant applied questions about the role of empathy in my theory and how best to consider things like autism and psychopathy. Chapter 5 will investigate subjectivity and access to reality in greater depth and detail further how certain capacities can be
shaped and developed. Chapter 6 will tie everything together, demonstrating the benefits of an interdisciplinary interpretation featuring IMS for assessing responsibility in real-world cases.
CHAPTER THREE

ACTING FOR REASONS

In this chapter, I will elaborate on the first part of the Integrated Mental Status condition (the ability to act for reasons) and address how it relates to Fischer and Ravizza’s Moderate Reasons-Responsiveness. After explaining Fischer and Ravizza’s view, and the standard criticisms, I will also consider whether the sort of mechanistic model that they propose has advantages over an agential model. Ultimately, I will argue that IMS adequately captures, and expands on, the best parts of Moderate Reasons-Responsiveness while avoiding the pitfalls of Fischer and Ravizza’s model. I will argue further that it is a benefit of IMS that it is able to do so without positing the sort of mechanism favored by Fischer and Ravizza.

**Fischer and Ravizza**

Fischer and Ravizza take a Strawsonian approach to defining morally responsible agency, or at least, targeting the set properly understood to be morally responsible agents. Someone is a morally responsible agent insofar as she is an appropriate candidate for the reactive attitudes; the agent is morally responsible so long as she is *rationally accessible* to the reactive attitudes (a potential candidate) even if she is not the actual recipient or target of any reactive attitudes (Fischer & Ravizza, 1998, 7).

Fischer and Ravizza explain their methodology/approach as seeking a “wide reflective equilibrium” through the consideration of cases (real and hypothetical) relevant to moral responsibility and weighing “considered judgments” and “intuitive judgments” against a range of related examples, facts about human nature, society, etc., to seek deeper patterns and truths that cohere with the knowledge we have of the natural and social sciences (Fischer & Ravizza, 1998, 11). Their intention is not to offer knock-down arguments for any part of their theory, but rather to appeal strongly to the intuitive plausibility of their view (Fischer & Ravizza, 1998, 11).

For Fischer and Ravizza, moral responsibility is something deeply tied to our personhood. We are involved in deep personal relationships, we believe that we have some control in our own lives, and we naturally hold each other accountable for things like choices, actions, and intentions. This conception of moral responsibility drives Fischer and Ravizza’s contention that their view of moral responsibility does not hinge on the truth or falsity of
determinism. They claim that though “…causal determinism is a potent doctrine, it does not seem to us that its truth should require us to abandon our view of ourselves as persons” (Fischer & Ravizza, 1998, 16). I will state here that I am very much in agreement with Fischer and Ravizza on this point and highly sympathetic to their overall project. I am also primarily interested in determining who is an appropriate candidate for our morally reactive attitudes\textsuperscript{36}; our disagreements, where they arise, are less about methodology and more about the content of the best (most intuitively plausible) account of moral responsibility.

**Guidance Control and Weak Reasons-Responsiveness**

The standard excusing conditions for moral responsibility are ignorance and force (coercion). Traditional accounts of moral responsibility often have positive conditions built into them in order to account for this. The **epistemic condition** captures the intuition that the agent is responsible only if she knows the particular facts surrounding her action, and acts with the proper sort of beliefs and intentions, and the **freedom-relevant condition (control condition)** insists that the agent must act freely and not as the result of undue force. In other words, in order to be morally responsible, the agent must understand what she is doing and the agent must be able to control her behavior in the right sort of way.

Fischer and Ravizza focus on the freedom-relevant condition and (explicitly) not the epistemic condition\textsuperscript{37} in developing their theory of morally responsible agency (Fischer & Ravizza, 1998, 13). Using a Frankfurt-style case, “Assassin,” Fischer and Ravizza contend that moral responsibility does not require “the sort of control that involves genuinely open alternative possibilities” (Fischer & Ravizza, 1998, 30). There are, however, two sorts of control that need to be distinguished from each other: guidance control and regulative control (Fischer & Ravizza, 1998, 31).

The relevant dimensions of guidance control, the control that Fischer and Ravizza take to be essential for moral responsibility, are whether the mechanism that the agent is acting from

\textsuperscript{36} Which is to say, determining who (given their context and mental status) it is appropriate to find responsible. Holding responsible is something that I wish to hold separate and while holding responsible will often follow neatly and appropriately from finding an agent responsible, the two come apart in important ways.

\textsuperscript{37} It should be clear that my view, the Integrated Mental Status condition, addresses concerns surrounding both the epistemic and control conditions. While the first part of IMS, the ability to act for reasons, borrows heavily from Fischer and Ravizza and, thus, will be concerned more (though not exclusively) with the control condition, basic social competence and access to reality (parts 2 and 3) speak to a number of epistemic concerns.
is the agent’s own and the degree to which that mechanism is reasons-responsive. Alternative possibilities are not required for guidance control, and so it is compatible with causal determinism. **Regulative control**, on the other hand, involves a *dual* power – the power freely to do some act A, *and* the power freely to do something else. It is the dual power to exercise guidance control in either of two or more ways (Fischer & Ravizza, 1998, 31), and it is this sort of control that requires alternative possibilities.

While the **freedom relevant condition** captures the appropriate intuition, theorists have long misunderstood the precise nature of the control required to satisfy it, according to Fischer and Ravizza. Frankfurt-style cases show that guidance control and regulative control can be teased apart, and when the facts of the case allow us to do this, it becomes clear that the relevant sort of control is guidance control (Fischer & Ravizza, 1998, 33). If it is true that guidance control is the type of control that grounds moral responsibility, then agents can have control without having alternative possibilities, and the challenge of causal determinism has no teeth; ruling out alternative possibilities will not entail a lack of control or corresponding lack of moral responsibility (Fischer & Ravizza, 1998, 34). Responsibility-undermining factors for Fischer and Ravizza, then, are those that prevent guidance control: hypnosis, brainwashing, direct manipulation, psychotic impulses, certain types of coercion, etcetera (Fischer & Ravizza, 1998, 35-36).

Additionally, Fischer and Ravizza propose an actual sequence account of moral responsibility. In an actual sequence account, what is important is what agents actually do and how their actions come to be performed. This can be understood in contrast to the sort of alternative sequences present in Frankfurt cases; the actual sequence account explicitly does not require alternative possibilities (Fischer & Ravizza, 1998, 37).

Fischer and Ravizza think that a mechanism-based approach is a better way to get at this than an agent-based approach (Fischer & Ravizza, 1998, 38). The relevant *mechanism* in cases where the agent is morally responsible is reasons-responsive; however, Fischer and Ravizza don’t offer any means of mechanism individuation beyond our intuitions about the case, nor do they go into much detail about what precisely a mechanism is.

In laying out their account of Moderate Reasons-Responsiveness, Fischer and Ravizza entertain and subsequently reject versions of strong and weak reasons-responsiveness. According to strong reasons-responsiveness: if (with the agent’s same mechanism operating)
there were sufficient reason to do otherwise, 1. the agent would recognize the sufficient reason to do otherwise and thus 2. choose to do otherwise and 3. do otherwise (Fischer & Ravizza, 1998, 41). Strong reasons-responsiveness is not the reasons-responsiveness that Fischer and Ravizza believe is necessary for guidance control or moral responsibility (Fischer & Ravizza, 1998, 42), but it does neatly draw out three potential points of breakdown for moral responsibility that are highly relevant. These potential “alternative sequence failures” (Fischer & Ravizza, 1998, 42) are:

1. Failure to be receptive to reasons
2. Failure of reactivity; failure to be appropriately affected by beliefs
3. Failure to successfully translate one’s choice into action

While Fischer and Ravizza determine that strong reasons-responsiveness demands too much, receptivity to reasons and reactivity to reasons remain central to their overall account.

Alternatively, for weak reasons-responsiveness, we merely need to hold fixed the actual kind of mechanism and require that there exist some possible scenario (or possible world, with the same natural laws as the actual world) in which there exists a sufficient reason to do otherwise, the agent recognizes the reason and the agent does otherwise for that reason (Fischer & Ravizza, 1998, 44). This gets at the sort of reasons-responsiveness necessary for moral responsibility, and Fischer and Ravizza believe that it is plausible to think that it is also sufficient (Fischer & Ravizza, 1998, 45-46).

**Moderate Reasons-Responsiveness**

Ultimately, Fischer and Ravizza determine that weak reasons-responsiveness is too weak a condition and strong-reasons responsiveness is too strong. They settle on a view that attempts to strike the appropriate balance between two extremes: Moderate Reasons-Responsiveness.

Weak reactivity to reasons may be all that is required for guidance control, and thus, moral responsibility, but certainly moral responsibility requires something more than weak receptivity to reasons. More than an ability to recognize a single sufficient reason to do otherwise, a morally responsible agent must demonstrate an appropriate pattern of reasons-recognition that demonstrates an ability to see how reasons fit together, why certain reasons are

---

38 Namely, that any/every sufficient reason to do otherwise would be recognized by the agent and move her to act in accordance with a choice to do otherwise.
stronger than others, and why a stronger reason should always be sufficient where a weaker reason is deemed sufficient (Fischer & Ravizza, 1998, 71). Regular receptivity is required for moral responsibility and regular receptivity to reasons requires an understandable pattern of reasons-recognition, minimally grounded in reality (Fischer & Ravizza, 1998, 73).

Only weak reactivity to reasons is necessary for moral responsibility because “reactivity is all of a piece” (Fischer & Ravizza, 1998, 73), meaning that if an agent’s mechanism reacts to some incentive to do other than she actually does, this shows that the mechanism can react to any reason to do otherwise (Fischer & Ravizza, 1998, 73). According to Fischer and Ravizza, “[t]his general capacity of the agent’s actual-sequence mechanism – and not the agent’s power to do otherwise – is what helps to ground moral responsibility” (Fischer & Ravizza, 1998, 73).

F&R distinguish between moral and prudential reasons in order to further the claim that certain borderline cases (smart animals, children, psychopaths) are not cases involving moral agency because they are not moral agents39. The distinction is as follows: prudential reasons are concerned with an agent’s long-term self-interest, while moral reasons issue from some sort of balancing of one’s own interests against the interests and rights of others (Fischer & Ravizza, 1998, 76). Ultimately, their theory requires a mechanism with regular receptivity to reasons where at least some of the reasons it is receptive to are moral (Fischer & Ravizza, 1998, 77); the mechanism need not be reactive to any moral reasons (Fischer & Ravizza, 1998, 79).

**Mechanism**

Before moving into my critical analysis of Fischer and Ravizza’s view (Moderate Reasons-Responsiveness), it is important to mention some additional background about Fischer and Ravizza’s conception of mechanism and the role that it plays in their theory. Fischer and Ravizza claim that mechanisms must be “temporally intrinsic” or “non-relational”; if mechanism M issues in act A, then M is relevant to the agent’s guidance control of A only if M’s operating does not entail that A occurs (Fischer & Ravizza, 1998, 47). No rule is offered for differentiating mechanisms or isolating the mechanism in any particular case, but Fischer and Ravizza do presuppose that for every action there is an intuitively natural mechanism that can appropriately be called the mechanism (Fischer & Ravizza, 1998, 47). They also argue that their theory applies equally well to cases of non-reflective behavior as cases of practical reasoning (Fischer &

---

39 They cannot be moral agents, because they do not understand moral reasons.
Ravizza, 1998, 86). This means that the mechanism need not be reflective or involve active reasoning in order to be weakly reactive to reasons and regularly receptive to reasons.

**Mechanism: Some Problems**

Moving into the critical portion of this chapter, I will begin with a few concerns about Fischer and Ravizza’s mechanistic model before moving on to other challenges for Moderate Reasons Responsiveness as an account of moral responsibility. I limit myself in this section to two major problems with Fischer and Ravizza’s employment of mechanism: the lack of a means of mechanism individuation and their contention that the mechanism in question can be non-reflective.

I am not the first to point out the problem of mechanism individuation in Fischer and Ravizza’s theory. Michael McKenna does an excellent job elucidating the problem in his 2001 book review of *Responsibility and Control*; I will briefly summarize his points here.

While Fischer and Ravizza get a lot of mileage about of their mechanism-based model, they offer no principled method of mechanism individuation and, thus, open themselves up to legitimate objections on incompatibilist grounds (McKenna, 2001, 95). Fischer and Ravizza’s account of reasons-responsiveness relies crucially on the idea of the “same mechanism” operating in other possible worlds (i.e. holding fixed the mechanism, in some possible world there is a reason(s) that that same mechanism will be receptive and reactive to if that mechanism is reasons-responsive), however, they offer no principled way to individuate mechanisms and rely exclusively on intuition to do the work (McKenna, 2001, 96). The way Fischer and Ravizza define mechanism is more stipulative than informative; they do not mean for it to capture “anything over and above the process that leads to the relevant upshot” (McKenna, 2001, 96).

Fischer and Ravizza’s appeal to intuition simply cannot do the work that they need it to. Their semi-compatibilist commitments combined with their approach of taking a ‘wide reflective equilibrium’ have left Fischer and Ravizza with a conception of mechanism that is ultimately too vague to appropriately undergird their Moderate Reasons Responsiveness.

Another concern is that Fischer and Ravizza claim that the sort of mechanism underwriting moral responsibility (namely the agent’s own reasons-responsive mechanism) can be non-reflective. Fischer and Ravizza argue that their theory applies equally well to cases of

---

40 I will discuss mechanism further when I argue for the benefits of an agential model over a mechanistic one.
non-reflective behavior as cases of practical reasoning (Fischer & Ravizza, 2001, 86). The mechanism need not be reflective or involve active reasoning in order to be weakly reactive to reasons and regularly receptive to them. It is not at all clear how a theory that requires, at base, a suitable pattern of reason-recognition, could be satisfied by a non-reflective mechanism. It is clear why Fischer and Ravizza want to make this claim (to allow us to hold responsible those who are acting from a non-reflective mechanism), but their example of a non-reflective mechanism and the brief discussion that follows aren’t particularly persuasive.

Fischer and Ravizza ask you to consider someone driving to work who gets off the freeway, as he always does, at University Avenue. Though the agent does this without any thought, he is on autopilot so to speak. Fischer and Ravizza claim that he is certainly responsible for turning off the freeway at University Avenue (Fischer & Ravizza, 2001, 86). Imagine that, contrary to the actual sequence, the University Avenue exit is blocked. By holding fixed the actual sequence (non-reflective) mechanism we can imagine a case where the agent (still on autopilot) takes an alternative route to his destination. This proves, according to Fischer and Ravizza that an agent’s actual-sequence action, though it comes from a non-reflective mechanism, can be moderately reasons-responsive (Fischer & Ravizza, 2001, 86).

Though this case seems obvious to Fischer and Ravizza (perhaps due to the familiarity of driving somewhere on autopilot), it is not obvious to me that the agent in the alternative sequence could have driven an alternative route without any reflection whatsoever. I have been driving for a number of years and I have never had the experience of non-reflectively following a detour or non-reflectively taking an alternate (infrequently used) route to a familiar destination when a primary route was unavailable. These situations, in my experience, require more reflection, not less, to get where I need to go.

I also believe that it is important to note here that Fischer and Ravizza use a decidedly non-moral case as their paradigm for non-reflective mechanisms that satisfy their moral responsibility requirements. Fischer and Ravizza suppose two possible interpretations and take a win-win approach that I find unconvincing. We can accept that we can hold fixed a non-reflective mechanism and get reasons-responsiveness (and so moral responsibility) or we can reject that a non-reflective mechanism is compatible even with reasons-recognition and so not reasons-responsive. If we take this second route (which I am inclined to), Fischer and Ravizza argue that it is very natural to apply the tracing component of their theory and see that agents can
be held responsible in cases where they are acting from a non-reflective mechanism when we substitute “behaviors originating from non-reflective mechanisms” for “trait actions” and trace the trait formation back to reasons-responsive mechanisms (Fischer & Ravizza, 1998, 97-98). While I find this tack more plausible than a non-reflective reasons-responsive mechanism, I believe that Fischer and Ravizza are attempting to dramatically downplay both the fact that their theory works best (and perhaps only) when explaining practical reasoning and the extent to which a lack of mechanism individuation inhibits a proper analysis of the role of any given mechanism in determinations of moral responsibility.

More Problems With Moderate Reasons-Responsiveness

I will now return briefly to the chief criticism that I leveled against Fischer and Ravizza’s view in Chapter 2 in the discussion of Mele’s agoraphobic (namely that insane/deeply troubled/compulsive agents can satisfy Moderate Reasons Responsiveness) and mention a few other points about their view that should give us pause.

Grumpy Mutt

Here is a case that I call Grumpy Mutt. Every day, on his walk to work, Jack passes an alleyway at whose entrance sits an old stray dog who never fails to growl as Jack walks by. One day, sick of the growling, Jack decides to kick the dog. Assume that the mechanism that Jack is acting from is Jack’s own and also that Jack is regularly receptive to a variety of sufficient reasons for not kicking the dog: he could seriously hurt the dog, the dog could seriously hurt him, someone might witness the kick and confront Jack or call the authorities…. Now imagine that the only case in which Jack decides that he won’t kick the dog, and doesn’t kick the dog, is the case in which he considers that he could scuff his shoe (assume his shoe is completely unremarkable) and so decides to refrain and refrains for that reason. Jack satisfies Fischer and Ravizza’s criteria for Moderate Reasons-Responsiveness (and therefore, presumably, moral responsibility) but if Jack, moved by the possibility of a scuff is not similarly moved by the possibility of torn clothing or torn skin, then it would seem that Jack is working from a faulty mechanism. Jack seems unbalanced and his reactivity to the possibility of a scuff looks like an idiosyncratic fluke rather than an indication of a general capacity of the agent’s mechanism to respond to reasons.
These sorts of cases make it clear that there is a problem with Fischer and Ravizza’s Moderate Reasons Responsiveness and that it seems to be the reactivity component. Before addressing how Fischer and Ravizza might go about strengthening the reactivity component of their theory, I want to voice a concern about the logic underwriting Fischer and Ravizza’s requirement of such a weak reactivity component. I find that notion that a mechanism’s general capacity for reactivity can be gauged by a single instance somewhat troubling, though Fischer and Ravizza maintain that reactivity is “all of a piece.” Consider the starter on my old Corolla. For the first 7 years that I owned that car, it started up without any trouble every time that I turned the key in the ignition. Then the starter began to fail (defective mechanism) and sometimes it would start up and sometimes it wouldn’t. Then it would most often not start at all, though it would start unpredictably and very occasionally. That the car eventually, most often, would not start, was evidence of a broken/faulty mechanism. The unpredictable and sporadic instances when the car would start up were not evidence that the mechanism was in perfect working order and would work in other cases.

It is clear that Fischer and Ravizza need to strengthen the reactivity component of Moderate Reasons Responsiveness in order to have a workable account. One way they could do this would be to maintain the asymmetry between receptivity to reasons and reactivity to reasons while emphasizing an explicit pattern of fit between the two. Regular receptivity to reasons already requires an ability to recognize patterns of reasons that conform (more or less) to reality, but a more explicit ranking/sorting/weighing of reasons could be built into this notion as well. This wouldn’t necessarily require the agent to consciously rank each reason, but it would establish a certain threshold, something more than mere sufficiency, beyond which the reason is strong enough/carries enough motivational force to produce action. Reasons that the agent would place above this threshold should generally, if not always, lead to a logical pattern of fit that indicates a general coherence between the mechanism’s ability to recognize reasons and its ability to act on them.

Though an improvement, this solution still cannot overcome the problems presented by Fischer & Ravizza’s mechanistic model. Ultimately greater consideration needs to be taken of the agent herself; part of the story about reasons-responsiveness, if it is to serve as a good basis

---

41 Fischer and Ravizza do mention a pattern of fit in discussing receptivity to reasons but it is left vague and they do little to draw a connection to reactivity.
for assessments of moral responsibility, must involve the agent herself – what she values, what motivates her, how she interprets reality, her mental status, etc. Not only do Fischer and Ravizza leave this out of their theory, but their theory attributes moral responsibility based on agent mechanisms that, while responsive to reasons, are responsive irrespective of an appropriate weighing and ordering of various reasons.

**Agent vs. Mechanism**

As with some other concepts employed by Fischer and Ravizza in their account of moral responsibility, that of an “agent’s own reasons-responsive mechanism” is left somewhat fuzzy around the edges. The way that this concept is employed seems to suggest that the sort of mechanism in questions is both distinct from and essentially a part of the agent. In attempting to unpack this phrase that Fischer and Ravizza rely on to do so much work, we are left with questions: if not the agent’s own mechanism, then whose? And, if not the agent’s rational capacities, then how could it be reasons-responsive? I will address each in turn.

**Whose mechanism?**

There are at least two senses in which Fischer and Ravizza could mean “agent’s own” here: 1) not belonging to some other agent/outside force and 2) belonging to something like an agent’s “authentic self.” The first sense is relatively uncontroversial and even obvious. There is general agreement that extreme coercion, evil neuroscientists, and secretly implanted brain-control devices, when active in the actual case, will have at least a mitigating and probably an exculpating effect on the moral responsibility of the victim of such forces. If all Fischer and Ravizza were trying to do with “agent’s own” is lend clarity in the face of Frankfurt-style cases, while it doesn’t seem that it is doing all that much work for their view, it would at least be fairly uncontroversial. However, this is not all that Fischer and Ravizza intend. Given, especially, their discussion of history and responsibility, it is clear that “agent’s own” is meant in something like the second sense as well.

**History**

Fischer and Ravizza insist that moral responsibility is a historical notion; it cannot simply be read off of current time-slice assessments of the agent, but rather, we must consider the

---

42 Credit for this question goes to a seminar discussion in a class taught by Michael McKenna.
43 See Chapter 7 of *Responsibility and Control.*
history underwriting current traits of the agent. Fischer and Ravizza explain the difference between epistemically and genuinely historical as follows: “[w]hen a phenomenon is epistemically historical but not genuinely historical, there is something in the snapshot features of the present that fully grounds the judgments in question and nothing is entailed about the past” (Fischer and Ravizza, 1998, 190). They go on to state that the distinctive thing about epistemically historical phenomena “…is that features of the past point us to snapshot features of the present; features of the past are useful in helping us characterize and understand the present, but all we ‘really’ need to understand the present is contained in the snapshot features of the present” (Fischer and Ravizza, 1998, 190).

Fischer and Ravizza claim that moral responsibility is more than merely epistemically historical, and here we reach a crucial point of disagreement. It seems to me that moral responsibility is most certainly epistemically historical in the way stated above, but that it is not genuinely historical. That is to say, while details about the past certainly shed light on our understanding of the present, we can read moral responsibility off of the so-called snapshot features of the case. The problem with other agential views about moral responsibility is not in choosing agent over mechanism or in asserting that moral responsibility can be read strictly off of the present features of the case, but in picking out which snapshot features are relevant to moral responsibility. My Integrated Mental Status condition, at base, can be understood as an attempt to latch onto the appropriate snapshot features.

It isn’t difficult to think of cases where agents with similar (lacking relevant dissimilarities) or same (holding fixed in considering alternative sequences) histories will evoke different ascriptions of moral responsibility. This is because moral responsibility is so deeply tied to the agent herself and not some mechanism that can be pulled apart from the agent, or an historical sequence which determined the agent’s actions. Moral responsibility can be epistemically historical insofar as history helps us to better understand and appreciate the snapshot features of the agent, but it is those snapshot features that determine whether the agent is morally responsible in any particular case.

---

44 One way to think about the difference between my approach and F&R might be that they explicitly ignore the epistemic condition for moral responsibility and I don’t.
The Twins

Imagine identical twins Jessie and Jenny. Jessie and Jenny grow up together in a white, upper-middle class household. Both girls are smart, attractive, well-liked, successful at most things that they try, and well-loved by both of their happily married parents. Their father dies when they are 10. Their mother, Marla, remarries 2 years later, and while he is initially charming, the girls soon realize that their stepfather, Harold, is conniving, abusive, and unpredictable. The girls are subjected to intense mental, physical, and sexual abuse; they also watch as their mother is abused but refuses to leave her new husband. Whenever they beg Marla to leave Harold, Marla insists that Harold loves all of them very much and that they all need to work harder not to upset Harold. The girls spend the next 6 years in a state of fairly constant fear, depression, and confusion until they are finally able to move out on their own.

Jessie begins her life as an adult damaged but determined. If she has learned anything, it is that she needs to work hard to never find herself in Marla’s position. She will work hard, she will support herself, she will insist on equality in all of her romantic relationships, and she will do her best to help who she can along the way, since nobody was there for her when she needed help the most. Jessie works her way through college and eventually starts a non-profit to help victims of abuse.

Jenny begins her life as an adult differently. She cannot relate to people; she cannot connect on a personal level; she hates herself for what has been done to her. Jenny frequently cuts herself in an attempt to feel something, anything, and when she can’t find the satisfaction that she needs from cutting, she turns to drugs in an attempt to escape from herself.

Despite a similar history (and I will just stipulate that it wasn’t in any relevant ways dissimilar) up until the point that they moved out of the house, Jessie and Jenny were very different agents with very different snapshot features when they turned 18 and moved out on their own. I don’t think that it is too much of a stretch to say that most people, knowing the girls’ shared past, will still want to find 18 year old Jessie more morally responsible in more cases than Jenny. The

45Related point: in psychologist Jonathan Haidt’s book, *The Happiness Hypothesis*, he offers a persuasive explanation of a phenomenon that he calls “post-traumatic growth.” People can be exposed to the same traumatic circumstances, but while some people experience (sometimes crippling) post-traumatic stress, others experience post-traumatic growth that makes them more resilient and adaptive to challenges in the future. He argues that there seems to be a critical window corresponding to age as well as methods for processing the trauma that increase the likelihood of this post-traumatic growth. In any case, it furthers the point that similar/same histories can result in very different future snapshot features of agents.
snapshot differences, in the agents themselves, are what is relevant for determining moral responsibility. Notice that despite different snapshot features of 18 year old Jessie and Jenny (which set them on radically different life trajectories) moral responsibility may still be described, in both cases, as epistemically historical. Knowing the history of the agent promotes a better understanding of the agent and her snapshot features, but the very same history can help explain any number of relevantly dissimilar snapshot features.

Just as similar histories can lead to very different snapshot features, very different histories can lead to similar snapshot features. I am not insensitive to the very real ways in which history shapes agents. It tends to do so in fairly predictable ways, otherwise, applied behavior analysis would be impossible. If an agent’s history has affected an agent to the extent that she should not be held responsible in a particular case, then that should be evident from an examination of the snapshot features. I argue that if, upon examination of the snapshot features, IMS is satisfied, then the agent in question may be morally responsible (to some degree) regardless of history. If, however, IMS is not satisfied, then the agent cannot be found responsible.

This partially explains why I think that an agential model that can track the relevant snapshot features of an agent’s mental status is ultimately more accurate and more useful for explaining moral responsibility than a mechanistic model that insists moral responsibility is more than epistemically historical.

**The Integrated Mental Status Condition and the Ability to Act for Reasons**

Criticisms notwithstanding, Fischer and Ravizza have laid important groundwork for understanding Reasons-Responsiveness and, by extension, the first component of my Integrated Mental Status condition. Perhaps the best place to start is with a return to Fischer and Ravizza’s analysis of where Reasons-Responsiveness can break down. Here are the 3 possible failure points discussed earlier in this chapter:

1. Failure to be receptive to reasons
2. Failure of reactivity
3. Failure to successfully translate one’s choice into action

The first point has to do with reasons-recognition; without access to available reasons, it is not possible to be moved by those reasons to act. Though I am currently focused on discussing
the first component of IMS, this seems a good place to point out why I refer to my condition as an *Integrated* Mental Status condition. For the purposes of explanation and analysis, I have broken my condition into components and subcomponents; however, it is important that we do not lose sight of the fact that it is the functional *integration* of all of these components within the agent that is at stake. I take it as an advantage of my IMS condition that it offers a reasonable explanation of various ways that a failure to be receptive to reasons could come about. One could fail to be receptive to reasons due to momentary or chronic inattention or myopia, but could also fail to recognize reasons due to a lack of basic social competence, or some perceptual or cognitive barrier(s) that prevent the relevant access to the reality of the situation. To be clear, my condition requires more than a general capacity to recognize reasons, it requires a capacity to recognize reasons that is at least minimally operational at the time in question.

Failure of reactivity to reasons means that, when presented with reasons, there is a failure to be appropriately moved to action. We might think of this as a motivational problem. Even when presented with a reason that should be sufficient to move a person to action, one is not motivated to act, or even to decide to act, based on that reason. Again, we can imagine very different ways in which this could come about. The psychopath might seem the prototypical candidate to exemplify this sort of break down; even recognizing reasons to act (in an academic sort of way) the psychopath may not *care*. A lack of empathy can certainly lead to a failure of reactivity, but so too can other motivational barriers. Consider someone who is severely depressed, or in shock, or has simply reached that point that many of us reach, on occasion, of ‘too tired to care.’ The Integrated Mental Status condition has the resources to allow for an analysis of each of these states (and virtually any others that come to mind) with respect to moral responsibility.

Many philosophers will think immediately of cases of so-called ‘weakness of will’ and the attendant debates. It is important to recognize, however, that there are many ways in which an agent can fail to successfully translate her choice into action. In Applied Behavior Analysis, a major consideration in determining why a person isn’t engaging in a particular behavior when there seems to be sufficient motivation (read: reinforcement contingencies in place) is whether there is a skills deficit (a person is unable to do something) or performance error (misfire with

---

46 This should be broadly construed.

47 I recognize that this might be a contentious claim, as some theorists would argue that psychopaths cannot recognize reasons of a particular sort (i.e. moral reasons).
respect to performing some behavior that is within the agent’s repertoire, due perhaps to misreading environmental cues, etc.) at work. Furthermore, there are environmental barriers to consider. I may recognize the need to call a nurse for my hospital roommate who seems in distress, be motivated to call for a nurse and decide to call for one, and still be unable to translate this choice into action due to the paralytic effects of a medication that I am on.

Now for a point that I wish to make about acting for reasons, the significance of which will become more apparent as we move forward. Virtually everyone has the ability to act for reasons in a limited sense. Most people have the ability to act for reasons in the weakly-reactive sense that Fischer and Ravizza claim is necessary for Moderate Reasons-Responsiveness. In fact, many people are weakly reactive to reasons without being regularly receptive to reasons. These are actually very important lessons that I have taken from my work with autistic children\[48\].

The reason why ABA therapy is so effective in changing people’s behavior (read: increasing reactivity) is that it capitalizes on the fact that there is at least one reason (consequence/reinforcer) that will motivate an agent to act. Delivering that reinforcement contingently can change behavior. The reinforcer powerful enough to induce said change may look very different from moral considerations and may not fit into any discernable pattern of fit about reasons for action. Perhaps the reason that James has for returning Emily’s greeting, from James’ perspective, is that he will receive 5 purple skittles for doing so. Over time, as those skittles are paired with other stimuli (praise, Emily smiling at the returned greeting), the previously neutral stimuli become conditioned reinforcers\[49\]. Where James was previously motivated only by a singular very idiosyncratic contingency (purple skittles), exemplifying weak reactivity, conditioning can grow the number of reasons that can motivate James to act in these sorts of cases, increasing both his receptivity and his reactivity to reasons. Eventually, the natural contingencies (ex. Emily smiling in return) will sufficiently motivate James to act (return greeting), even in the absence of the possibility of receiving purple skittles. Reactivity can be, and often is\[50\], strengthened over time.

---

\[48\] This should in no way be taken to indicate that autistic children are unique in this respect.

\[49\] This is straightforward operant conditioning; think Pavlov.

\[50\] Just consider the way that very young children learn. Initially they will be far more motivated to act for a reaction from a caregiver than for any further-reaching concerns.
This is also why holding responsible, even in the absence of finding responsible, can be very important. To be clear, when I say hold responsible, I mean allow to contact (some) of the usual consequences for a particular action, and not something like blame the agent for something that she is not responsible for. Holding responsible in the absence of finding responsible is instructional and should be done without rancor. It is a means of increasing both receptivity and reactivity over time. James does not satisfy IMS when only skittles can move him to act, but if he is made to say hello (skittle access is contingent on saying hello), contacts reinforcement for saying hello, contacts additional stimuli that become conditioned reinforcers, and does this over enough trials that the skittles can be withdrawn and he is still motivated to act, then James might satisfy IMS at some future time when presented with similar situations.

Problems Solved/Benefits of an Integrated Mental Status Condition

In concluding this chapter, I want to very briefly highlight what I take to be the major advantages of my Integrated Mental Status condition (an agential model) over Fischer and Ravizza’s mechanistic model for moral responsibility. With an Integrated Mental Status condition, there is no need to posit anything in addition to the agent to explain moral responsibility of that agent. On my model, there is no problem of mechanism individuation or assessing whether the mechanism is the agent’s own in the relevant sense. There is no question of how the mechanism relates to the agent; there just is the agent. While it would be disingenuous to say that my Integrated Mental Status condition is a simple means of determining moral responsibility, I do believe that it is fair to say that it is not vague in the ways that many of Fischer and Ravizza’s most central concepts are vague. With IMS, assessments of moral responsibility are contextual and may be complex, but the standards for satisfying IMS are, at least, useful and will be explained in more depth in the coming chapters.

IMS has the additional benefits of avoiding problems like the Crazy Problem (laid out in Chapter 2), addressing the concerns inherent in positing a rationality requirement while capturing the important/relevant features of the agent that a rationality requirement (alone) cannot. Though I have no desire to take on the free will debate here, an agential view like mine could plausibly work within the sort of semi-compatibilist framework that Fischer and Ravizza propose, and so that potential benefit need not be lost.
In Chapter 4, I move on to the Basic Social Competence condition of IMS and explain how it compares to Walter Glannon’s concept of Normative Competence.
CHAPTER FOUR

UNDERSTANDING BASIC SOCIAL COMPETENCE

My proposal of an integrated mental status condition for morally responsible agency is not completely unprecedented in the philosophical literature, Walter Glannon insists that a "satisfactory account of the conditions under which a person can be responsible for her actions, omissions, and other events requires a careful analysis of the ways in which her mental states and events play a causal role in the etiology of her behavior" (Glannon, 2002, 31). Glannon’s project in *The Mental Basis of Responsibility* (2002) is somewhat more ambitious than mine; it seems that his “…unifying concern is to make a case for compatibilism with an emphasis on the mental basis of responsibility, with respect to both theoretical reasoning and practical reasoning” (Jacobs, 2003, 711). Though I have compatibilist sympathies, an argument for compatibilism is beyond the scope of this project. My aim is to present a necessary condition for morally responsible agency that rests with the mental status of the agent. To that end, it is important to consider Glannon’s account of normative competence and its overlap with my account of IMS, especially with respect to Basic Social Competence. In what follows, I will briefly outline Glannon’s Normative Competence, emphasizing the similarities with my own account but ultimately arguing that our points of departure are critical, as evidenced by our disparate stances with respect to understanding empathy, tracing accounts of moral responsibility, and the moral responsibility of psychopaths. I will then move into a discussion of the Basic Social Competence aspect of IMS, focusing critically on the roles of empathy and attentiveness in assessments of moral responsibility.

Normative Competence v. Integrated Mental Status Condition

Glannon’s Normative Competence consists of three parts: “…the cognitive capacity to have and respond to reasons, the affective capacity to have and respond to emotions, and [the] volitional capacity to express intentions in choices and actions…” (Glannon, 2002, 31). Glannon embraces a reasons-responsiveness view, like that of Fischer and Ravizza in order to satisfy the cognitive component of Normative Competence, but he argues that their view fails to properly take into account the role of emotions like fear, joy, empathy, shame, and remorse in agential deliberation and action (Glannon, 2002, 31). In his discussion of affective capacities, Glannon
attempts to draw out the interrelatedness of an agent’s affective, cognitive, and volitional capacities. Glannon relies on Jon Elster’s distinction between occurrent and dispositional emotions (Glannon, 2002, 35) and points out that while an agent has little control over her occurrent emotions, she can cultivate certain dispositions and develop integrative strategies for responding to various occurrent emotions. Cognitive capacities work with affective capacities, according to Glannon, in a feedback loop, and the “capacity for both cognition and emotion as interacting, interdependent mental faculties gives us reflective self-control over the springs of our actions” (Glannon, 2002, 36).

My view shares many similarities with Glannon’s. We are both offering capacity views that endorse a graduated view of moral responsibility based on mental states of the agent. Like Glannon, I emphasize the importance of a modified version of Fischer and Ravizza’s Moderate-Reasons-Responsiveness (see Chapter 3) in identifying core aspects of the cognitive features required of an agent to qualify as morally responsible. I agree further with Glannon that it is important to understand the interrelated nature of an agent’s cognitive, affective, and volitional states and, importantly, that neuroscience can aid us greatly in efforts to better understand agents and, in some cases, assess moral responsibility.

That said, there are also important differences between our views. To begin with, we parse the landscape differently. The capacities underwriting Glannon’s Normative Competence are cognitive, affective, and volitional. While, these capacities are important for my view, they are spread across the components of IMS; the capacity to act for reasons has cognitive, volitional, and affective components (ex. where emotion plays a role in reasons-recognition), basic social competence has cognitive, affective, and volitional components (ex. impulse control), and accessing reality has, at least, cognitive and affective components. Perhaps this is owing to what I take to be different approaches to the question of moral responsibility. Glannon is offering an account of how an agent’s mental states work together to achieve reflective self-control and, thus, moral responsibility. I am most concerned with how the mental status of an agent can cause moral responsibility to break down, in identifying the nature of those breaks, and identifying a threshold above which agents may be found somewhat morally responsible and below which they are not found responsible at all.

While Glannon believes that moral responsibility comes in degrees and acknowledges that agents with certain psychological disorders should be considered less responsible, or not at
all responsible, for actions springing from their disorder, he is more rigid about ascriptions of moral responsibility for those without psychological disorders claiming that “…although acting ‘in the heat of passion’ may affect intent and change a legal judgment from first degree murder to manslaughter, it does not affect the fact that the person would deserve a complete, not diminished, ascription of responsibility for the action” (Glannon, 2002, 36). This is because Glannon believes that the capacity for reflective self-control ultimately determines whether an agent is morally responsible and that capacity, for Glannon, seems to be something that an agent either has or doesn’t have (whether or not they are exercising that capacity).

My account is too contextual to allow for a generalization like the one above; whether an agent ‘in the heat of passion’ was fully morally responsible, partially responsible, or not responsible at all would depend entirely on the agent’s mental status at the time of the action and whether the agent satisfied each part of IMS. A further point of departure between my view and Glannon’s, is that he endorses tracing accounts of responsibility, namely that if we can trace the action, X, of an agent with diminished capacities to some prior action, Y, of that same agent when her capacities were intact, then the agent is fully responsible for X. Glannon writes, “[t]here are instances, however, where a person may act in a state of psychotic delusion and still be responsible for his action. This would be the case if a schizophrenic were able but refused to take antipsychotic medication. We could trace his deluded state of ignorance and unfree action to an earlier free act of refusing to take medication, at which time he could foresee the likely consequences of not taking it” (Glannon, 2002, 49). Many philosophers appeal to tracing accounts in order to ascribe full moral responsibility to agents acting with diminished capacities for one reason or another. I take it as a strength of my view that I can offer a satisfying account of moral responsibility that does not rely on tracing. In Chapter 6, I will offer an in-depth analysis of classic tracing cases in order to demonstrate how agents under the influence (of drugs, trauma, psychoses, etc.) can be assessed without resorting to tracing.

**Empathy**

Glannon does little to define empathy (or any of the emotions critical to his account) but does claim that “it is appropriate to characterize empathy as an emotional state with a cognitive antecedent, rather than as a mental state with cognitive and affective aspects” (Glannon, 2002, 36). This goes to Glannon’s characterization of the role of cultivating dispositions and mental strategies for responding to reasons.
Knowing that I have harmed someone by acting out of fear or anger and feeling remorse or guilt for having done so can motivate the belief that cultivating a general disposition on which to act will avoid harm and remorse in future circumstances. Reminding myself that the person I have harmed has needs and interests is likely to generate the disposition of empathy toward him and other persons. This illustrates how most higher-level, ‘moral’, emotions have cognitive antecedents (Glannon, 2002, 36).

I will not argue the importance of cultivating moral dispositions, indeed, this is of importance to my non-tracing approach to ascriptions of moral responsibility as well as my contention that it is sometimes most appropriate to hold someone responsible even in the absence of finding that person morally responsible. That said, it is unclear, even from Glannon’s explanation, why we should consider empathy to be an emotional state with a cognitive antecedent rather than a multi-faceted and dynamic mental state. Glannon seems to take the road to empathy to involve a sort of internal cognitive prompting that results in empathic feeling. This internal cognitive prompt, however, seems to have been motivated by emotions (i.e. the prior remorse and guilt of the agent). The interplay of cognition and affect are key to understanding empathy, even on my own view, but Glannon offers no good reason for thinking that empathy must have a cognitive antecedent like that described above and may be guilty of over-intellectualizing the internal processes of agents. Some agents may require cognitive prodding toward empathic feeling, but Glannon fails to consider that some almost certainly do not. It seems clear that empathy could be properly said to have an emotional antecedent in the same way. Consider the following case:

When checking her e-mail one morning, Tara sees a message from a college friend. They had lost touch, but she is excited to see the message. When she opens it, she sees that this message has been sent to many of her college friends and that it concerns yet another friend, Mali. Mali’s parents and four younger siblings had been on vacation visiting family in India. They were in a car accident. Both of her parents and her two younger sisters were killed. Her two young brothers, one of whom has special needs, are in critical condition. Mali has flown to India to deal with the remains of her parents and sisters and to await a time when her brothers are stable enough to travel. Mali’s family in India is distant and poor. There is no other family in the U.S. for her to lean on. She is likely, at the age of 24, to become the sole guardian of two adolescent boys, one of whom will require continuous care for the rest of his life. It is all too much for Tara, she feels like she has been punched in the chest, and before she has made it through the entire message, she finds that she needs to stop and breath, to fight the tears.
streaming down her face. It actually isn’t until after several moments of processing that Tara even gains the awareness that all of this pain and sadness is for Mali. Tara never even met Mali’s family, hasn’t spoken to Mali since college graduation, and has, herself, lost nothing. Nevertheless, Tara feels for Mali before she is consciously aware of it.

In this case, it seems that the emotion comes first and cognition follows. This goes to why it is important to conceive of empathy as having both an emotional and a cognitive aspect rather than a firmly delineated antecedent of one type or the other. If Tara was overcome with deep emotion after reading about her friend Mali, but lacked the cognitive capacity to ever meaningfully connect those feelings with Mali, to understand that she was feeling for Mali and not herself or anyone else, then Tara would not be experiencing empathy. Glannon’s account generally takes our understanding of concepts like empathy for granted, but a concept so central to moral responsibility requires more in-depth examination.

**Psychopaths**

I find Glannon’s treatment of psychopaths both problematic and inconsistent. Glannon claims that it is our capacity for empathy that makes our reasons for action moral (Glannon, 2002, 39), that the properties common to all psychopaths are impulsivity and lack of empathy (Glannon, 2002, 56), and that psychopaths completely lack the capacity to respond to moral reasons (Glannon, 2002, 57). Nevertheless, Glannon claims that psychopaths are, at least partially, morally responsible. Glannon summarizes his argument below:

The argument can be formulated more precisely as follows. If the psychopath has a general understanding of what it is to harm people, and this understanding implies regular receptivity to moral reasons for or against certain actions, then the psychopath is capable of being regularly receptive to moral reasons. The philosophical and clinical analyses I have cited suggest that the psychopath does have this understanding and therefore is capable of being regularly receptive to reasons. Receptivity is a necessary (but not sufficient) condition for complete responsibility for one’s actions. In addition, one must have the capacity to respond to reasons for or against actions. Psychopaths lack this second capacity. But if moral responsibility is a matter of degree, if regular receptivity to moral reasons is enough to make one partly responsible for one’s actions, and if psychopaths have the capacity for regular receptivity, then they can be partly responsible for their actions (Glannon, 2002, 57).
To begin, it is not at all clear that psychopaths are regularly receptive to moral reasons, and it is not clear from the philosophical and clinical analyses that Glannon cites that the psychopath has anything close to the understanding requisite for receptivity to moral reasons. Glannon supports the claim that psychopaths are regularly receptive to moral reasons by citing Murphy’s claim that psychopaths have some sense of what it means to wrong people (Glannon, 2002, 57) and Carl Elliot claiming that even if a psychopath is incapable of understanding what it means to be emotionally hurt, from observing others who have been emotionally hurt “it seems likely that he would be able to form at least some minimal concept of what it means to hurt someone”\(^{51}\) (Glannon, 2002, 57). However, Glannon also cites Hervey Cleckley on the previous page as saying:

The [psychopath] is unfamiliar with the primary facts or data of what might be called personal values and is altogether incapable of understanding such matters…He is, furthermore, lacking in the ability to see that others are moved. It is as though he were color-blind, despite his sharp intelligence, to this aspect of human existence. It cannot be explained to him because there is nothing in his orbit of awareness that can bridge the gap with comparison. He can repeat the words and say glibly that he understands, and there is no way for him to realize that he does not understand\(^{52}\) (Glannon, 2002, 56).

I think that it is enough to say that it is unclear, even given the evidence that Glannon provides, whether psychopaths can be regularly receptive to moral reasons. Even so, there remains the puzzling question of how psychopaths could be even partially morally responsible if they completely lack the capacity to respond to/act for any moral reasons. Fischer and Ravizza are careful not to include an explicit requirement of reactivity to moral reasons (Fischer & Ravizza, 1998, 79) for moral responsibility, though they do require regular receptivity to moral reasons. And while they take it to be an “illuminating” (Fischer & Ravizza, 1998, 79) result of their theory that some psychopaths could be found morally responsible and others not, what they and Glannon refuse to address head-on is just how an agent (or an agent’s mechanism) that is incapable of ever reacting to any moral reasons can be consistent with moral responsibility. Talking about degrees of moral responsibility makes sense when discussing agents with varying

\(^{51}\) This is Glannon citing Carl Elliot’s “Diagnosing Blame: Responsibility and the Psychopath”, *Journal of Medicine and Philosophy* 17 (1992), 199-214, 208.

or diminished capacities, but Glannon fails to explain how moral responsibility can be attributed in cases where an agent *completely* lacks a capacity that has some plausibility as being necessary.

Glannon’s rationale for finding psychopaths (somewhat) morally responsible is not terribly convincing. Glannon points out that Wallace, in *Responsibility and the Moral Sentiments*, argues that recognitional and motivational capacities are jointly necessary for moral responsibility and that it would be unfair to hold psychopaths at all responsible for their actions (Glannon, 2002, 57). To counter this, Glannon points to psychopathic serial killers like Ted Bundy and Gary Mark Gilmore. Glannon reasons thusly

...although they may have been unable to reason about the consequences of their actions or to empathize with their victims, there was enough planning prior to the murders they committed to indicate a sufficient degree of the capacity to recognize reasons against performing their actions (Glannon, 2002, 58).

It seems clear from the above passage that Glannon is equivocating on reasons and is failing to distinguish between prudential and moral reasons. Psychopaths are not, necessarily, irrational. There are a vast range of reasons for psychopathic killers to act with careful premeditation in carrying out their murders, not the least of which are that planning will increase the chances of success and decrease the chances of being apprehended and punished. These aren’t moral reasons, but even Fischer and Ravizza, on whose version of reasons-responsiveness Glannon leans so heavily for his analysis, insist that general receptivity to reasons requires that the agent recognize a range of reasons at least some of which are moral (Fischer & Ravizza, 1998, 77). Glannon fails to prove that psychopaths can be regularly receptive to reasons (in the relevant sense), and so, even if we grant that reactivity to moral reasons is not required, it remains unclear how psychopaths could be morally responsible on Glannon’s view. Glannon distinguishes appropriately between responsibility and criminal liability (Glannon, 2002, 58) but does not seem to recognize that there are different ways for an agent to be responsible. It is important that we not collapse moral responsibility into responsibility more broadly, for moral responsibility is a distinctive sort of responsibility, and maintaining that distinction is important when considering the case of the psychopath and whether such an agent can be morally responsible.53

53 This is, of course, assuming that psychopaths truly lack even the capacity for empathy – a thesis that I find somewhat problematic for reasons that will be discussed in chapter 6.
Basic Social Competence

I will start from the assumption that most of us have a sense of the qualities that separate the socially adept from the socially inept. The most socially adept people tend to be charismatic, attentive, perceptive, insightful, empathetic, articulate, thoughtful, and friendly. The less-than-adept tend to have these qualities in lesser degrees, or not at all. The socially adept are a minority, and moral responsibility does not require that an agent robustly embody all of these qualities. I do believe, however, that it requires, at least, what I call Basic Social Competence (BSC); without BSC, agents lack a type of understanding crucial for acting for moral reasons: how to identify, interpret, and respond to beings worthy of moral consideration.

BSC does not require that the agent possess all the qualities of the socially adept, but it does require, in some small measure, possession of those qualities that are most fundamental: attentiveness and empathy. In this chapter, I will demonstrate why these qualities are essential for ascriptions of moral responsibility and, through the use of illustrative examples, work to demonstrate the degree to which an agent must possess these qualities in order to satisfy BSC.

Assessing the Morally Relevant

If we take morality to consist of the behavior of morally responsible agents with respect to morally relevant (sentient) beings54, which I do, then understanding morality requires an understanding of morally responsible agents as well as morally relevant beings. It also seems quite clear to me that it requires that morally responsible agents have some understanding of morally relevant beings.

Navigating Social Context

So much goes into navigating social context that those of us who are able to do so without much conscious analysis should consider ourselves lucky. There are endless cues for affect, tone, relationships, and mood that must be perceived and integrated into an understanding of the whole. There is social etiquette to be considered and weighed against what you know of the expectations of individuals. There are expectations of performance: posture, expression,

54 I do take this to roughly capture the scope of morality, though I grant that it may be regarded as controversial by some and that “sentient” will be the most likely sticking point. Some will believe that it goes too far, others, not far enough. Many environmentalists, for example, will believe that how we treat the environment (independent of the sentience of any inhabitants of that environment) falls quite firmly within the scope of morality.
words, tone of voice, volume, proximity to others, which must all be appropriately synchronized in order to fully convey meaning.\(^{55}\)

Consider, for a moment, all that goes into reading an expression of sarcasm. Now imagine how much more difficult it would be if you were unable to distinguish between various facial expressions and tones of voice. The words, taken on their own, will mean precisely the opposite of what the speaker intended to convey. When Caroline asks “How do I look?” and Jeff responds “Absolutely hideous!” the ability to pick up on subtleties could mean the difference between thinking Jeff has a fine sense of humor or that Jeff is verbally abusive.

BSC does not require finesse in interpreting and conveying all of the finer points of human communication; miscommunications happen all the time for reasons that have nothing to do with a general deficiency in social competence (fatigue, anger, insensitivity, being in a hurry, carelessness, cultural differences, inebriation, etc.); however, those with a diminished capacity for social interaction face very real challenges not only in understanding and making themselves understood but in identifying the morally relevant and acting morally.

As mentioned in chapter 2, BSC requires a basic understanding of beings worthy of moral consideration, a basic understanding of social expectations and norms, the ability to adapt in the face of new social situations, and some measure of control over emotional impulses. Understanding beings worthy of moral consideration means both an ability to identify sentient life and, on some level, value it. A basic understanding of social expectations and norms has a great deal built into it; included are the ability to understand and make oneself understood with (some kind of) language and a sense of, admittedly culturally variable, things like privacy, respect, personal space, and humor.

The ability to adapt to new social situations, for the purposes of BSC, does not require a highly adaptive personality that can take anything in stride, but does require the ability to cope with breaks in routine that still fall within the scope of normal (read: not extreme) social interactions. The adaptability I have in mind here is of a limited variety and does not require the ability to handle things like extreme trauma with aplomb. It is also worth stating that the need for this adaptability is only at play in unusual (for the agent) circumstances; because we are only concerned with whether an agent was morally responsible in a particular case at a particular time (and not whether an agent is morally responsible \textit{simpliciter}), it is possible for the same agent to

\(^{55}\) For a nice discussion of this, see Peter Railton’s “Normative Guidance.”
be morally responsible for something that happens in the context of their everyday routine but not responsible for something that happens outside of that. This may be the case for some agents with certain kinds of autism, for example.

Emotional Impulse Control

Emotions affect our ability to function within a social context. Though the emphasis in many schools of philosophical thought, from the ancients on down, has been on controlling/overcoming emotions, that is not the point I wish to make here. Emotions are important; they allow us to understand other people and build relationships. Emotions can precede more deliberate reasoning about situations, prodding our moral compass or alerting us to situational problems that have gone unrecognized. Emotional impulses can also override more deliberate reasoning, causing us to hurt other people as well as ourselves, or make bad decisions. Emotions are powerful. The Integrated Mental Status Condition requires that the agent have some level of control over their emotional impulses. Given that the threshold (IMS condition) is meant to capture the requirement for minimal moral responsibility, the bar is set relatively low here. The agent need only have the capacity to reflect on the emotional impulses, at work in the particular case, prior to action, in order to satisfy this part of IMS. I acknowledge that the presence of this capacity will be difficult, if not impossible, to adequately judge in real-world cases, but I will present a case (similar to several real-world cases I have witnessed) in the hopes of clarifying how this capacity remains worthy of consideration.

Audrey is an 11-year-old girl. She does well in school. She has some problems with attention but she works hard and is generally eager to please. She comes from a family that is loving and supportive. Even so, anywhere from 2 to 20 times per week, Audrey will have a complete and total meltdown. Something will happen that she does not like and she will begin repeating herself over and over, rocking her body back and forth, bite her arm, slap her head, lash out physically at people and objects around her, and visibly shake for several minutes up to several hours. Sometimes she has to be swaddled tightly in a blanket and held for an hour or more before she can calm down. This only ever happens at home in the presence of her family. Audrey hasn’t had a public melt-down in more than 5 years.

---

56 Bad decisions can be understood as decisions inconsistent with the agent’s short-term interests, long-term interests, or well-being.
So what can be said about Audrey and emotional impulse control? Clearly, she has a problem with it. However, in certain environments, Audrey seems to exercise more control over her emotional impulses than in others. It is far more plausible to assume that Audrey has the capacity to exert some control over her emotions in public, than to assume that she has never once had something upsetting happen to her in public in the last 5 years. So, let’s say that Audrey has the capacity to exert some control over her emotions in public; can we then say that she has this capacity at home, but fails to utilize it? Not, necessarily. The safer claim is that given Audrey’s capacity to exert control over her emotions in public, it should be possible to cultivate a capacity to exert control over her emotions at home, if appropriate actions are taken. As it stands, satisfying all other parts of IMS, Audrey is a candidate for (some level of) moral responsibility for how she responds to upsetting situations in public settings and may or may not be responsible for how she responds to upsetting situations in the home. She is, however, a better candidate for being responsible for her responses to upsetting situations, at home, in a future where she has been taught various techniques for self-calming and reflecting on her own emotional responses, than would be similarly situated Audrey who had the melt-downs described across all environments.

I won’t say much more here about the ability to control emotional impulses. The level of control that I have in mind is fairly common sense and identifying a precise level of control is not entirely necessary, as emotional control will be most critical to our assessments of moral responsibility in cases where the agent completely lacks it. It is important to keep in mind that, as with other components of IMS, because we are concerned with moral responsibility in context, the same agent may demonstrate the requisite emotional control in some cases but not in others.

**Empathy and Attentiveness**

Underwriting all aspects of BSC are two key capacities: empathy and attentiveness. When either of these capacities is compromised, so too is an agent’s moral responsibility. A deficit in these capacities might be chronic or acute, and the severity of that deficit will determine our attributions of moral responsibility, but if an agent completely lacks either of these capacities, then she cannot be morally responsible. This is not to say, for example, that a generally unobservant person cannot ever be morally responsible. Both of these capacities can
be developed, expanded, and refined. Joe can train himself to pay closer attention, however, should Joe be truly oblivious\textsuperscript{57} to an assault happening in his vicinity, it is not appropriate to say that he is morally responsible, in that case, for failing to call the police or take some other action against it. If, however, an agent lacks even the seeds necessary to begin developing one of these capacities, then such an agent cannot be properly called morally responsible at all.

**Empathy**

There is disagreement in both psychology and philosophy about how empathy should be understood. Oftentimes the meaning of empathy is merely taken for granted. Some psychologists focus on *empathic accuracy* (how accurately an agent can guess at what another agent thinks/feels/believes) while others emphasize the role of empathy as a *purely emotional* phenomenon that arises from our encounters with others (Stueber, 2008, 12). Some philosophers, notably Michael Slote, have emphasized the importance of empathy for ethics, but their attempts to define empathy can range from unclear to problematic, “Unfortunately, it is not entirely clear … just what Slote means by empathy, as he identifies this at different points as (1) the phenomenon of ‘having the feelings of another (involuntarily) aroused in ourselves’ (15), (2) a distinctive feeling that is itself ‘involuntarily aroused in response to situations or experiences that are merely heard of, remembered, or read about’ (17), and (3) an act of ‘deliberately adopt[ing] the point of view of other people [in order to] see and feel things from their perspective’ (17) …” (Smith, 2010).

I will not offer the definitive definition of empathy here. Such a project is well beyond the scope of this chapter. That said, it is nevertheless important to offer a working definition of empathy for the discussion to follow. Tentatively, *empathy is the capacity to understand someone else’s internal state without losing one’s own frame of reference*\textsuperscript{58}. I believe that this adequately captures major schools of thought in both psychology and philosophy while remaining in line with our common sense notions about empathy. This is a good starting place for discussion, but there are various ways to understand ‘understand’ in the definition above, and

\textsuperscript{57} It does not matter why he is oblivious in this case or that he has, at other times, demonstrated a capacity to attend to similar things. He may be morally responsible for other things that may well be related to his inattentiveness, but for not calling the police when he was not aware of a need to call the police, we cannot find him responsible.

\textsuperscript{58} I am indebted to Lacy Orme, a social-worker, for this definition.
there are at least two different ways that are critical for BSC: cognitively and emotionally\(^59\). I propose that, for what follows, we understand empathy thusly:

\[
\begin{align*}
E &= \text{the capacity to understand someone else’s internal state without losing one’s own frame of reference} \\
E_{\text{cog}} &= \text{the capacity to accurately infer, from available cues, the content of another person’s thoughts and feelings}^{60} \text{ without losing one’s own frame of reference} \\
E_{\text{aff}} &= \text{the capacity to experience an appropriate and commensurate affective response to the perceived mental state of another without losing one’s own frame of reference}
\end{align*}
\]

An agent must have both \(E_{\text{cog}}\) and \(E_{\text{aff}}\) if she is to be properly understood as having empathy. Thus, a deficit in empathy can occur in several places: a) the agent could lack the requisite perceptiveness to recognize and interpret cues and clues about the mental state of another, b) the agent could lack the ability to connect emotionally with the perceived mental state of another, or c) the agent could lose her frame of reference and fail to appropriately distinguish between herself and another. Any of these deficits is sufficient to disqualify an agent from satisfying BSC. Such deficits may be temporary, as when counselors or hospice workers experience so-called ‘empathy fatigue’ and find themselves unable to muster an appropriate affective response to tragedy due to the stress of constantly having to empathize with patients, or may be more permanent, as with an autistic person who cannot pick up on social cues, or a psychopath who is incapable of experiencing the appropriate emotional response to the perceived mental state of another. Whether the deficit is temporary or permanent, the effect for IMS remains the same, as

\(^{59}\) This is in keeping with Baron-Cohen, Wheelwright, and others who believe that empathy requires these dual components

\(^{60}\) This is adapted from Ickes’ 1997 definition of empathy where the emphasis is on empathic accuracy.
satisfying IMS depends on the mental state of the agent at the time of the action in question and not the dispositional mental status of the agent in general or an appeal to tracing.

A genuine and total lack of empathy is, on my view, incompatible with moral responsibility. This is not to say, however, that psychopaths cannot be held responsible, full stop. Psychopathy can coincide with high intelligence, excellent reasoning skills, perceptiveness, attentiveness, an ability to understand and act for reasons, and an adequate sense of reality. It is quite possible for a psychopath to satisfy all other aspects of IMS (and to an extremely high degree), but lack of empathy is a disqualifier for moral responsibility even though the high functioning of some psychopaths allows us to hold them responsible for their actions. Indeed, depending on the context of the case, holding the psychopath responsible may be the best thing, not only for society, but for the psychopath.

There is, I believe, an appropriate expectation that morally responsible agents should be capable of some level of remorse when they act immorally. While true psychopaths may regret certain actions, a lack of empathy makes them incapable of remorse in the relevant sense.

It is not difficult to see how this would be the case: Marsha is on the first manned space mission to Mars. When she disembarks she walks around on the surface and begins taking readings and collecting samples. Should scientists eventually come to realize that what appear to be red grains of sand are, in fact, a tiny yet spectacularly intelligent and sensitive alien species, we would not find Marsha responsible for stomping all over them when she could not recognize them as appropriate targets for moral consideration. Upon discovering that she had unintentionally snuffed out countless millions of this newly discovered species, we would expect...

---

61 I am using this term loosely but it should be understood as meaning both acts and omissions of moral relevance.

62 This is not to say that the dispositional mental status of an agent is irrelevant to moral responsibility or other pressing matters in ethics and philosophy more generally, however, assessing what role, if any, an agent’s dispositions will have on moral responsibility is beyond the scope of this project.

63 Importantly, only if we define psychopaths to mean agents completely lacking the capacity for empathy.

64 More on this in Chapter 6.

65 This is not to say that people are not frequently blamed for a lack of remorse (or lack of demonstrated remorse) but that, if an agent is truly incapable of such feelings, it is probably inappropriate to find them responsible for the lack.

66 Again, defined as lacking even the capacity for empathy.
Marsha to feel bad about it, however. If she did not, and in fact was incapable of ever doing so because she lacked the inborn capacity to empathize with anyone at all, then Marsha isn’t morally responsible for crushing the Martians but also cannot be appropriately found morally responsible for anything else. An utter lack of empathy, a complete inability to feel remorse, is a disqualifier when it comes to ascribing moral responsibility. While we may find such agents culpable in other senses, depending on the status of their other mental faculties, they can never appropriately be found morally responsible.

**Attentiveness**

Thus far I have focused on empathy and said little about perceptiveness and attentiveness. The capacity for each is as important for BSC as the capacity for empathy, and the exercise of each is critically important to ascriptions of moral responsibility. In order to clarify how and why this is the case, I will borrow from Lawrence Blum’s “Moral Perception and Particularity” (Blum, 1991). Blum says rightly that, “[a]n agent may reason well in moral situations, uphold the strictest standards of impartiality for testing her maxims and moral principles, and be adept at deliberation. Yet, unless she perceives moral situations as moral situations, and unless she perceives their moral character accurately, her moral principles and skill at deliberation will be for naught and may even lead her astray” (Blum, 1991).

Attentiveness, for the purposes of this discussion, is *the active capacity of the agent to discern, unprompted, morally relevant stimuli from all stimuli available to the agent.* Can the agent adequately parse her environment so that the morally salient features of her environment fall within the scope of her awareness?

**Blum’s Moral Perception**

In “Moral Perception and Particularity,” Blum presents 3 cases in order to tease out various ways that perception is important in our moral assessments. I believe that a brief examination of two of these cases will serve to clarify, somewhat, the roles of empathy and attentiveness in assessments of moral responsibility.
Case 1: John and Joan

In this case, John and Joan are both seated passengers on a subway train. Some of the passengers are standing, as there are no vacant seats, but the train is not uncomfortably full. One of the standing passengers is a woman in her thirties holding two full shopping bags. John is cognizant of the woman but is not paying her any particular attention. Joan, on the other hand, is keenly aware of the woman’s discomfort.

Blum discusses these perceptual differences with respect to salience. What is salient to John, in this case, is merely that the woman in question is holding two bags, while what is most salient for Joan is the woman’s discomfort. It is entirely possible that if Joan were to prompt John somehow, say by asking “Does that woman look uncomfortable to you?” he would realize that she did and that he had been aware of this on some level all along without ever bringing it to the forefront of his consciousness. If this were the case, then “[h]is deficiency is a situational self-absorption or attentional laziness,” (Blum, 1991) and would be the sort of thing that he could actively work on improving.

Even so, if we assume that John was genuinely oblivious to the discomfort of the woman with the shopping bags, then he would fail to satisfy BSC in this case and it would not be appropriate to find him morally responsible for failing to offer up his seat to her. His failure, in this case, would be one of attentiveness and his failure to discern the morally relevant stimuli in the situation. That said, there may well be other grounds for criticism of John. If we assume that John has no limitations that would prohibit him from attending with care to his surroundings, that he is aware of his tendency to zone out and miss morally relevant details in his environment, and that he is not bothered by this tendency and makes no effort to change, then his “situational self-absorption or attentional laziness” (Blum, 1991) may be something that we can find him morally responsible for. It is another step, however, to say that he is morally responsible for failing to offer a seat to a woman when he was completely unaware of her discomfort.

Case 2: Theresa and Julio

Theresa is the administrator of a department and one of her subordinates, Julio, comes to her about accommodations for a disability that causes severe pain in his leg. While Julio is very clear with Theresa about the nature of his disability and the type of accommodations that he needs (and is legally entitled to) and Theresa understands, on an intellectual level, what is
required, she is in no way moved by Julio’s pain. For some deep-seated psychological reason, Theresa firmly associates feeling pain with weakness and feels contempt for Julio, rather than sympathy. The persistent feeling that Julio is being overly dramatic and needs to pull himself together leads Theresa to consistently offer less than is required to accommodate Julio’s disability. The deficiency in Theresa’s response to the situation is her failure to be in touch with part of the moral reality she is confronted with bound up with an additional character deficiency (Blum, 1991, 705). Perception is not limited in this situation to identifying the morally distinct elements at play, it is also a matter of Julio’s disability “weighing adequately within one’s (Theresa’s) view of the situation” (Blum, 1991, 705).

Whereas John’s primary problem seems to be one of attentiveness, the source of Theresa’s problematic behavior seems to stem from a deficit of E_{aff}. The degree of her moral responsibility in this case would depend on the degree to which her capacity for E_{aff} is diminished. If it is completely diminished, then so too is Theresa’s moral responsibility. That said, moral responsibility is not the only consideration in this case. It seems clear that Theresa is fully legally responsible, as Julio’s superior, for failing to offer him adequate accommodation in the workplace. Additionally, assuming that Theresa is not a psychopath and incapable of appropriately empathizing with Julio in this case due to personal deep-seated psychological reasons, she may be morally responsible and criticizable for failing to attempt to work on her psychological issues.

**Considering Deficits**

There are reasons why we don’t find a toddler morally responsible for pulling painfully on the ears of a dog. Most of those reasons stem from the fact that toddlers lack the basic social competence required for moral responsibility. Toddlers, as a general rule, lack the requisite attentiveness and empathy to be found morally responsible. They fail to understand that the dog is capable of feeling pain (E_{cog}) and that their actions are causing pain (E_{cog} + attentiveness); they fail to recognize when the dog is in pain (E_{cog} + attentiveness), and crucially, they cannot connect emotionally with that pain in such a way that they come to understand it as pain and recognize it as harmful (E_{aff}). Lack of empathy and attentiveness (more so than lack of reasoning ability, though that certainly is a factor) are the best candidates for finding that young children are not morally responsible in most cases. The egocentrism of young children is well documented in the psychological literature, and an inability to grasp points of view beyond one’s own is a failure to
meet BSC. The age of reason in the Catholic tradition, the age at which children begin to be morally responsible for their actions, is seven years old. It would be equally appropriate to call this the age of empathy, however, for this is the approximate age at which most children are able to begin to understand things from another person’s point of view. Attentiveness is harder to pin down; while many children may intermittently satisfy BSC from a fairly early age, consistent attentiveness is generally not achieved in very young children and when (if ever) it is achieved seems to be highly variable across individuals.

There are many reasons why an individual may fail to satisfy the Basic Social Competence component of the Integrated Mental Status Condition or satisfy BSC only in limited contexts. Many of the most common reasons will influence other aspects of the Integrated Mental Status Condition as well. For this reason, I will take up a discussion and treatment of autism, psychopathy, and mood disorders (bipolar disorder, major depressive disorder) in Chapter 6, after laying out the final component of the Integrated Mental Status Condition, access to reality, in Chapter 5.
CHAPTER FIVE
REALITY AND PERCEPTION

If I am to say anything at all about accessing reality, I am going to have to start with some assumptions about reality that some may view as contentious. Even so, I believe that most of what I say will seem common sense to most people, or at least be consistent with how most people live their lives.

Assumptions

There is a Fact of the Matter

I am assuming the existence of an objective reality, whether such reality is (fully) perceptible to human beings or not. I accept that humans, as agents, as subjects, are (generally speaking) limited to subjective experience, but I do assume that they are subjectively experiencing the same objective reality. Rather than defend this view, I am going to take it as a primitive; it is foundational, not only for my argument, but for the way that human beings go about their daily lives.

Filters and Lenses

Despite the (assumed) existence of an objective reality, there can be little question that individuals experience this reality very differently from one another. This is due to a combination of factors including, but not limited to: human psychological tendencies (sharpening and leveling, availability heuristic, etc), education and upbringing, culture, personal history, emotional state, physical disability, and mental ability/disability/disorder.

While I am talking about figurative filters and lenses, I think that the most helpful and accessible way to begin this discussion is with an extended analogy about the role of actual filters and lenses with respect to visual perception. For this, I will utilize figures 5.1-5.16 below.
Figure 5.1 is an un-doctored photograph that I took of the Gulf of Corinth with a standard point-and-shoot digital camera. For the purpose of this analogy, Figure 5.1 represents the objective reality of that time and place. In figure 5.2, the colors have been boosted; it is (arguably) the prettier picture and actually comes closer to matching the way I remember that view in my mind: bright blue sky, deep blue water, and lush, vibrant, vegetation. When I took in this view, I was happy, relaxed, and enjoying the vacation of a lifetime in Greece; I remember everything from that time as bright and beautiful and readily admit that Figure 5.1 may be the more accurate representation of that view on that day.

Figures 5.2-5.3 have been manipulated in various ways that have altered the level of color or detail but the view itself remains intact. Figures 5.4-5.5 are manipulated further with a loss of certain details and an exaggeration of others, but the major elements of the landscape remain visible.

Figure 5.6 begins to look like a somewhat different landscape with details of the mid-ground lost and elements in the background (that do not exist in Figure 5.1) gained.

Figure 5.7 is an exaggeratedly Technicolor impression of Figure 5.1 and represents the beginning of a layered filter progression culminating in Figure 5.12, an image that cannot be readily identified as anything, much less a landscape.

**Some Filters and Lenses (Can) Enhance While Others (Can) Obscure**

Notice how much more visible the small terracotta-roofed building on the hillside is in Figure 5.2 as opposed to Figure 5.1. While Figure 5.2 deviates from the original in noticeable ways, it does so in such a way that details are more easily discerned and colors more easily recognized. Just so, not all of the metaphorical lenses and filters that influence our perceptions of reality serve to obscure; in certain contexts, they can be viewed as an asset. Someone who approaches others from a default standpoint of caring and patience is much more likely to pick up indicators about the feelings of others, just as an individual with medical training is much more likely to pick up on early symptoms of a serious illness than someone without such a background.

As is made clear by the photo progression above, however, some filters and lenses shadow, obscure, and even obliterate our perceptions of reality. The deviation, at a certain point, can become so great that it is impossible to even guess at the original. While Figures 5.1-5.2 depict a pretty, and even tranquil, landscape, Figures 5.4-5.6 evoke a more surreal or forbidding
sense despite the many features held in common. It is easy to see how being immersed in the landscape of 5.2 could inspire peace and contentment and how immersion in 5.5 could inspire anxiety and even fear. In this analogy, acquired biases like racism, sexism, and even patriotism, or emotional/personality/psychological disorders (like social anxiety, obsessive-compulsive disorder, bi-polar disorder) can color perceptions of reality in a similar way.

Our circumstances and environment also play a significant role in our perceptions of reality. There are many recent psychological studies that point to significant differences in perception between social classes. It has been found that those people who come from a lower socio-economic background are, in general, more attuned to their social environment, more cognizant of the emotional status and personal circumstances of the people around them (Stellar Manzo, Kraus, & Keltner, 2012). It is posited that this is a sort of vigilance required by a more threatening environment. A major coping mechanism for this, it is theorized, is that people of lower socio-economic class “may manage external threats by initiating cooperative relationships, which promote greater attentiveness, awareness, and concern for others” (Stellar et al., 2012). Stellar et al. (2012) conducted 3 separate studies to test their hypothesis that people of lower socioeconomic status would exhibit greater compassion for others than their higher class counterparts.

Study 1 tested the standing disposition of participants to experience compassion. Participants reported key demographic information, information relevant to determining socio-economic status, and a measure employing the Dispositional Positive Emotion Scale (DPES). The study did find that lower-class participants had a greater disposition toward compassion than higher class ones. Additionally, “…controlling for gender, ethnicity, and spirituality, social class did not predict scores for dispositional joy, contentment, pride, love, amusement, and awe …[t]hese latter findings suggest that social class uniquely predicts compassion, and that this association is not due to a class-related tendency to feel higher levels of all positive or approach-related emotions” (Stellar et al., 2014, 452).

67 See: (Gallo, 2005), (Kraus, 2009), (Kraus, 2010), (Kraus, 2011)

68 The Dispositional Positive Emotion Scale (DPES) is a well-validated self-report measure. It measures the trait-like tendency to feel positive emotions including joy, contentment, pride, love, compassion, amusement, and awe on an enduring, trait-level basis (Stellar et al., 2012, 451).
Study 2 sought to discover whether the class differences with respect to compassion went beyond self-report. Heart rate deceleration is a physiological response that has been associated with greater external focus and orientation toward others (Stellar et al., 2014, 452), so, in addition to gathering self-reported data, researchers monitored the heart rates of participants in response to two videos (one neutral, one of suffering). Again, lower class subjects showed greater compassion across the board, when asked about both the neutral and the suffering video, level of compassion induced by the suffering video, and the increase in compassion from neutral to suffering video. This was the case for both self-report and physiological monitoring (Stellar et al., 2014, 453-454).

In Study 2, social class predicted participants’ compassion related subjective reports and autonomic responses associated with orienting toward and engaging with others. Although lower-class individuals typically show greater threat-related increases in heart rate and blood pressure than upper-class individuals in response to ambiguously negative stimuli (e.g., Chen & Matthews, 2001; Chen et al., 2004), in Study 2, lower class individuals showed greater heart rate deceleration in response to the suffering of others. This heart rate deceleration, which has been associated with a greater outward focus, engagement, and feelings of sympathy (Caccioppo & Sandman, 1978; Eisenberg et al., 1989), may be part of a bodily calming response that prepares individuals to engage with another person in a more affiliative and caretaking fashion. Overall, the results from Study 2 are in keeping with the complex emotions theorized to be part of a tend-and-befriend response to threat, and reveal that lower-class individuals appear to exhibit quite different autonomic responses depending on whether the threat is to the self or another person (Stellar et al., 2014, 454).

Study 3 tested class differences in compassion within the context of a competitive mock interview. Researchers were interested in seeing whether lower class participants would exhibit greater compassion for their interview partners. There is some evidence to suggest that lower class individuals are better able to read distress in others, through recognizing and interpreting various social cues (Stellar et al., 2014, 455), so researchers were also interested in how well participants would be able to gauge the emotional state of their interview partners. They found that “overall, the results suggest that, due to their orientation to the social context and their perceptions of others’ distress, lower-class individuals report experiencing more compassion compared to their less contextually oriented upper-class counterparts” (Stellar et al., 2014, 456).
It is postulated that lower class individuals must be more attentive to their environment, because their environment is often more hostile and dangerous than that of their upper class counterparts and that this greater outward focus makes them more finely attuned and ready to react not only to perceived threats, but to suffering observed in others. If this is the case, then it serves as a prime example of the sort of lens that I have in mind. If we take the results of these studies to heart, then a lower socioeconomic class can serve as an enhancing lens with respect to perceiving morally significant features of the environment and that, conversely, a privileged background may blind someone to these same features. A significant upshot here is that, if this is the case, it must be by way of a cultivated disposition (rather than innate differences between agents); this lends support to the notion that dispositions of this sort can, in fact, be cultivated through the manipulation of environmental factors.

A Certain Level of Obscurity Constitutes a Break with Reality

Figures like 5.11-5.12 are so far removed from figure 5.1 that it is fair to say that the view is more than merely obscured. Words like “dandelion,” “bush,” or even “smudge” would be much more likely to come up as descriptors than “view of a body of water” or “landscape,” if you were to put these latter images in front of someone. This level of departure from the original constitutes a break with reality. There can be no expectation that a person viewing figure 5.12 can describe the color of the water, the shape of the vegetation, or the location of the small terracotta-roofed dwelling on the hillside. Carrying this analogy through, perceptual distortion of this magnitude can be brought on by: certain drugs, severe hallucinations and delusions, extreme trauma, mental disorders like schizophrenia, or malfunctioning of some sensory or processing apparatus. An agent in the midst of a break with reality cannot be found morally responsible for actions directly informed by such perceptions.

---

69 Perhaps, by way of cultivating some psychological trait or disposition.
70 This is very much in line with the arguments that I will make about holding responsible and Applied Behavior Analysis.
Breaking with Reality

Below, I believe, is a fairly comprehensive list of categories that will be helpful in discussing the moral significance of a break with reality. For our purposes, a break with reality may be caused by:

1. Physical Impairment: permanent or temporary
2. Cognitive Impairment: permanent or temporary
3. Emotional Impairment: permanent or temporary
4. A combination of such impairments (ex. from trauma)

I take it to be relatively uncontroversial that these five categories provide for what we might call extenuating circumstances where moral assessment is concerned, and that, if it is severe enough (i.e. results in a break with reality), most people will allow that permanent physical, cognitive, and emotional impairment have exculpating power where moral responsibility is concerned and that certain cases within traumatic contexts might also make ascriptions of moral responsibility inappropriate.

The controversial sub-set of cases that fall within these categories are the temporary impairments that seem to result from a straightforward causal chain that begins with some un-coerced, un-impaired, intentional action on the part of the agent. The conviction that such agents are responsible, even for actions that occur while they are in the midst of a break with reality, motivates both our everyday assessments of praise and blame, and backward-tracing accounts of moral responsibility in the philosophical literature on the subject. We blame people, for example, for actions undertaken while under the influence of psychotropic drugs, even if those actions were a direct response to severe hallucinations. Walter Glannon argues that a person suffering from schizophrenia, I will call her Mary, can be held morally responsible for her actions, even when those actions were in direct response to a schizophrenic break with reality, if, at some earlier time, knowing the possible consequences, Mary refused to take her medication (Glannon, 2002, 49).

---

71 This is not to say that it is necessarily exhaustive.
72 Permanent, in this context, should be taken to mean on-going, long-standing, or chronic and not as implicating some necessity of that state or the impossibility of ever emerging from that state.
I am highly skeptical of backward tracing accounts of moral responsibility. If an agent is motivated at time $t$ by a perception of reality at time $t$ that is so radically different from the objective reality as to constitute a break with reality, I don’t believe that an agent can be found morally responsible for their actions at time $t$. It may well be appropriate to find and hold the agent fully responsible for acts and omissions that occurred prior to her break with reality, to weigh the significance of those actions in light of the well-known associated risks and, perhaps, to subject her to consequences as if she were responsible. That said, it does not make sense to insist that an agent can be morally responsible at a time when she lacks appropriate access to reality and her actions are (mis)informed by that impairment.

I acknowledge that this is one of the more controversial claims that I will make, so I will work to clarify why I take this to be so with a return to Glannon’s case of the schizophrenic who chooses to go off meds.

**Schizophrenic Mary**

Let us begin with the schizophrenic woman that I have chosen to call Mary. Mary is a diagnosed schizophrenic, and at some time (say $t_x$) Mary forgot to take her psych meds. A few days later (say $t_y$), feeling more clear-headed and hopeful than she had in a long time, Mary made the decision to stay off of her psych meds. Assume that Mary was not operating under any misperceptions about her own state of mind or the possible consequences of going off her meds. She recognized the possibility of a psychotic break, but was also hopeful that she could avoid that outcome (she had read some promising things about the benefit of a healthy diet and antioxidant vitamins in treating schizophrenia).

We will imagine several possible worlds where Mary makes the very same choice to go off of her meds. In the first world (W1), in the days following $t_y$, Mary begins to feel more in tune with the world around her, more cognizant of her surroundings and the suffering of the people around her. She decides to start volunteering at a local soup kitchen and, at time $t_z$, she is volunteering. In the second world (W2) Mary feels like she comes to a deeper understanding of herself and this leads her to start an inspirational blog chronicling her struggle with schizophrenia on which, at time $t_z$, she is setting up a way for her readers to donate resources to various organizations that assist the mentally ill. In the third world (W3), Mary becomes
increasingly paranoid until she is utterly convinced that there is a conspiracy bent on her demise. She begins to see threats around every corner and becomes desperately afraid for her life. When Mary’s new neighbor, Simon, recognizes her in a parking garage at the mall and starts moving to intercept her, Mary is convinced that her enemies are closing in and that her life depends on getting to the bottom of the conspiracy. At time $t_z$, she goes on the offensive, punching Simon and demanding to know who he is working for.

My intuition is that most people will argue that Mary is morally responsible for her philanthropic actions in W1 and W2. Further, I think that most people will want to say that Mary is morally responsible, not because she made the decision to go off her meds at $t_y$, but because, at time $t_z$, she (“of sound mind” i.e. satisfying IMS) made the choice to volunteer at the soup kitchen, or help others with mental illnesses. I believe that most people will go one of two ways in assessing Mary’s moral responsibility in W3, either they will say that she is not morally responsible for her actions at time $t_z$, due to her state of mind at time $t_z$, or they will argue that she is morally responsible for her actions at time $t_z$, because she made a choice at time $t_y$ that resulted in the delusion that caused her actions at time $t_z$.

Those who want to hold Mary responsible in W3 at time $t_z$, want to do so because at time $t_y$, Mary chose to stop taking her psych meds knowing that this increased her chances of having a psychotic episode. Importantly though, Mary’s psychotic break was not a foregone conclusion, but one of many possible outcomes. Yet, of the three possibilities presented in this case, it is the only outcome where we attempt to locate the source of her moral responsibility at time ($t_y$) in a past action and, perhaps, the only scenario (of the 3) where we would consider her actions at $t_y$ (to go off meds) to be morally problematic.

The Asymmetry of Backward Tracing Accounts

Backward tracing is a stop-gap measure. Backward tracing accounts of moral responsibility are generally employed as a means of assigning blame when we have strong feelings about a negative consequence and few resources for explaining why someone should be held responsible for it. I have difficulty conceiving of a case wherein we resort to backward tracing as a means of praising someone who we cannot otherwise justify praising for some good consequence. You may be tempted here to think of a case like that presented to me by Randolph Clarke:
Suppose I know that my wife finds my snoring to be the funniest thing in the world, and I see that she needs a good laugh. I make myself fall asleep, knowing that I’ll snore and then she’ll be amused. Would this be a case of tracing praiseworthiness for an outcome to an earlier praiseworthy action? In fact, isn’t every case of praiseworthiness for a non-action outcome a case of tracing back to a praiseworthy action? Clarke is right to point out that we do trace praiseworthy non-action outcomes, back in time, to earlier praiseworthy actions. Clarke is also very careful in his wording when he specifies “non-action outcomes.” Some clarification of my view is clearly in order. While I am committed to a current time-slice assessment of an agent’s mental status for determining moral responsibility at a particular time (which is to say, looking at the agent’s mental status at the time of the relevant action/omission), I am not committed to an analysis of action divorced from time. Chronology is important; consequences will always be temporally separated from the actions that cause them and the length of this temporal separation is widely variable. Tracing non-action outcomes to a previous action, be it praiseworthy, blameworthy, or morally neutral is unproblematic on my account.

The problematic tracing cases, as I see it, are those that seek to trace responsibility for not only non-action outcomes, but further actions plus the outcomes of those further actions, to some earlier responsible action. The problem is that the outcomes of actions, the responsibility for which is ascribed based on the fact that the responsibility for some antecedent action can be traced to some prior responsible action relies on a chain that distances those outcomes not only temporally, but probabilistically with respect to likely results and thus the predictive powers of the agent at the time of the original responsible action.

This fundamental asymmetry, that we look to backward tracing accounts to do so much work for blame and so very little for praise should, at the very least, raise some eyebrows. Not all asymmetry is bad, but when it is applied to explain our morally reactive attitudes (that are ultimately grounded in a very personal and extremely complicated mix of emotion, experience, reason, and intuition) we should have a care. If we are justified in treating moral blame as fundamentally different from moral praise for the purposes of certain cases, then we should be able to offer reasons for this. I’m not saying that there aren’t any, but I am saying that I can’t think of any.

---

73 Thanks to Randy Clarke for this editorial comment from May 31, 2013.
Appealing instead to IMS to determine moral responsibility offers another way to consider such cases that doesn’t require this sort of asymmetry and still accounts for most of our intuitions about appropriate morally reactive attitudes. If it is the case that I can do away with backward tracing accounts of moral responsibility and this potentially problematic asymmetry, while allowing for a reasoned analysis of our moral intuitions, I take that to be an advantage of my theory over competing theories that rely on backward tracing to do the heavy lifting when it comes to explaining moral responsibility in tough cases.

**Attributability in Tough Cases**

Let us now return, briefly, to the case of Mary. I discussed earlier how her responsibility for attacking her neighbor might be explained on a volitional account of moral responsibility, through the use of backward tracing. I also, briefly, offered an objection to backward tracing as an effective strategy for assessing moral responsibility. Now, I would like to briefly revisit Mary’s case and consider how our intuitions might be better explained by employing IMS.

If you will recall, Mary is a schizophrenic who decided to go off of her medications. She did so with a plan to try to manage her condition through other means (diet/supplement therapy). In W3, Mary, at time \( t_z \), attacks her neighbor, Simon, having mistaken him for a player in a conspiracy to kill her. According to IMS, Mary is not responsible for attacking Simon, because, at the time in question, her perception of reality was too distorted for her to satisfy IMS. Mary would similarly be found not responsible for the attack on Smith’s rational relations view, because her action was not reflective of her own evaluative commitments, but rather, the result of her altered mental state. Even though Mary would be found not responsible for attacking Simon at time \( t_z \), there are other actions, attitudes, and judgments of Mary’s that may be appropriately subject to our moral evaluations.

This particular case leaves room for us to go various ways in our assessment of Mary. Schizophrenia is unpredictable, and the extent to which various psych meds control symptoms is highly variable from person to person; furthermore, the side effects of such medications, especially the long term side-effects, can be highly damaging. This should be kept in mind, lest we think of the case simply as Mary making the choice of whether or not to have a schizophrenic episode. That said, Mary was aware that she was more likely to have an episode off of her meds and so her reasons (the relationship between her reasoned judgments and the attitudes and evaluative commitments that are in some part constitutive of who she is as a person) for making
that choice matter. If she believed (and didn’t merely hope without thinking there was much of a chance) that she could do well off of the meds with alternative treatment plans, and that being off of her meds could be an important step in becoming a better person, then we are likely to view (and judge) her decision to stop taking her medication far differently than if she believed that she would have an episode, knew that she would likely be a danger to others when that happened, and decided that a few days of clear-headedness were worth more than the safety of the people around her. Though, importantly, how to judge Mary for the decision to go off her meds is separate and distinct from the question of how we should judge her for attacking Simon.

In Conclusion

Rather than taking on perception, in general, I have dedicated most of this chapter to taking on what I take to be one of the most controversial sorts of cases involving perception and ascriptions of moral responsibility, namely those involving temporary impairment of the agent that, one could argue, results from a straightforward causal chain that begins with some un-coerced, un-impaired, intentional action on the part of the agent. I have worked to show that not only are these causal chains less than straightforward, but that relying on such chains with backward tracing accounts of moral responsibility is an ad hoc method of ascribing responsibility and that my own condition, IMS offers a plausible alternative.

With all of the major components of my Integrated Mental Status condition laid out, I will turn in Chapter 6 concluding remarks about the most important aspects of my view and how it might be applied in various sorts of cases.
CHAPTER SIX

FINDING RESPONSIBLE, HOLDING RESPONSIBLE, & APPLYING IMS

This is my concluding chapter, and much of it will be devoted to the application of the Integrated Mental Status Condition, as a tool for analyzing moral responsibility in various types of cases. Before moving on to particular cases, however, I will briefly recap the takeaway points of this dissertation. This section should do double duty as both conclusion and consolidated introduction to the framework that will underwrite the analysis to follow.

The Take-Away Points

This is a theory about moral responsibility.

In the Integrated Mental Status condition, I outline the aspects that I take to be most relevant to answering the question “What is required for an agent to be morally responsible?” and those relevant aspects have to do with an agent’s mental status. I am however, positing that bit of theory within a set of constraints about moral responsibility that I take to be equally important.

Responsibility is not an all or nothing affair.

Determinations of moral responsibility aren’t binary. Responsibility exists on a spectrum, or continuum. The Integrated Mental Status Condition is an attempt to isolate and talk usefully about that part of the spectrum where there is a shift from not responsible to somewhat responsible, but I am not advocating a hard line approach. There are infinitely many degrees of responsibility that could be targeted for discussion. I do believe that my Integrated Mental Status Condition tracks something important and that the greater the degree to which it is satisfied, the more morally responsible an agent can be said to be.

Moral responsibility is deeply contextual…

…and so ascriptions of moral responsibility must be indexed. I appeal to a current-time-slice notion of moral responsibility; the claim is that we cannot sensibly ask questions like “Is
Suzy a morally responsible agent?” only questions like “Was Suzy morally responsible in this particular case?”

**A moral philosophy without the conceptual resources to talk about real people is missing the point.**

There is a danger in philosophy of reaching an untenable level of abstraction. When ethics and moral philosophy stray too far from questions about how people should actually live, it can become little more than an academic exercise. I am aiming for something more than that with this project. This is why I have appealed to other disciplines that put human beings at the center of their inquiry and practice. It is also why I have chosen to eschew certain things, like agent mechanisms, that I take to be philosophical abstractions, in my own theory. While I would like very much to think that my theory could apply to, as yet unknown, advanced alien races, my goal here is simply to talk about real people, the people that we are, the people that we know.

**The Integrated Mental Status Condition is an integrated mental status condition.**

I’ve spent this entire dissertation breaking down my condition into its component parts in order to better explain the content of the condition. While I have not yet thought of a better way to offer a comprehensive treatment of the condition, such an approach can have misleading connotations. The reason the condition is called the Integrated Mental Status Condition is because satisfying the condition requires the successful integration of all of these component parts within the mental life of the agent. While I believe the components capture what I want to capture, the divisions that I have drawn in an attempt to explain them, are somewhat artificial. Cohesion and interdependence is essential for the condition to be satisfied and when asking whether an agent is responsible for a particular action at a particular time, it makes sense to ask, “Were the relevant aspects of their mental status integrated at the time?” The relevant aspects dissected and considered as the ability to act for reasons, basic social competence, and access to reality should be read as guidelines for the sake of clarity rather than a hard and fast checklist.

**Sometimes we should hold people responsible, even when they are not.**

This is a big one, and it runs counter to many current conventions so I will spend a bit of time explaining this. I have laid out what I take to be a necessary condition for moral
responsibility in the sense of finding someone morally responsible; that is, making a
determination that they were, in fact, morally responsible in a particular case. It is often taken as
a further step to then, hold someone responsible for their actions. Even so, the contingency
standardly runs: If found responsible, then hold responsible (or choose not to). Indeed, much of
the analysis of moral responsibility has been done with the express purpose of determining when
it is appropriate to hold someone responsible.

I will now argue that holding someone responsible for their actions need not be
contingent on finding that person morally responsible for those actions (only on finding them
technically responsible) and that, in some cases, it may be the best course of action to hold
responsible in the absence of moral responsibility. Teasing apart the relationship between
finding and holding responsible, will allow for a more nuanced reading of many ethical tough
cases.

Before moving forward, I wish to make it clear, that I am not conflating notions of
responsibility when I suggest that we hold someone responsible even in the absence of finding
them responsible. I am not merely suggesting, for example, that we can hold someone
responsible in a legal sense, even if we do not find them morally responsible. Nor, that we can
hold someone responsible in moral situations in the straight-forward cause and effect fashion that
we might say someone is responsible in situations without moral implications (ex. Joseph is
responsible for turning on the light switch). What I am suggesting is something very much like
what Hanna Pickard calls responsibility without blame, or “detached blame” (Pickard, 2011).

**Responsibility Without Blame**

In her 2011 article, “Responsibility Without Blame: Empathy and the Effective
Treatment of Personality Disorder,” Hanna Pickard distinguishes between what she calls
“detached blame” and “affective blame” and argues that while “detached blame,” is essential for
the successful clinical treatment of personality disorders, “affective blame” can be highly
detrimental to treatment (Pickard, 2011). Effective clinical treatment, according to Pickard, calls
for a conceptual framework that not only delineates detached and affective varieties of blame,
but appropriately separates the concepts of responsibility, blameworthiness, and blame that are
often inextricably linked in the philosophical literature.
Accepting responsibility, Pickard argues, and I agree, is essential for treating personality disorders. Not only do individuals need to accept responsibility for their behavior, clinicians need to accept that they are capable of being responsible.

For one cannot rationally resolve to change that which one believes one is powerless to change (Pearce and Pickard 2010; Pickard and Pearce forthcoming). Effective treatment for PD [personality disorder] depends on clinician and service user believing that the service user has choice and at least a degree of control over their behavior: they are to that degree responsible agents. (Pickard, 2011, 213-214)

Pickard believes that this is possible if clinicians adopt a stance of detached blaming and eschewing a Strawsonian approach that would reduce morality to morally reactive attitudes. We can judge that an agent is responsible for her behavior, that she had “conscious knowledge, choice, and a degree of control over [her] behavior,” (Pickard, 2011, 2015) but in a way that leaves out the “characteristic ‘sting,’” so closely associated with accounts of blame that can be so clinically counter-productive (Pickard, 2011, 2016).

Pickard’s article is chiefly concerned with individuals suffering from personality disorders, but the concept of detached blame, that she presents, has broader implications. It is this sort of blame that can be instructive and constructive. It is a way of acknowledging mitigating factors, without obliterating agency and, in so doing, offers a way forward. It is my position that the various aspects of the Integrated Mental Status condition, that I have proposed, can be developed over time, actively doing so, however, requires a presupposition of agency and the ability to take on a conceptual framework, like Pickard’s, where excessive affect can be left out of the equation. Relying too heavily on our emotions to parse the moral landscape, can have pernicious effects “[b]lame is like fear. It can fly in the face of considered judgments about what is true of the blamed object and what should be true of the blaming object (Pickard, 2011, 217).

Applying the Integrated Mental Status Condition

In what remains of this chapter, I will discuss the Integrated Mental Status condition within the context of cases/populations that have presented various sorts of challenges for ascribing moral responsibility on traditional accounts. I will begin with and elaborate most on those with autism. It is the population that I have the most first-hand experience with, and also the one that I believe would be voted “most likely to be excused,” of the cases I will present. I
will then address the case of the classic psychopath, and elaborate on my skepticism of true psychopathy hinted at in Chapter 4. I will briefly revisit the case of Mary, also from Chapter 4. Finally, I will discuss cases of voluntarily altered consciousness through the use of various substances (drugs, alcohol) ending with the class of cases that I would expect would be voted “least likely to be excused.”

**Autism**

Autism is a neurodevelopmental disorder characterized by “severe and pervasive impairment in reciprocal socialization, qualitative impairment in communication, and repetitive or unusual behavior” (Levy, Mandell, & Schultz, 2009, 1627). In addition to these core impairments, comorbid disorders, including intellectual delays, inattention or symptoms consistent with attention deficit-hyperactivity disorder, externalizing behaviors like aggression, affective difficulties, sleep disruption, and sensory differences are common in children with autism (Levy, Mandell, & Schultz, 2009, p.1629). Autism spectrum disorders have a wide degree of variability and include autism, Asperger’s syndrome74 and PDD-NOS (pervasive developmental disorder not otherwise specified). There is such a range within the autism spectrum that it can be problematic to generalize, but Asperger’s is typically less severe than autism and, of the three subgroups, PDD-NOS encompasses the greatest degree of variability.

Over the last two years, I have been involved in the direct one-on-one implementation of ABA therapy with autistic children. The consistent presentation of consequences (holding responsible) for actions, even, and perhaps, especially in the absence of a pre-existing understanding of consequences of an action is critical for building that understanding. Contacting natural consequences is the way in which animals learn. Considering the situation of autistic children, who either do not intuitively take away all the lessons that natural contingencies have to offer or are denied contact to natural contingencies, for one reason or another, illuminates a sort of break-down of what we might consider ‘natural learning processes.’

---

74Asperger’s has been removed from the most recent DSM, however, the designation remains in use, even in clinical settings.
ABA techniques are effective for teaching and altering behavior across populations but it has gained popularity over the past decade, largely due to its effectiveness in teaching individuals on the autism spectrum. One reason why ABA is so effective is that virtually everyone satisfies what Fischer and Ravizza would describe as Minimal Reasons-Responsiveness. There is almost always at least one thing that has the power to alter behavior (technically future-behavior; not everyone is rule-governed to the extent that they could reason things out if merely told about different consequences). In ABA, these things are called reinforcers and punishers. A reinforcer is something (some stimulus) that, when presented as a contingent consequence, increases or maintains a particular behavior, over time. A punisher is some stimulus that, when presented as a contingent consequence, decreases or eliminates the frequency of a behavior over time. Importantly, you cannot make a list of things that are reinforcers and things that are punishers. Reinforcers and punishers can only be determined for a particular agent at a particular time because they are different from person to person and will change for any particular person over time (perhaps in the course of a moment). What functions as a reinforcer for one person, will serve as a punisher for some people and a neutral stimulus for others.

In the case of autistic children, reinforcers, (those consequences that will increase the frequency of a behavior over time) might be incredibly idiosyncratic. Idiosyncratic and restricted to the point that she or he may be only very weakly reactive to reasons and the reason in question might lead us into something like the “Crazy Problem” introduced in Chapter 2, or “Grumpy Mutt,” from Chapter 3. The important point that I want to make here, is that, even if this is where we start, and we grant that IMS is not satisfied and so, the agent is not morally responsible at this time, that isn’t/needn’t be the end of the story.

Colby

Let’s take the case of Colby, a nine-year-old boy with autism. Colby, prefers to be alone and avoids the company of his peers. His interests tend to be highly restricted, and at the current time, his only interest seems to be the Hercules beetle, a variety of rhinoceros beetle found in South America. Not much else seems to matter to Colby and attempts to get him to engage with his peers, result in aggressive outbursts. You can explain to Colby that hitting and kicking hurts

---

75 Indeed, there is a whole branch of ABA, OBM (Organizational Behavior Management) that capitalizes on this when using ABA techniques in business settings.
the other people, that his outbursts are upsetting to his teachers and parents, that other children will not like him, that he could be accidentally hurt when restrained, that he will not be allowed to attend school in a regular classroom and, even understanding and accepting these things, he is un-moved by them. Which is to say, none of these reasons motivate Colby to change his behavior.

Now, suppose that Colby is offered the opportunity to read a bright glossy book on the Hercules beetle, if he can sit in a small group of his peers without hitting or kicking for a specified amount of time and suppose that he sits with his hands and feet to himself for the specified period and earns the book. This reason, that access to a desired book was contingent on sitting appropriately, had the motivational force to change Colby’s behavior in a way that none of the other reasons did.

At this stage, and in this case, Colby is not going to satisfy IMS for various reasons. Though he has proven weakly reactive to reasons, he does not have the requisite ability to act for reasons, his basic social competence is lacking, and he may also be dealing with difficult perceptual barriers. That said, we needn’t assume that Colby can never satisfy IMS, due to his autism.

Colby’s weak reactivity to reasons confirms that he has at least one reinforcer. The great thing about reinforcers, is that it is possible to make more. Conditioned reinforcers can be created using pairing and simple operant conditioning. As verbal praise is repeatedly paired with delivery of the desired book, verbal praise can become a conditioned reinforcer and once social reinforcement is on the table, the door is opened to developing those areas of social competence where Colby is deficient. As aspects of the natural environment are paired with reinforcement and the number of Colby’s reinforcers grows, Colby will become more than weakly reactive to reasons and, if the right things are paired in shaping his behavior, he will become increasingly susceptible to the reasons listed earlier. As his book is paired with proximity to peers (and, as time goes on, peer interaction) he will find the company of his peers less aversive and more reinforcing on its own merits. If we treat Colby as an agent, and allow him to contact consequences, encourage him to take responsibility for his own actions (even when we don’t find him morally responsible), and blame him for his missteps, in Pickard’s sense of detached blame, then there is a possibility that, at some point in the future, Colby will satisfy IMS in certain
cases. If we don’t treat Colby as an agent, and treat him as a diagnosis, instead, the chances of him ever satisfying IMS in any case is extremely unlikely.

**Autism and Empathy**

I have often heard that a key aspect of Autism (and one that occasionally gets them lumped in with psychopaths) is that those on the Autism Spectrum lack empathy. In my experience, this is a somewhat misleading statement. Often those on the spectrum have difficulty with communication, self-expression, labeling and identifying emotions, reading social cues, and engaging those around them, but this should not be interpreted as a blanket lack of feeling or caring or, worse, a lack of the capacity for feeling and caring. I’ve come across multiple children with an Autism diagnosis that could be described as overly sensitive or emotional. I’ve seen children, on the spectrum, who initially seemed shut off, become quite social. Where they tend to need the most help is with processing. The learning associated with building up the ability to process things that do not process intuitively, requires many teaching trials and the force behind that teaching is holding responsible even in the absence of finding responsible. Importantly, this process, over time, can have the effect of making them more responsible (potentially to the point of satisfaction of an Integrated Mental Status condition).

**Schizophrenia**

**Mary Revisited**

Let us revisit Mary, the schizophrenic women, from Chapter 4, who decides to stop taking her medication and ends up accosting her neighbor. I have previously argued against backward tracing accounts that would find Mary morally responsible for attacking her neighbor, Simon, because she made the conscious choice to go off of her medications at a prior point in time knowing the potential consequences for doing so. I would now like to consider Mary’s case with respect to the Integrated Mental Status condition and the theoretical framework that I have in mind. Ultimately, I will argue that while we can find and hold Mary responsible for the decision to stop taking her medication, we cannot find Mary responsible for attacking Simon, given her failure to satisfy IMS at the time of the attack, however, we may still hold Mary responsible without blaming her.
In assessing whether Mary is morally responsible, we need to be specific about when and for what we are determining responsibility, and from there, appeal to IMS. If Mary satisfied all parts of IMS, was able to act for reasons, exhibited basic social competence, and had reasonable access to reality, when making the decision to go off of her medication, then we can say that Mary is responsible for that choice. To be clear, in saying that Mary is responsible for making a choice, it is an acknowledgement of her autonomy and ability to make a decision to do something with potentially dangerous and somewhat unpredictable consequences. It is not to say that Mary is responsible for attacking Simon.

At the time that Mary, attacked Simon in the parking garage, she did not satisfy IMS. This is primarily due to the fact that Mary, at the time of the attack, lacked the requisite access to reality to satisfy the third part of IMS. Assuming it was Mary’s genuine perception that she was acting in order to save her own life, we cannot say that Mary was morally responsible for attacking Simon when she did. This is not to say that we cannot still hold Mary responsible in an important sense. It would not be outlandish to expect that, once she was mentally stable, Mary should apologize to Simon. If Simon was seriously injured, Mary should bear witness to that and offer what assistance she can. Though she was not responsible, contacting consequences, has the potential to shape her future choices. When, in the future, she is weighing the risks of going off of her medication, this incident with Simon should weigh considerably into her decision making.

And, if, in the future, while satisfying IMS, Mary makes the decision to go off of her meds, we might be justified in judging that decision far more harshly than her original decision to go off her meds. This would be so, because, being receptive and responsive enough to reasons to satisfy the ability to act for reasons component of IMS, concerned enough about her fellow human beings to satisfy the basic social competence component, and in touch with reality to the extent that she knew what she was doing and satisfied the access to reality component, Mary chose to do something that would very likely be dangerous for those around her, rather than just potentially dangerous.

**Voluntarily Altered Consciousness**

Mary chose to stop taking a medication with a number of unpleasant side-effects knowing that doing so could potentially alter her mental status. I have argued that she is responsible for that choice, but that saying she is responsible for that choice, is distinct from
saying that she is responsible for what happened after that choice. Mary, however, was dealing with schizophrenia and faced with a choice featuring two relatively unappealing options. How should we assess the responsibility of individuals who actively seek to alter their mental state through the use of drugs and alcohol?

The answer to this question, might be disappointing to some. Responsibility should be assessed in the same way. Questions about the moral responsibility of an agent should be narrowed to an act/omission at a particular time and then it must be considered whether the agent satisfied IMS at that time. If the answer is no, then the agent is not morally responsible in that case. It still may be the best course of action to hold them responsible, to the extent that doing so can help shape future behavior, but it would not be accurate to say that they are morally responsible if IMS is not satisfied. That being said, there may be a whole host of acts/omissions that occurred when the agent was sober and satisfying IMS that they are fully responsible for, and these may have far greater moral significance that we often credit.

The Drunk Driver

It is an open question, as far as I am concerned, whether alcohol consumption alters a person’s mental state to the extent that IMS would no longer be satisfied. Certainly, drinking alcohol can impair reasoning, make one less sensitive to certain relevant reasons (and perhaps overly sensitive to others), reduce reaction time, affect mood, lower inhibitions, and dull perception, but even so, consuming (even a great deal) of alcohol might not take an individual below the threshold of satisfying IMS. It has not, as far as I am aware, ever altered my mental state to this extent, however, I remain open to the possibility that for some people and some quantities of alcohol, it could.

If we assume that Darrell is one such individual and, on a particular Tuesday night, Darrell drinks himself into a state where he does not satisfy IMS, then we cannot say that Darrell is responsible for his acts/omissions during the time that he does not satisfy IMS. This may not sit well with many, especially those who are considering the damage that could result if Darrell found himself behind the wheel of a car in this state. It may still be appropriate for Darrell to contact serious consequences for driving when so heavily intoxicated. But let us consider some of the things that Darrell is responsible for in this case.
If it is possible for alcohol to disqualify an average individual, with a normal metabolism, from satisfying IMS, it almost certainly takes a great deal of alcohol. Prior to intoxication, while fully satisfying IMS, there were many things that Darrell could have and should have done to avoid ending up drunk, behind the wheel of a car.

Darrell could have avoided getting drunk. I have a policy when I am out and have to drive home – I can have no more than one drink and I must be able to wait at least one hour before driving. It is a simple rule. While I am confident that I could safely operate a vehicle with more than one drink and less than one hour, under my belt, this policy keeps me from every having to try to judge whether I’ve had “too much.” It is easy to lose track of both drinks and time the more that you have, however, if you have only one drink, it is fairly simple to keep track of time. And so, Darrell could have adopted a similar policy and been safe to drive himself home.

If Darrell was determined to get drunk, there were still a number of safety precautions that he could have taken. He could have arranged for a designated driver. He could have handed his keys over to the bartender and requested that a cab be called when he was ready to leave. Alternatively, Darrell could have chosen an environment to maximize his safety and that of those around him. He could get drunk at a place that he would have no pressing reason to leave while intoxicated, be it a friend’s house, a hotel, or his own home. The list goes on.

I believe that it is common in our society to heavily censure those individuals who, while driving drunk, injure others while being far less critical of individuals who drive drunk and make it home safely or even those that drive drunk and are charged with a DUI. What we should consider is that, accidents and injury are largely a matter of moral luck. Failing to take appropriate action to avoid such an easily avoidable, extremely hazardous situation, is a far greater moral failing that we often consider. We should hold the drunk driver responsible (even if IMS is not satisfied) in such a way that the moral significance of these antecedent choices is made more salient rather than claim that the drunk driver only is morally responsible if someone gets hurt and then, only for the hurt.\footnote{I have crafted this case about drunk driving, but similar logic prevails for other cases (ex. date rape under the influence).}
Psychotropic Drugs

Stating that there are some drugs that can alter mental states to the extent that IMS cannot be satisfied is relatively uncontroversial. If you are actively hallucinating, for example, you will fail to satisfy the access to reality component of IMS. As with all cases, if an agent does not satisfy IMS, the agent is not responsible for acts/omissions during the time that IMS was not satisfied.

As with the case of the drunk driver, above, there may still be a host of antecedent acts and omissions that the agent is responsible for. I will assume here that there is nothing inherently immoral about taking psychotropic drugs, which is to say, I will not argue that drugs are morally worse than alcohol because they are illegal and alcohol is not. Assuming the act of taking a psychotropic drug, itself, is morally neutral, how we judge responsibility is going to depend on context. A college student taking LSD in a controlled lab setting as part of a psychology experiment probably has very little to worry about with respect to moral responsibility (in this case).

A single mother, with two young children, taking LSD while home alone with her children, has a great deal more that she can be found responsible for. Even if she fails to satisfy IMS while high, if she made the decision to take the LSD and did so, while satisfying IMS, and without first considering and making provision for the safety of her children, to include the risk associated with being caught/charged for engaging in an illegal activity, then there are many morally relevant things to find and hold her responsible for.

In Conclusion

I have presented what I take to be a necessary condition for morally responsible agency that privileges the mental status of the agent, above other conditions, within a conceptual framework that takes moral responsibility to exist on a continuum and assessments of moral responsibility to be deeply contextual and localized to a time. I have drawn from the fields of philosophy, psychology, and applied behavior analysis to make an argument for how finding and holding responsible can and should come apart in an analysis of moral responsibility.

77 This, of course, is not without philosophical precedent. Holding responsible is an essential component of a proper moral education. Aristotle advocates a certain “fake it till you make it,” strategy in his Nichomachean Ethics that would be friendly to what I have in mind.
REFERENCES


BIOGRAPHICAL SKETCH

Megan Marie McGrew was born in Frankfurt, Germany on April 9, 1983 to Randy R. McGrew and Robin L. McGrew (now Delarosa). She has two younger siblings, Ryan McGrew and Sarah McGrew. Megan’s family moved frequently during her early years, courtesy of the United States Army. Some of her most formative years were spent living abroad, in Belgium, as part of a close-knit multinational community at the Supreme Headquarters Allied Powers Europe (SHAPE).

Megan attended St. Mary’s College of Maryland from 2001-2005 and graduated magna cum laude with a double-major in Philosophy and History and a concentration in Women, Gender, and Sexuality Studies. During her time at St. Mary’s, Megan worked as an archaeological technician and archaeological field school teaching assistant for the Historic St. Mary’s City Foundation.

In 2005, Megan moved to Tallahassee to begin her graduate school career in the Department of Philosophy, at Florida State University. As a graduate student, Megan worked as a teaching assistant as well as a graduate instructor; she also spent several years as the president of the Society for Women’s Advancement in Philosophy (S.W.A.P.). In 2009, she received her M.A. in Philosophy from Florida State University and advanced to doctoral candidacy that same year.

Megan’s early doctoral work focused primarily on issues of ethics and moral responsibility and Megan became increasingly interested in issues concerning mental health. In 2012, Megan made the move from Florida to Georgia and began working within the mental health field while continuing to work on her dissertation. She currently works as a BCaBA (Board Certified Assistant Behavior Analyst) specializing in therapeutic interventions for children with Autism. Megan currently lives outside of Atlanta with her two dogs and her endlessly supportive partner, Dave Donegan.