

4-15-2019

# Managing Madness: The ethics of identifying and treating mental illness

Mason Seely  
*Santa Clara University*

Follow this and additional works at: [https://scholarcommons.scu.edu/lib\\_ugrad\\_research](https://scholarcommons.scu.edu/lib_ugrad_research)

 Part of the [Mental and Social Health Commons](#), [Other Psychiatry and Psychology Commons](#), [Philosophy Commons](#), and the [Psychology Commons](#)

---

## Recommended Citation

Seely, Mason, "Managing Madness: The ethics of identifying and treating mental illness" (2019). *Library Undergraduate Research Award*. 3.  
[https://scholarcommons.scu.edu/lib\\_ugrad\\_research/3](https://scholarcommons.scu.edu/lib_ugrad_research/3)

This Article is brought to you for free and open access by the Student Scholarship at Scholar Commons. It has been accepted for inclusion in Library Undergraduate Research Award by an authorized administrator of Scholar Commons. For more information, please contact [rsfroggin@scu.edu](mailto:rsfroggin@scu.edu).

Managing Madness:

The ethics of identifying and treating mental illness

A thesis submitted in partial satisfaction of the  
requirements for the honors college bachelor's degree program

at

Santa Clara University

by

Mason Seely

Thesis Committee:

Professor Erick Ramirez

Professor Don Riccomini

## **Table of Contents**

**Introduction** (3-4)

**Chapter 1: Defining Psychological Dysfunction** (4-5)

The Biopsychosocial Model (5-7)

Wakefield and The Harmful Dysfunction Model (7-12)

The Sick Label: An Ethical Evaluation (12-14)

**Chapter 2: Responsibility, Rules, and Blame** (15)

Legal versus Moral Responsibility (15)

Responsibility and Mental Illness (15-16)

Attributable versus Moral Responsibility (16-17)

Factor 1: Normative Competence (17)

Cognitive Component (17)

Semantic Capacity (17-18)

Affective Capacity (18-19)

Volitional Component (19-22)

Factor 2: Fair Opportunity (22-23)

Translating Responsibility and Normative Analysis into Clinical Actions (23-24)

**Chapter 3: Clinical Considerations** (25)

Case Study: Alcohol Dependence Disorder (25-30)

Responsibility and Blame (30-31)

**Conclusion** (31-32)

**Bibliography** (33-34)

## Introduction

In the 1960s, diagnostic medicine underwent a revolution. Traditional diagnostic methods used in clinical practice, like interpreting case histories, were discredited as illegitimate science and replaced by methods that looked more like standard scientific models used in research (Horwitz, 2007). After this transition, providers' primary objective became cataloguing symptoms and making discrete diagnoses versus collecting extensive patient histories and ascribing behaviors and actions to underlying psychological needs (Horwitz, 2007).

This change in standards was problematic for dynamic psychiatrists who based their craft around the analysis of case histories, yet needed to maintain their relationship with medicine in order to retain their credibility as professionals that treat mental illness. Without a strong association with medicine, there was nothing that distinguished psychiatry from other forms of social work that also performed psychoanalysis as part of their job (Horwitz, 2007). This meant that dynamic psychiatrists had to change their clinical approach to better fit the new research model spurring the transition from dynamic to diagnostic psychiatry (Horwitz, 2007).

The diagnostic psychiatric model emphasized the precise identification and treatment of discrete, universal mental illnesses. Consequently, an effort ensued to create a standard diagnostic manual of syndromes for a variety of mental disorders, resulting in the birth of the Diagnostic and Statistical Manual of Mental Disorders (DSM). However, because clinicians were going to comprise the bulk of people who used the manual, the DSM needed to satisfy both the conditions of a research model and the interests of physicians in maintaining their client base<sup>1</sup> (Horwitz, 2007). This social and financial incentive led to the transformation of "problems of the living" treated in dynamic psychiatry's upper-class clientele into discrete mental illnesses along with conditions like schizophrenia in the DSM (Horwitz, 2007). Therefore, the list of mental illnesses found in the DSM are not exclusively based on science and in fact continue to be molded by changing cultural values and social action. For example, homosexuality was removed from the DSM-II in 1973 in the midst of changing social opinions and values about the appropriateness of homosexuality (Horwitz, 2007).

Although the DSM is structured to fit the biomedical model, it is molded by social and political interests. This is problematic because western health authorities market the text as a guide to mental illness that emerged as a result of scientific discovery. Standardizing the use of the DSM is tempting because if it is true that the diagnosis and treatment of mental illness is a mathematical algorithm where 'x' number of specific symptoms equals 'y' disease, the manual is a great asset to clinical practice.

However, although the theory that there are specific symptoms and behaviors that universally correlate with discrete mental illnesses is seductive, it is not reality. Regional cultural values shape the opinions, behaviors, and norms for individuals in a society meaning that the symptoms and behaviors associated with mental illness are dependent on

the social context in which they are being evaluated. Trying to use a strictly objective model for mental illness is therefore problematic because social values are an intrinsic part of the diagnosis process<sup>2</sup>.

Moreover, there are other reasons why a more scientific model for mental illness is insufficient. Western societies “typically contain legislation that implies a special relationship between mental illness and responsibility” (Edwards, 2009). The mental illness label comes with certain assumptions about a person’s rational and volitional capacities and thus their ability to act in a reasonable manner and control certain behaviors. When we ask questions about whether a person is mentally ill, we are simultaneously asking if we believe that the person should not be responsible for actions related to his condition. This decision is intrinsically normative as debates over what we ought to do requires ethical analysis (Edwards, 2009).

In this essay, I describe the benefits of using a normative conception of mental illness and argue that this model is the best way to achieve appropriate psychiatric diagnoses. First I outline several popular scientific frameworks for understanding mental illness and point out the problems that arise in these more objective models. Then I discuss the benefits of a more normative account and expand on current ethical models by providing an account of the capacities and conditions necessary for responsibility. Finally I marry my account of responsibility with a normative account of mental illness and illustrate how this framework creates space for different degrees of patient responsibility by applying it to specific patient case studies. Ultimately, I aim to create a more appropriate and comprehensive model for recognising and treating psychiatric dysfunction in clinical practice. Physicians need to shift away from using more objective accounts of psychological dysfunction in practice and embrace the ethical implications of the mental-illness label. By carefully balancing normative and pragmatic considerations, providers can create more effective and just therapeutic regimens that are tailored to individual patients’ circumstances and needs.

### **Chapter 1: Defining psychological dysfunction**

There are many different scientific, philosophical, and ethical models used to understand what constitutes mental illness. These conceptions of psychiatric disease vary by how much importance is placed on physical, mental, and social determinants of health as well as their focus on mind-independent symptoms (behavior, neurological changes, physiological changes) versus mind-dependent symptoms (dysphoria, delusion, rationality). In all accounts for defining psychiatric disorder, there is a common question raised and addressed as to whether diagnosing mental illness is possible using objective measurements versus normative contextual analysis.

The latest edition of the DSM, the predominant authority for identifying and categorizing mental illnesses in clinical practice in the United States, determines what constitutes a mental disorder using the parameters below:

a mental disorder is a syndrome characterized by clinically significant disturbance in an individual's cognition, emotion regulation, or behavior that reflects a dysfunction in the psychological, biological, or developmental processes underlying mental functioning. Mental disorders are usually associated with significant distress or disability in social, occupational, or other important activities. An expectable or culturally approved response to a common stressor or loss, such as death of a loved one, is not a mental disorder. Socially deviant behavior (e.g. political, religious, or sexual) and conflicts that are primarily between the individual and society are not mental disorders unless the deviance or conflict results from a dysfunction in the individual, as described above (American Psychiatric Association, 2013, 1)

There are several key issues with the criteria defined by the DSM-5. What constitutes a "clinically significant disturbance" is very subjective and dependent on patient and provider factors. The source of an individual's mental dysfunction that causes this disturbance is attributable to biological, psychological, and social factors in undefined proportions, which leaves the causative agent for a patient's mental illness broad and vague (Ghaemi, 2003). Additionally, judging if a specific response is culturally appropriate or acceptable to a given stressor is subjective and non-universal. This becomes problematic for psychiatry, which must use objective measures and science to form diagnostic criteria in order to be considered a branch of medicine (Horwitz, 2007). Finally, what is meant by the term "dysfunction" and how to assess whether an individual has an internal dysfunction that causes abnormal behavior (mental illness) or is deviant for alternative reasons is not specified (Wakefield, 1992).

Different models for identifying what constitutes a mental disorder deviate or compliment the DSM to different degrees. Some of the frameworks are incorporated within the DSM-5 (i.e. the Biopsychosocial model) while others like the Harmful Dysfunction model are in response and to it. Below, I outline some of the predominant accounts used to understand and define psychological dysfunction with a consideration of the benefits and potential problems associated with each lens. Finally, I end by suggesting that a deeply normative model of mental illness (and mental illness ascriptions) best reconciles our descriptive and normative practices, including the development of mental illness categories, the removal of old ones, and diagnosis/treatment in therapeutic practice .

### **The Biopsychosocial Model**

Within the current medical profession, the biopsychosocial (BPS) model of mental disorder has taken the position of authority for determining the causes and requirements of mental illness and shapes DSM-5 criteria as well as treatment approaches. Under this framework, mental illness is understood to be the product of biological, psychological, and social factors that all contribute to the formation of psychiatric disease (Ghaemi, 2003). Nassir Ghaemi, an academic psychiatrist, asserts that although this holistic understanding of illness is comprehensive and valuable, in practice, such a broad approach has evolved into "a confusing set of assumptions about the content of mental disorders... and does little but

assert that all illnesses have components that are, unsurprisingly, biological, psychological, and social” (Ghaemi, 2003).

One of the major issues with the biopsychosocial model is that it does not identify the degree to which each component (biology, psychology, sociology) contributes to the underlying cause of disease. The BPS model is like a “list of ingredients, as opposed to a recipe. To cook a meal, it is not sufficient to simply know the list of ingredients. One also has to know how much of each ingredient to use, and in which order” (Ghaemi, 2003).

The lack of clarity about how much different factors contribute to psychiatric disorder is concerning considering how much the BPS model is used in clinical practice and the spectrum of treatment options currently offered for mental illness. The BPS model “only lists relevant aspects of psychiatry; it is silent on how to understand those aspects under different conditions and in different circumstances” (Ghaemi, 2003). Without this kind of specificity, both psychopharmaceutical and psychoanalytic methods are theoretically valid therapeutic options for any given mental illness. If, as assumed by the BPS model, all psychiatric illnesses “are biological, psychological, and social... then it would seem to follow that everyone should receive both biological and psychological treatments” (Ghaemi, 2003). However, this is a “false and faulty belief that stems directly from the biopsychosocial model” and in clinical practice is a harmful oversimplification<sup>2</sup> (Ghaemi, 2003).

For example, take two patients who are both diagnosed with major depressive disorder using the characteristic symptoms of this condition outlined in the DSM-5. The DSM-5 states that an individual must be experiencing five or more of the eight symptoms below during two consecutive weeks and at least one of the symptoms must be depressed mood or loss of interest/pleasure.

1. Depressed mood most of the day, nearly every day.
2. Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day.
3. Significant weight loss when not dieting or weight gain, or decrease or increase in appetite nearly every day.
4. A slowing down of thought and a reduction of physical movement (observable by others, not merely subjective feelings of restlessness or being slowed down).
5. Fatigue or loss of energy nearly every day.
6. Feelings of worthlessness or excessive or inappropriate guilt nearly every day.
7. Diminished ability to think or concentrate, or indecisiveness, nearly every day.
8. Recurrent thoughts of death, recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide (APA, 2013)

To receive a diagnosis of depression, these symptoms must cause the individual clinically significant distress or impairment in social, occupational, or other important areas of

functioning. Additionally, these symptoms cannot be a result of substance abuse or another medical condition.

Patient one has been experiencing all eight symptoms nearly every day ever since her brother and father died in a car accident. Her depression is caused by a traumatic life event in which she lost a significant portion of her family, and her condition is sustained because of the economically depressed neighborhood she lives in, limiting her job opportunity and hope.

Patient two has not experienced this kind of life event and lives in a stable environment. He has felt sad and worthless for as long as he can remember. His family reports that they can't remember a time when he was ever happy, and that his weight, daily behaviors, and overall pleasure have remained constant. Patient two is diagnosed as depressed, but unlike patient one, his depression is caused by an inherent reduced dopamine receptor function.

A physician abiding by the BPS model would likely assume that the causes for patient one and patient two's depression were a complex mixture of biological, psychological, and social components and prescribe similar treatments to each person though the source of the condition is not the same between them. Patient one's depression is primarily due to social and psychological forces that leave her isolated and hopeless while patient two suffers from an imbalance of neurotransmitters (i.e. a concentration of neurotransmitters that significantly deviates from normal neuroanatomical ranges). Treating both patients with antidepressant medications and therapy doesn't lead to equal or optimum therapeutic effect for either patient though the disease label is identical. As noted above, using DSM-5 criteria, both patients would be diagnosed with the same underlying illness. Given their very different etiologies, this assumption may be unwarranted.

Using the BPS model, current clinical assessment of mental illness is often more focused on linking syndromes to discrete disorders versus identifying what is causing the symptoms because it is preemptively assumed that multiple and complex factors cause psychiatric disorders (Horwitz, 2007). Consequently, the BPS framework leaves clinicians with a spectrum of causal factors that are too broad and ill-defined to effectively treat the specific source of mental illnesses, and treatment becomes more focused on mitigating symptoms than discovering why these symptoms are occurring. The biopsychosocial model does a good job acknowledging that there are numerous forces that can lead to mental illness. However, without clearly defining to what degree each component influences the causal root of disease, and without a clear specification of what mental disorders *are*, it is an incomplete model for assessing and treating psychiatric disorders.

### **An alternative: Jerome Wakefield's 'Harmful Dysfunction' Analysis**

Jerome Wakefield believes that problems with the BPS model and the DSM-5 arise at a fundamental level as neither offers a clear reason why specific syndromes are mental disorders. He asserts that a correct and clearer understanding of disorder "is essential for

constructing ‘conceptually valid’ diagnostic criteria that are good discriminators between disorder and nondisorder” (Wakefield, 1992). Without a concrete definition of dysfunction, the BPS model and the DSM-5 lack a solid foundation for identifying why and when a set of behaviors and cognitions are considered disordered and this creates a problem at the earliest point in the diagnostic process. It is illogical to talk about what forces contribute/cause mental illness without a specific conception of when a syndrome is a psychological dysfunction.

Wakefield aims to expand the scope of his discussion by creating a unifying account of disorder that can accommodate physical and mental illnesses. His “harmful dysfunction (HD)” model offers clear criteria to identify medical disorders based on the premise that illnesses are best recognized as having subjective and objective components. Wakefield asserts that neither lens alone sufficiently explains why we call certain conditions disorders. Thus, he concludes that “a disorder exists when the failure of a person’s internal mechanisms to perform their functions as designed by nature impinges harmfully on the person’s well being as defined by social values and meanings” (Wakefield, 1992). In Wakefield’s HD model, dysfunction “is a scientific term based in evolutionary biology” and describes “the failure of an internal mechanism to perform its natural function for which it was designed” while harmful is meant to capture the “consequences that occur to a person because of the dysfunction and are deemed negative by sociocultural values” (Wakefield, 1992). The HD framework contains an objective etiologic criterion as an attempt to be value-neutral and requires that a dysfunction be harmful to social function in order to acknowledge the role of culture in shaping conceptions of illness and health. This combination of scientific and normative analysis creates a system for identifying disease that is biologically valid and can be used in multiple contexts.

Under the HD model, a disorder is more than statistical deviance or the presence of “lesions” (abnormalities in structures) (Wakefield, 1992). Wakefield argues that there are many instances of statistical deviance that we don’t consider illness, like acute intelligence, as well as statistically normal conditions, like obesity, that we do consider illness and consequently statistical frequency is a poor criterion for disorder (1992). Similarly, the presence of a lesion doesn’t always indicate illness, like having a fully functional heart positioned on the right side of the body, and there are plenty of cases where illness exists even in the absence of a lesion like senile pruritus (itchy and scaly skin) (Wakefield 1992). These examples demonstrate that most strictly objective analyses of disorder are incomplete. Etiological conceptions of disorder, like Wakefield’s, have an advantage over statistical or lesion-based views; however, Wakefield himself argues that this is not enough to ground a theory of mental or physical illness. This is because we must also consider cultural and social differences in how dysfunctions are understood and accommodated. Because the illness label is not universal, Wakefield concludes that categorizing a condition as a disorder is not neutral or independent of social values and perceptions and requires some degree of normative evaluation.

However, Wakefield is not satisfied with a purely subjective conception of psychiatric disease. He claims that “the fact that all disorders are undesirable and harmful according to

social values shows only that values are part of the concept of disorder, not that disorder is composed only of values” (Wakefield, 1992). He supports his claim by providing examples of many undesirable conditions like pain during teething or externally driven instances of misfortune like poverty that are not socially desirable yet not considered disorders (Wakefield,1992). Additionally, he argues that a purely subjective approach only identifies what symptoms are indicators of illness but not the cause of these symptoms (Wakefield, 1992). Recognizing these problems, Wakefield concludes that a good theory of disorder requires a consideration of designed biological function. He ultimately concludes that illness, both physical and mental, is the result of evolved (etiological) mechanisms not performing the job they were designed to perform and that therefore harm the person with the disorder (Wakefield, 1992). For example, if a heart could not circulate blood than we would say that it is not performing its etiological function in the cardiovascular system. Such a disorder would lead to straightforward harms to a person and thus would be an illness in Wakefield’s view.

To account for cases of mental disorder that have no clear physical dysfunctional mechanism, like delusions, Wakefield asserts that etiological mechanisms exist at both the physical and mental levels (Wakefield, 1992). He argues that certain thought processes, moods, or other psychologically related behaviors negatively or positively influence an individual’s fitness and consequently are subject to the process of evolution (Wakefield, 1992). In the delusion example above, Wakefield may argue that having intense delusions could be the result of a mental mechanism responsible for perception malfunctioning, and this lack of access to reality could be a harmful to an organism’s ability to perform essential functions like obtaining food. Therefore, delusions could be a harmful dysfunction rooted in a malfunctioning mental mechanism.

Using the harmful criterion in conjunction with the etiological requirement, Wakefield is able to clarify that a dysfunctional evolved mechanism only results in a disorder when the defect causes an undesirable outcome for the individual. For example, if the heart circulates blood through an alternative means compared to normal circulation, yet did so in a manner that still supported the organism’s physiological needs, then this heart would not be considered diseased because the abnormality doesn’t disrupt the organism’s ability to maintain homeostasis or participate in social/physical environments (Wakefield, 1992).

Within philosophical literature, Wakefield’s HD model is perhaps the most agreed-upon conception of disorder as it balances both objective and normative factors that contribute to the dysfunction label. However, this model is better equipped to define physical disorders and is incomplete and problematic for mental disorders for several reasons. Wakefield’s discussion of mental mechanisms allows for the inclusion of psychological disorders like delusions to fit under the HD model, but identifying the evolved function of these mental mechanism is difficult and imprecise. Two major opponents of the HD model, Dominic Murphy and Robert Woolfolk, describe some of the major obstacles to identifying dysfunctional evolved mental mechanisms, including the existence of spandrels and the mismatch between etiological design and current environment. Wakefield’s own response to these objections illuminates the weaknesses in his model.

A spandrel refers “to any undersigned side effect of design” and in the context of biology would capture any feature that was not selected for directly but instead is the result of a different evolved characteristic (Wakefield, 2000). The human chin is the classic example of a biological spandrel as the chin itself serves no adaptive function and is only present as a byproduct of other evolved functions and mechanism involved with chewing and respiration. Although a chin is useful, even potentially relevant to evolutionary fitness, it does not come directly from evolutionary processes and instead is an indirect result.

Woolfolk and Murphy object that “if mental spandrels exist, then there are mental mechanisms that are the byproducts of evolution, but have themselves never possessed adaptive functions,” and therefore under the HD analysis, these mechanisms cannot be dysfunctional because they lack etiological functions, yet they still appear capable of causing pathology (Wakefield, 2000). Wakefield assumes that useful capacities are likely directly evolutionarily developed, but this isn’t necessarily true. For example, the capacity to understand math may be an indirect product of pattern recognition. Finding patterns in nature could be an etiological capacity as it could allow humans to remember what kinds of plants/animals were dangerous or could be used for food. While similar processes may aid someone in understanding math, the ability to do math isn’t the direct result of evolution. Therefore, someone who is incapable of doing math can never be considered disordered because the capacity to do math is not an etiological function. If an evolutionary story can be told that is irrelevant to why a trait exists, this begs the question as to how do we know what capacities are products of evolution and thus can be dysfunctional? Woolfolk and Murphy recognize this problem and argue that distinguishing between evolved mental mechanisms and spandrel mental mechanisms is imprecise and cannot be done only with biological science. Additionally, if mental spandrels produce disorder, then the HD analysis doesn’t account for these diseases and is consequently incomplete.

Wakefield responds to these objections as misunderstandings of the HD analysis but his examples seem to demonstrate more problems than offer solutions. For example, Wakefield claims that “failure of a spandrel implies a disorder when and only when it implies the failure of a naturally selected function” yet offers no clarification or examples of how to differentiate between evolved and spandrel mental mechanisms (Wakefield, 2000). His counterexamples are mostly limited to physical cases like the inability of the nose to hold up a pair of glasses as an instance of a failed spandrel that doesn’t imply disorder (Wakefield, 2000).

When trying to apply the same reasoning to dyslexia, Wakefield tacitly appeals to normative criteria to determine if an evolved dysfunction is present. Recall that Wakefield’s appeal to etiological function was intended to capture the value-neutral, properly scientific, essence of diagnostic criteria. By using normative conceptions of personhood to deduce the evolutionary purpose of a mechanism, he undermines his project. He claims that when people fail to learn how to read because they “lack educational opportunity, or are unmotivated” or are otherwise impacted by other external factors that this inability doesn’t indicate disorder (Wakefield, 2000). However, if a person fails to learn how to read

“even under optimal learning conditions,” then “we infer that there is something wrong with some internal neurological mechanism that when functioning as designed supports the capacity to read” (Wakefield, 2000).

A few things must be noted about this response. For one, Wakefield cannot mean that humans have evolved the capacity to read as reading is a relatively new ability for the species. Thus, Wakefield is assuming that there are etiological mechanisms that when normally functioning support the ability to read, and he concludes that the inability to read suggests that one of these mechanisms is dysfunctional. This argument is weak and problematic for several reasons. Wakefield uses his own definition of etiologic dysfunction in his response but never explains how or why we are able “to infer that there is... some internal mechanism that when functioning as designed supports the capacity to read” (Wakefield, 2000). He assumes that there are evolutionary purposes for mechanisms that aid in reading but never proves this point and doesn’t offer clear ways to distinguish between evolved versus spandrel mental mechanisms. Wakefield uses a socially desirable skill, reading, to determine the presence of dysfunction, but this makes the objective etiological requirement contained in the HD framework dependent on social values. He depends on normative considerations in order to evaluate both the objective and subjective criteria in his HD model and consequently falls short of capturing all instances of dysfunction using normative *and neutral* requirements. The presence objective and ethical components is essential to Wakefield’s argument, but this requirement is clearly violated in this example revealing gaps in the HD model.

What is needed is clarification as to why under certain circumstances we would want to consider certain behaviors and cognitions the result of a mental disorder. In his argument, Wakefield hints at the necessity for contextual evaluation when he classifies dyslexia that results from poor education, negative attitude, etc., as normal but dyslexia that isn’t caused by other harmful contributing sources, as disordered. Wakefield's inability to clearly explain why lacking socially useful capacities, like the ability to read, is evidence of an evolved function (despite lacking other comorbid psychological problems) is evidence that values are playing a neglected or misunderstood role in even his view of illness. Wakefield tries to avoid using unscientific values by leaving the scientific concept of dysfunction value-neutral. But, given that values are part of both sides of his analysis, this is a problem.

The necessity for normative analysis when classifying dysfunction shouldn’t be framed as an opportunity for error and injustice. There is a “nonarbitrary justification for our classification of mental illnesses, but it is based in ethical truths rather than value-free features that are unique to mental illness” (Edwards, 2009). A greater appreciation of external factors and individual characteristics aids diagnosis and should be evaluated when determining the presence of mental illness. With a clear system for specifying when we ought to consider an individual disordered and understanding of what this disease label means, we increase our ability to capture psychiatric dysfunctions in a way that best supports broader social values and individual interest. In my analysis, I will be embracing the role of value, especially in the application of illness, and explain the benefits and reasons for doing so below.

## **The sick label: an ethical evaluation**

The appropriate response to different kinds of dysfunctions invokes values and requires an “ethical reflection upon our legal and social rules” (Edwards, 2009). For one, we must consider what we are assuming about the individual’s capacities when we diagnose them with illness. The “sick label” is a normative label that “*ethically justifies* certain consequences, such as the ill person adopting the role of patient and social denial of that person’s responsibility for some of the person’s actions” (Edwards, 2009). When we allow someone, via diagnosis, to take on the sick role, we are saying that the individual is a passive victim in respect to the disease or disorder in question and that he or she cannot control certain actions/thoughts that are regulated by the condition. We assume that illnesses are external agents that act on patients beyond their control that force them to behave in specific ways that may contradict social and legal rules for appropriate action. Because the individual is being controlled and coerced by their condition, they they are not morally or legally responsible for any actions or cognitions caused by the disorder that violate social or legal norms.

Given the overlap of behaviors and emotions that can be associated with both mental illness and deviant behavior, we must recognize why we discourage these behaviors and qualities and define under what conditions it’s appropriate to say that these actions and characteristics are not within the reach of the individual. There needs to be an account of psychiatric disorder that differentiates “mental illnesses from the vast array of irrational and pre-rational drives and personality traits for which we usually wish to hold the bearer morally responsible” (Edwards, 2009).

For the kind of value-laden theory I wish to advocate, mental illness is not a consistent set of symptoms; “it is a label that stipulates how people should respond to the condition” (Edwards, 2009). When a person has a psychiatric disorder, we reduce the individual’s responsibility for actions and cognitions controlled by the illness, but we do so at the cost of the individual’s autonomy and ability to take ownership of his or her behavior and characteristics. Mental illnesses are “dysfunctions in personhood” and make it so an individual is “unable to to fully function as a person owing to impairments of processes and capacities that are necessary for being a person” (Edwards, 2009). Consequently, we do not grant these individuals the same degree of autonomy or responsibility that we would a person with full capacities and abilities.

It should be noted that diagnosing physical illnesses is not primarily ethical. Physical illnesses often have more concrete and binary associated syndromes making them easier recognize and properly diagnose. Physical diseases are almost always cases of “biological dysfunctions” while mental illnesses are often identified because they are “social dysfunctions” making ethical considerations more integral to the decision of whether someone is mentally ill (Edwards, 2009).

The mere fact “that a biological process is impaired is not sufficient for reduced autonomy or responsibility,” instead it is the effect that the malfunctioning mechanism has on “one’s

existence not just as a living organism but as a social being with a mental life that is capable of being the subject of rights, duties, and responsibilities-- that is relevant to autonomy and responsibility" (Edwards, 2009). When using a normative framework for classifying mental illnesses, it is important to specify under what circumstances and for what actions a person with the sick label has reduced autonomy and responsibility. For example, someone with heart disease is not considered mentally incompetent and unable to make autonomous decisions. However, someone who suffers from a discontinuity in character, lack of capacity for rational agency, or inability to respond to reason is not able to access, interpret, or act on the relevant information necessary to make informed decisions about his or her specific condition. Therefore, a person who exhibits any of these characteristics can be a proper target for the mental illness label (Edwards, 2009).

A person who has a mental illness is not morally and/or legally responsible only for actions and cognitions *directly* impacted by the condition (Nelson and Ramirez, 2017). Imagine that an individual tears down and defaces the front of a stop sign. If the individual who tore down the sign is anorexic, she would still be held responsible for destroying public property because the disease (anorexia) doesn't directly impair the mental and physical processes utilized in deciding and acting to destroy a sign. Society would rightfully hold her morally and legally responsible for her actions and blame/punish her for tearing down the sign.

However, what if we discovered that the person who tore down the sign suffered from schizophrenia, and that his condition generated delusions that signs were secret agents meant to kill him? In this case we may choose to consider him non-autonomous and thus not hold him responsible for his actions. This is because his condition makes it so he cannot access reality in a way that allows him to obtain rational and reasonable information about the purpose and intentions of the signs and respond accordingly; the condition impedes his access to reasons directly involved in this decision-making process. A person should have reduced responsibility and autonomy for thoughts and actions that are directly impacted by a dysfunctional mechanism and can be granted the sick label in respect to these behaviors and conceptions (Nelson and Ramirez, 2017).

So far in this discussion, we have looked at what the sick label means for the individual in question. However, it is important to recognize that these choices are influenced by social values and that our final conclusions should reflect broader rules and virtues (Edwards, 2009). As a person in society, an individual's actions and thoughts are evaluated in relationship to a larger environment with certain rules and virtues. When using a normative account of mental illness, the answer as to whether a condition is a mental illness changes based on the ethical context in which the decision is made and requires a careful consideration of what is just at both micro and macro scales.

Rules and values that we hold in society are a balance of individual right versus social duty/safety. Actions like stealing, violence, and lying are not tolerated in a civil society because they endanger the ability of the group to work cohesively under formal and cultural standards of compassion and nonviolence. We enforce these principles by teaching children their responsibility as citizens to obey social laws and morals under the

assumption that people are capable of controlling actions and conceptions involved in these rules and values (Edwards, 2009). For example, when we punish a child for acting violently and impulsively, “we teach them that they must not only avoid committing rationally calculated evils, but also must actively avoid adopting character traits that lead to evil acts” (Edwards, 2009).

When we deem that a certain mental condition or action is caused by psychiatric dysfunction, “we clarify that it is not a legitimate choice for personal development” or behavior (Edwards, 2009). Persons are morally responsible for developing and maintaining positive personality characteristics “*to the extent that such development can in fact be taught*” to the individual (Edwards, 2009). It is logical to create and enforce formal and informal rules for behaviors within the control of the individual because control goes hand in hand with autonomy and responsibility.

Given the intimate link between autonomy/responsibility and mental illness, a clear framework and process of determining autonomy and responsibility would be beneficial to normative conceptions of psychiatric disorder. However, creating a straightforward or well-defined system for determining the degree of agency or control that an individual has is challenging. Indeed “rationality is the subject of such an enormous quantity of philosophical debate that it is likely impossible to describe without begging some questions” (Edwards, 2009).

Though difficult to craft, an account of agency and responsibility that allows a clinician to define these properties on a case-by-case basis is needed when using a normative framework of mental illness. With an enhanced understanding of what capacities are necessary for responsibility and agency, we could better identify when and what mental conditions influenced these capacities, and consequently assign reduced responsibility and autonomy (i.e. the sick role) more appropriately. In the following chapter, I aim to create this framework and marry it to a normative model of mental illness. By joining these two accounts, I ultimately craft a stronger and more comprehensive ethical framework for recognizing mental illness and assigning the sick role.

### **Chapter 1 Endnotes**

- 1) Although I focus on the DSM throughout this document, it should be noted that the International Classification of Diseases (ICD) manual has undergone similar, and in some cases even more problematic, changes
- 2) In the United States, there has been a growing movement in clinical research of mental illnesses to abandon the criteria used within the DSM for different mental illness categories. In this process, the emergence of new research domain criteria (RDoC) for mental illnesses has started to take shape. Researchers investigating and crafting these new criteria aim to define more precise, reliable, and objective guidelines for diagnosing psychiatric diseases. With a better understanding of the roots and etiology of mental illnesses, clinicians can prescribe more appropriate treatment.

## **Chapter 2: Responsibility, rules, and blame**

Questions about responsibility, rationality, and agency are at the center of countless philosophical debates. These are difficult topics because they run parallel to ideas about personhood and require extensive normative analysis (Edwards, 2009). In this chapter, I focus on models of responsibility based on features like rationality, receptivity to reason, and reactivity. After articulating the distinctions between these capacities, I offer an account of responsibility that is a combination of normative competence and fair opportunity.

My criteria create a model of responsibility that, when married to the normative conception of mental illness, provides a comprehensive and more ethical approach to diagnosing and treating mental illness. This framework challenges current clinical methods by suggesting that patients diagnosed with the same condition using DSM-5 criteria can have different degrees of responsibility for their condition, thus not equally entitling them to the sick role. What constitutes proper therapeutic response should thus be based on a case-by-case analysis of different individual and societal ethical considerations. The ultimate goal is to demonstrate that there is room for improvement and growth in current psychiatric practice that can be partially achieved by incorporating a more moral perspective in clinical contexts.

### **Legal versus Moral Responsibility**

Societies create formal rules and laws in order to enforce behaviors and standards that embody their core values and morals. In the justice system, “criminal punishment is the authorized deprivation of an agent’s normal rights and privileges, because he or she has been found guilty of a criminal act,” and punishment as a form of blame requires that the agent has committed “culpable wrongdoing” or is responsible for the action being punished (Brink and Nelkin, 2011).

Although “criminal law reflects central assumptions about moral responsibility,” legal and moral responsibility are not the same thing (Brink and Nelkin, 2011). This distinction is important but beyond the scope of this discussion. In the sections below, I focus on moral culpability and create a framework of capacities and situational conditions necessary for responsibility.

### **Responsibility and mental illness**

The sick role has direct impact on the degree of responsibility an agent carries for specific actions, under the assumption that individuals are passive victims to their condition (Edwards, 2009). When we clarify what capacities are required for responsibility, it helps reveal what conditions are mental illnesses (i.e. the conditions that impact these abilities). Insights into concepts described below like normative competence and fair opportunity

will give us better tools for making appropriate decisions as to when a condition ought to be labeled a mental illness and a person be allowed to adopt the sick role. Therefore, attaching a framework of responsibility to a normative account of mental illness creates a stronger and more comprehensive framework.

When trying to decide if an agent should be held responsible for deviant behavior, we are not asking whether the agent did or didn't act in alignment with moral norms or even if the agent had first order knowledge (semantic understanding) that his actions violated these norms. What needs to be determined is if the agent had the capacities necessary to recognize, understand, and react to the ethical consequences of his actions as well as a fair opportunity to exercise these capacities in a manner that matched his will. Responsibility must be predicated on "the possession, rather than the use," of volitional and cognitive capacities and a fair opportunity to act on these abilities (Brink and Nelkin, 2011). My proposed model of responsibility as a combination of normative competence and fair opportunity is in service of the larger goal of this paper to craft guidelines for determining when and why a condition should be considered a mental illness and a person be allowed to adopt the sick role.

### **Attributable versus moral responsibility**

In discussing *moral* responsibility, it's important to distinguish this type of responsibility from attributable responsibility. Moral responsibility is a subtype of attributable responsibility. These two can be teased apart by looking at emotional responses to certain actions.

Reactive attitudes are "emotional responses directed at oneself or another in response to that person's conduct" and include love, hate, pride, etc. (Brink and Nelkin, 2011). Within this broad class of emotions is a subgroup of moralized reactive attitudes that "reflect assumptions about responsibility" (Brink and Nelkin, 2011). Take for example happiness and gratitude. If one day you showed up to work and there were cookies at the office you would likely feel happy that this treat had magically appeared. However, without a clear source of where these cookies came from, it would be odd to feel gratitude as this emotion is best directed towards a specific agent who acted in way that benefited you. If Sally from the office made you cookies, you would likely be grateful to Sally as Sally made them specifically for your enjoyment and behaved in a way that impacted you positively. In this scenario Sally's actions (giving you the cookies) matched her "quality of will in the right way" as she intended to give you the cookies and thus she is responsible for the action and a proper target for gratitude as a moral reactive attitude (Brink and Nelkin, 2011).

However, if you later found out that Sally wanted to give those cookies to Steve, but accidentally put them on your desk instead of his, you may be happy that you have cookies but you likely wouldn't feel gratitude towards Sally because Sally didn't intend to give you the cookies. This example shows how responsibility requires both attributability (the agent identified is the agent that did the action) and accountability (the action performed was the action the agent meant to perform). Reactive attitudes are appropriate "just in the case the

targets of these attitudes are responsible” and importantly it is the “responsibility of the targets that makes the reactive attitudes toward them fitting” (Brink and Nelkin, 2011). An agent must be the actor for a given behavior and the action must reflect the agent’s intentions in order to be a proper target for moralized reactive attitudes like blame and praise. When we understand that the behavior of an individual is regulated by certain cognitive capacities and situational factors, we create room for different degrees of responsibility and thus different degrees of ‘blameworthy-ness’ and ‘praiseworthy-ness’ that an agent is appropriately able to possess.

When ascribing moral responsibility, intuition can provide us with some seemingly obvious answers. We do not blame leaves for falling or lions for hunting gazelles, yet we do hold most adults responsible for their actions and have varying degrees of responsibility we assign to adolescents. How can we begin to tease apart the different degrees of responsibility that are appropriate in each of these cases? Below, I address this question by describing the two necessary components of responsibility, normative competence and fair opportunity, and the subcomponents and capacities included in these two factors that comprise responsibility.

### **Factor 1: Normative Competence**

In a nutshell, normative competence is a combination of cognitive and volitional capacities that reflect the ability to tell right from wrong and behave in a manner that fits this moral understanding (Brink and Nelkin, 2011). Both the cognitive and volitional units of normative competence require semantic and affective capacities making them nuanced and multidimensional. Below, I expand on the capacities captured within these two components of normative competence.

#### **Cognitive component**

Responsibility is partially predicated on normative competence which “requires the cognitive capacity to make suitable normative discriminations, in particular, recognize wrong doing” (Brink and Nelkin, 2011). Typically, cognitive capacities are assumed to be completely semantic in nature, requiring only formal knowledge to be well developed. However, this conclusion is misguided. Emotional shortcomings or the inability to step away from certain emotions when they arise may compromise cognitive capacity which in turn can impact responsibility (Brink and Nelkin, 2011). Understanding how cognitive capacity is comprised of both semantic and affective components is best achieved when each dimension is analyzed separately.

#### **Semantic capacity**

The semantic component of cognitive capacity is usually easier to recognize than the affective component. We would not hold a dog and an adult equally responsible for intentionally killing a cat because the dog is not sentient or self-aware, thus doesn’t possess the intellectual capacities to understand moral reasons. The ability to

acquire and have this kind of first order knowledge is assumed to be restricted to mankind.

However, the importance of semantic capacities can also be demonstrated by comparing two adults. If two individuals, one a man and one a woman, both hit other people intentionally though they had the affective and intellectual capacities to understand this was wrong, then they would be equally responsible for hitting other people. However, if we later learned that the man was drunk when he hit the other person, though we may blame him for getting drunk, we do not assume he is as blameworthy as the sober female for hitting someone because we know that consuming alcohol reduces an individual's semantic capacities. How responsible and therefore blameworthy the man and the woman are *in regards to hitting other people* is different because of differences in their intellectual capacities at the time the action was committed.

### **Affective capacity**

Emotions are critical components of attitudes that are essential in motivating or guiding the action of agents. Moral value “is not identified with a natural quality objectively present in morally considerable beings” instead it is “projected by valuing subjects” with requires some degree of empathetic or affective capacity (Callicott, 2010). Understanding moral reasons requires “the capacity to appreciate their force as moral reasons, which involves the capacity for a certain kind of immediate emotional response” (Greenspan, 2003). If an agent lacks certain affective capacities, he or she may lack the ability to recognize/experience the “motivational effect” embedded in moral rules and consequently lack “a full appreciation of moral reasons” (Greenspan, 2003).

For example, psychopaths are “not personally affected or inhibited by others’ emotional responses in the normal fashion” and consequently fail to properly learn socially accepted moral rules because they never associate past wrongdoings with “emotional anxiety, as needed to bring their past failures to bear on practical reasoning” (Greenspan, 2003). These agents “understand reasons generally but are just insensitive to a certain class of reasons,” moral reasons, which involve “something like the appropriate assignment of value” to objects (Greenspan, 2003). Normal self-control involves “the ability to inhibit action more or less automatically, on the basis of emotional responses that reveal moral significance”; thus agents who lack experience in emotional empathy “have a kind of ‘moral learning disability’ limiting them to more roundabout means of self-control such as reflection on the likely consequences of lawbreaking” (Greenspan, 2003).

However, agents that do have complete affective capacities still may be candidates for reduced responsibility. In order to be normatively competent, a person must not “simply act on their strongest desires, but be capable of stepping back from their desires, evaluating them, and acting for good reasons” (Brink and Nelkin, 2011). For

example, the degree of blame we assign to an adult for hitting another person is different from the degree of blame we place on an eight year old for hitting another person. Why? Both of them are the agents that did the hitting, and they have both been formally taught that hitting another person is wrong, thus knowing at a semantic level that hitting another person is considered deviant behavior.

In this case, the different amounts of culpability we assign to the adult and the child reflects our assumptions about the capabilities of each person to control his emotions. We do not believe that the cognitive mechanisms in place that allow for an individual to control and act above their desires are as mature in an eight year old as they are in an adult. Consequently, the eight year old doesn't have a fully developed normative competence and is not blameworthy or responsible to the same degree as the adult is. Normative competence requires both cognitive capacities "to distinguish right from wrong and volitional capacities [discussed in greater detail below] to conform one's conduct to that normative knowledge" (Brink and Neilkin, 2011).

Full cognitive capacity is a combination of affective and semantic components. Deficits in either dimension of cognitive capacity can lead to a compromised normative competence which in some situations results in an agent having at least reduced responsibility for the action. An agent "can be held responsible for failing to react to reasons he does recognize, making him at least partially responsible for his actions" (Fischer and Ravizza, 1998). Because a failure to recognize such reasons varies with each situation, case-by-case analysis is critical when we understand complete cognitive capacity as a mixture of affective and semantic components. Because cognitive capacity is essential to normative competence, a required component for responsibility, proper conclusions about the appropriate degree of responsibility to assign to an agent are best achieved when each aspect of cognitive capacity is individually and comprehensively assessed. The space for different degrees of responsibility only grows as we begin to untangle the volitional component of normative competence as discussed below. Zooming back out to the overall goal of this chapter to better understand responsibility in order to make better conclusions about the appropriateness of the mental illness label, we see that differential responsibility suggests that the sick role may also come in degrees.

### **Volitional component**

Though cognitive and volitional capacities are integrated and often co-dependent, they are nonetheless distinct and neither is individually sufficient for complete normative competence. Like cognitive capacity, volitional capacity requires semantic and affective appreciation in order to be fully functional.

The most common framework used to distinguish the two dimensions of normative competence (cognition and volition) is the "reasons-responsiveness" model (Brink and Neilkin, 2011). In this model, proposed by Fischer and Ravizza, the cognitive and volitional aspects of reasons-responsiveness are separated from one another in terms of

“reasons-receptivity” and “reasons-reactivity” respectively (1998). A good way to understand these two qualities is to apply them to a machine.

There are three critical processes that must occur for a machine to do its job. A device must be able to 1) properly receive input from the outside world relative to the task at hand 2) process this information in a logical fashion that properly reflects the information contained in the input and 3) be able to translate its correct interpretation of the input into action that is in accordance with the information gathered (Brink and Nelkin, 2011). For example, let’s use a phone as our prototype machine. In order to send a text, phone must have a functioning keyboard (receive relevant input), software that maintains the proper order of the message and packages it properly (process the input properly), and the ability to send this message to the correct address in the desired format (act in accordance with proper interpretation of information). A break or malfunction in any one of these steps would inhibit the phone from accomplishing the task of sending a text properly.

The reasons-receptivity component is a mixture of components one and two in the example above. Both the receiving and processing capacities of a mechanism are critical to its overall receptivity (Brink and Nelkin, 2011). Reasons-reactivity is analogous to the ability of a mechanism to act in accordance with its understanding of the information gathered and interpreted. If we recognize that normative competence is a key component of responsibility and understand the brain as a system that requires specific capacities to accomplish normative competence, then we can see how all three abilities (receiving input, processing input, acting on input) are integrated and important when assessing responsibility (Brink and Nelkin, 2011). Since an error in any dimension of reasons-responsiveness can individually compromise normative competence, it can consequently impact an agent’s responsibility and thus their right to the sick role.

What is important to note in this model of normative competence is that it allows for a spectrum of competency. When assessing the functionality of a person’s cognitive and volitional capacities, we are essentially checking how well the individual components that contribute to competency are working in order to make judgments about a person’s normative competence. This then raises important questions as to how functional each of the mechanisms involved in cognitive and volitional capacities need to be in order for a person to be considered normatively competent.

When determining if/how much damage to the mental and/or physical mechanisms involved in cognitive and volitional capacities impact normative competence and thus responsibility, three different dimensions of the broken mechanism need to be assessed. These components are 1) the location of the mechanism (in relationship to the action), 2) the scope of the problem (what mechanisms are impacted and to what extent), and 3) the duration of time that the problem did or has existed (Brink and Nelkin, 2011). All of these aspects provide us with unique facets of information that can impact the ultimate conclusion about an individual's level of responsibility and thus how appropriate it would be for him to adopt the sick role.

For example, say that you are walking your dog and a man walking past you kicks your dog in the side. If this man was normatively competent, then you would rightfully hold him completely responsible for his action. However, suppose that this man suffers from a condition that weakens mechanisms involved in specific scenarios that require normative competence. Is he still responsible for kicking your dog or should he be allowed to adopt the sick role? It depends. If the mental mechanism that was causally responsible for the decision to kick your dog is broken at the time of the action in such a way that makes it completely dysfunctional (i.e., it is not reasons-responsive), it seems clear that he is certainly not entirely responsible for kicking your dog or even partially responsible. The malfunctioning mechanism is directly related to the action, the scope of the problem is severe, and the damage was present at the time the man kicked your dog, drastically compromising the man's normative competence and thus eliminating his responsibility for the act.<sup>3</sup>

Similarly, the presence of any malfunctioning mechanism doesn't reduce responsibility for all actions. If the mechanism is not involved in the decision or action pathway to kick the dog, then the man should be responsible for kicking your dog. Because the malfunction does not impact his normative capacity in respect to kicking your dog, he is responsible for the action of kicking the dog even though he has a malfunction that may implicate his normative capacity in respect to other actions. If a student's foot is broken, he can still write and should be expected to write because the site of the damage (the foot) doesn't affect the ability of the student to use his hand to write.

When applied to mental illness and the use of the sick role, two important facts become clear. One, patients can be held responsible for prior actions even if the mechanisms involved in those actions are now damaged by their condition. Additionally, a patient is still responsible for actions that are not impacted by their condition.

What if the broken mechanism is a part of the decision or action pathway to kick your dog, but the mechanism is only partially broken or altered? In this scenario it is harder to assign absolute responsibility or lack of responsibility since the mechanism is not completely implicated in the damage. Here it seems best to say that the man who kicked your dog is partially responsible to the degree of functionality the mechanism still possesses (Brink and Nelkin, 2011). If a student broke his pinky on his dominant hand, we may still expect him to be able to write but we would likely have different standards and expectations about the volume and penmanship the student can achieve. But, if his thumb was broken versus his pinky, we may further decrease our expectations of what the student is capable of writing as the thumb is a more crucial piece of the hand involved in the ability to write--that is, analogically, a more integral part of the mechanism involved in the decision/action pathway.

Overall, the cases above demonstrate how each dimension of a mechanism (duration, location, and scope) impacts the overall mechanistic function in relationship to the action in different ways (Brink and Nelkin, 2011). Thus when trying to determine responsibility, it is important to measure all three aspects of the malfunctioning mechanism in question.

Questions about mechanisms in relationship to responsibility are important because they provide important considerations in the ethical decision-making process of whether or not a person ought to be able to adopt the sick role. Because mental illness labels prejudice our assumptions about a person's autonomy and responsibility in regards to certain actions, having a better conception of the degree of responsibility an individual ought to have gives us better insight into how appropriate the sick role is and to what extent an individual should be able to adopt this role (Edwards, 2009).

## **Factor 2: Fair opportunity**

So far I've described the internal cognitive and volitional capacities required for an actor to have normative competence and consequently responsibility. However, the social context within which a person acts must also be evaluated as external conditions can influence the capacities necessary for responsibility. This consideration is the second component of responsibility; fair opportunity. Fair opportunity embodies many of the social factors and circumstances captured in the biopsychosocial model, and it tells us important information about to what degree an individual had the ability avoid and develop certain undesirable conditions like addiction.

A guiding principle underlying responsibility is "the fair opportunity to avoid wrongdoing" or an individual's ability to act in accordance with their intentions given the external/internal circumstances under which they acted (Brink and Nelkin, 2011). In order to have a fair opportunity to "avoid wrongdoing, it must be true that when one commits wrong, one could have done otherwise" (Brink and Nelkin, 2011). The "ability to do otherwise" is made up of various cognitive and volitional capacities that are not impaired in any significant relevant way at the time of action (Brink and Nelkin, 2011). Thus specific "impairments of cognitive and volitional capacities and specific kinds of external" threats to our ability to conform to relevant norms undermine the fair opportunity to avoid wrongdoing (Brink and Nelkin, 2011).

Because a person's actions and decisions are influenced by internal and external capacities and conditions, responsibility therefore is most appropriately thought of as a combination of normative competence and fair opportunity. For example, if a woman with full normative competence hit her child while in the park, we would hold her responsible for this immoral action. However, if we later discovered that the woman was held at gunpoint by a man who told her that if she didn't hit her child, he would shoot the child, we likely wouldn't hold her responsible for hitting her child though she was the actor who did the hitting. Under this extreme environmental pressure and threat, the woman was not in a position to properly respond in a manner that matched her normative competencies, and thus she was not given a fair opportunity to avoid the immoral action.

It can be unreasonable to hold a person accountable for actions beyond their control and for external conditions that mitigate cognitive or volitional capacities required for normative competence. Situational control and normative competence are therefore independent and "individually necessary and jointly sufficient" factors in responsibility

(Brink and Nelkin, 2011). There are degrees of each that are required in responsibility and “falling short in either dimension is excusing” (Brink and Nelkin, 2011). Using this framework of responsibility, each component is a scalar quantity that is context dependent and assessed separately for individual cases.

### **Translating responsibility and normative analysis into clinical actions**

Thus far I have argued that a normative framework for recognizing psychological dysfunction is more appropriate than other popular accounts used in clinical practice like the BPS model. When we determine that a person is mentally ill, we recognize their right to adopt the sick role that affects their responsibility and autonomy (Edwards, 2009). These features are “qualities of personhood” that dictate our assumptions about a person’s rights, duties, and ability to interact with society. Thus our decisions about when a person ought to be able to adopt the sick role inform societal assumptions about that individual’s agency as well as broader social rules and standards about what we believe individuals can and should be accountable for (Edwards, 2009).

In the second chapter, I expanded on existing ethical accounts by fleshing out and outlining the capacities necessary for responsibility as this information allows us to make more informed conclusions about when a condition ought to warrant an agent less responsible for their actions (i.e., when a person ought to be able to adopt the sick role). My framework combines normative competence and fair opportunity as necessary and jointly sufficient criteria for responsibility and creates space for different degrees of accountability based on the functionality of different mechanisms involved in these capacities.

By incorporating and creating a spectrum of responsibility into an ethical account of mental illness, I created space for the possibility of varying degrees of the mental illness label. This conception of the sick role as a normative label that lies on a gradient is distinct from current ethical frameworks that contain a more binary account of the sick role as a label that either is or is not appropriate (Edwards, 2009).

It should be noted that this spectrum model of psychiatric disorders may not be a good fit for some conditions. Some illness continuums may still be best represented as binary or categorical. For example, schizophrenia is a well-defined condition with distinct, universally observed symptoms. For this condition, it makes sense to diagnose a person as either schizophrenic or not schizophrenic versus partially schizophrenic because the disease has clear, severe symptoms and impacts a broad scope of capacities and abilities. Contrarily, for a condition like anxiety, it may be more important to determine to what degree an individual suffers from the condition as the sick label has dramatic impacts on the individual’s assumed autonomy as well as the treatment options given to the individual. Importantly, this label also influences how the individuals see themselves in relationship to the disease and their identity (Charland, 2004). Specifying the degree of anxiety leads to better conclusions about the person’s capacities and accountability, leading to better treatment as well as nurturing a more appropriate self-concept.

In this final section, I apply my crafted requirements for responsibility along with a normative conception of mental illness to case studies of psychiatric dysfunction contained within the DSM-5. I highlight how differential conclusions about responsibility and the appropriateness of the sick label can occur even when the conditions are considered equal under current DSM-5 criteria. From this marriage between the normative sick label and a fleshed-out conception of responsibility arises a new and potentially more effective approach for diagnosing and treating mental illness in clinical practice that requires a deeper assessment of individual history and context.

### **Chapter 2 Endnotes**

3) The presence of a malfunctioning mechanism in the present does not mitigate responsibility for prior actions. If you learn that the man's action/decision pathway broke a week after he kicked your dog, we would still hold him accountable for this action as the current state of the mechanism does not implicate his responsibility for prior actions that occurred when the mechanism wasn't damaged.

## Chapter 3: Clinical Application

### **Case Study: Alcohol use disorder**

The DSM-5 states that alcohol use disorder is a detrimental pattern of alcohol use occurring over a year that causes “clinically significant impairment or distress as manifested” when two or more of the following symptoms are present (APA, 2013).

- Alcohol used in larger amounts or over a longer period of time than intended
- Persistent desire or unsuccessful attempts to cut down or control alcohol use
- Significant time spent obtaining, using, and recovering from the effects of alcohol
- Craving to use alcohol
- Recurrent alcohol use leading to failure to fulfil major role obligations at work, school, or home
- Recurrent use of alcohol, despite having persistent or recurring social or interpersonal problems caused or worsened by alcohol
- Recurrent alcohol use despite having persistent or recurring physical or psychological problems caused or worsened by alcohol
- Giving up or missing important social, occupational, or recreational activities due to alcohol use
- Recurrent alcohol use in hazardous situations
- Tolerance: markedly increased amounts of alcohol are needed to achieve intoxication or the desired effect, or continued use of the same amount of alcohol achieves a markedly diminished effect
- Withdrawal: there is the characteristic alcohol withdrawal syndrome, or alcohol is taken to relieve or avoid withdrawal symptoms (APA, 2013)

The severity of the condition as mild, moderate, or severe is dictated by the number of symptoms observed and early remission is achieved when none of the criteria are experienced (except craving) for at least three months (APA, 2013).

Alcohol use disorder is one of the more controversial diagnoses within the DSM-5 as different theories about what contributes to the condition have differential consequences on how much accountability the individual is assumed to have for their actions. On one end of the spectrum lies the moral model of addiction which views “drug use as a choice” and addicts as people with “bad character” (Pickard, 2017). Under this model, addicts are seen as fully responsible for their dependence because they are assumed to have a fair opportunity to not use drugs, but *choose* to continually use them in pursuit of pleasure no matter the consequences to themselves or others. Their choice to start using and to continue using reflects deficiencies in their personality that are socially condemned, and addicts stigmatised as “bad” (Pickard, 2017). The moral model of addiction is the predominant framework present in most societies: Alcohol dependence and other forms of drug use are cross-culturally severely stigmatised with studies showing that social disapproval of addiction is higher than social disapproval of “leprosy, HIV positive status,

homelessness, dirtiness, neglect of children, and a criminal record for burglary” (Pickard, 2017). Individuals who believe in the moral model of addiction think that it is right to blame addicts for their behavior because their deviant actions are voluntary and substance abuse is best prevented and discouraged in society when viewed as a choice.

Critics of this framework assert that blaming addicts for their addiction is not the best way to motivate them to stop using and worry that cultural stigma associated with addiction creates detriments to self-identity that prevent addicts from seeking out the resources they need to curb their addiction (Pickard, 2017). Thus, shaming addicts for their behavior or holding them responsible for their addiction is not the proper pragmatic response. Instead, they define addiction as a disease and drug use as involuntary. Addiction is “a chronic, relapsing neurobiological disease” and addicts “literally cannot help using drugs and have no choice over consumption” despite personal and social consequences (Pickard, 2017). Under this framework, addicts are justified in adopting the sick role as “helpless victims” to their addiction (Edwards, 2009).

Somewhere in between the moral and disease model lies the learning model of addiction. The creator of this framework, Marc Lewis, sees both the moral and disease model of addiction as problematic (Pickard, 2017). The moral model allows society to shame addicts as lazy, self-indulgent, and irresponsible, but the disease model is not better as it “wrongly pathologized both the brain and the person” (Pickard, 2017). Lewis asserts that changes in the brain that result from addiction are a sign of “neuroplasticity” rather than pathology and are completely normal responses to continued drug use that are akin to other kinds of learning and habit formation (Pickard, 2017). He also does not believe that the disease model is the more pragmatic approach because overcoming addiction requires “a sense of agency and empowerment” that is not possible if addicts are viewed as incapable of controlling their behaviors without some kind of “cure” (Pickard, 2017). Research supports this view as addicts who believe that they are dependent on professional help or other medical remedies to curb their behavior (i.e. subscribe to the disease model of addiction) are more likely to relapse 6 months post-treatment (Pickard, 2017).

However, demographic patterns of addicts suggest that addiction is not entirely random and that certain social factors and environmental exposures increase the risk of developing addictive behaviors because they impact what choices are available (Pickard, 2017). People can have more or less options “genuinely available to them, and more or less capacity for control” (Pickard, 2017). This fits nicely with the concepts of volitional capacity and fair opportunity described in my previous chapter about responsibility. Ultimately, differences in addicts’ abilities to control their behavior (differences in volitional capacity/fair opportunity) lead to different degrees of responsibility assigned to them for developing and maintaining their addiction.

Edwards in his normative analysis of mental illness touches on alcohol addiction and offers the following thoughts:

if the sufferer is completely unable to seek treatment without outside assistance, it may be too unfair to apply moral blame for a situation that the sufferer cannot change, and that would give good reason to classify alcoholism as a mental illness. However, if the barrier to seeking treatment is less than a total barrier, then there are strong considerations both in favor of applying the label and in favor of withholding it. One would need to determine which of the interests represented by the criteria were more important—for example, inculcating values through the expression of moral blame, or protecting sufferers from harm by declaring the condition an illness rather than a legitimate part of their persona. [Edwards, 2009]

In this excerpt, we see how the ethical decision to legitimize the sick role is largely dependent on the context and the specifics of individual cases. In his work, Edwards lays out a series of ethical considerations to help clinicians assess when, and why, to diagnose a patient as having one or more mental illnesses.

- a. Is the condition harmful to the person who has it?
- b. Is there any reason for legitimizing the condition as a character trait that one can choose to develop or maintain?
- c. Is the condition one that can be discouraged through the inoculation of appropriate moral values during childhood?
- d. Will applying moral responsibility to the condition uphold broader moral values in one's ethical system
- e. can one have insight into the condition's effect upon oneself and if so, how difficult is it to take an active role in seeking treatment for oneself? [Edwards, 2009, 83-84].

In this list, we see that appropriateness of the mental illness label is a balance between individual and broader societal considerations. It's important to emphasize the need for case-by-case analysis when assigning responsibility since responsibility lies on a spectrum and is dependent on internal capacities in conjunction with external factors that may impact an individual's fair opportunity. Since the particulars of a patient's history are critical to how we assign the sick label and responsibility, let's spell out a specific case.

Mr. J is a 42 year old man who has been drinking at least six drinks per day for the last four years though he continually tries to stop his compulsive behavior. He started drinking more after his father died of pancreatic cancer, but over time this coping strategy has transformed into chronic behavior. His maladaptive behavior has interfered with his ability to hold down a job and tend to his family, and therefore causes him significant distress. Recently, Mr. J's wife decided that she can no longer cope. Frustrated by his continued use, inability to keep a job, and worried about the influence his drinking is having on their children, she has chosen to file for a divorce.

Using DSM-5 criteria, Mr. J would be diagnosed with alcohol use disorder. He has tried, but is unable, to curb his drinking. His behavior causes him significant distress, he has sustained this pattern for over a year, and it interferes with his ability to function in social contexts. Given his case, what do we believe about the sick role and Mr. J's responsibility?

Our instinctive interpretation of different aspects of Mr. J's case will give us differential answers about how to respond and usually reflect our biases towards the moral or disease model of addiction. We can use some of Edward's questions to help guide our thinking about Mr. J's responsibility. For example, is Mr. J's condition harmful to him? Given Mr. J's distress and interpersonal tension his condition is causing, it seems clear that the condition is detrimental to his personal and social well-being. Is there any reason to legitimize Mr. J's dependence as a character trait that he can choose to develop or maintain? Here we are asking whether when Mr. J began drinking after his father died, do we believe that he had a fair opportunity to try healthier coping mechanisms but instead resorted to drinking. Using the moral model of addiction, we would blame Mr. J for his addiction with the assumption that his drinking is something he can choose to do, and therefore, Mr. J continues to drink out of lack of self-control. However, if we believe that a combination of the moral and learning models of addiction is accurate, we would assume that Mr. J felt that he had no other options after his father died except to try and numb the pain with alcohol and unfortunately developed a dependence to it over time that now makes him incapable of voluntarily stopping.

These two views imply different things about whether Mr. J's drinking reflects faults in personhood (traits/behaviors that we can punish people for adopting) and his responsibility for developing and maintaining his addiction overtime. In both views, it does seem clear that we recognize that the man did make the initial decision to start drinking more when his father died. Therefore causal responsibility is not different between the two accounts. What varies is the *blame* placed on the man for starting to drink.

However, learning new information about Mr. J's history can change our assumptions about responsibility by influencing our answers to Edwardian criteria . Below, I spell out different backgrounds for Mr. J to demonstrate how different ethical, social, and environmental, factors can impact our analysis about responsibility even when the drinking behavior and consequences are the same.

Biography A: Mr. J grew up in a middle class neighborhood with strong social networks within his household, school, and community. His parents made clear efforts to teach Mr. J important social values, and his father had a friend in college who died in a drunk driving accident making him particularly sensitive and explicit about the importance of responsible drinking behaviors. No one in Mr. J's immediate family drank in excess and Mr. J has continued to have familial financial and emotional support throughout his lifetime.

Biography B: Mr. J grew up in a poor neighborhood with high rates of crime, violence, and drug consumption. Due to work constraints, Mr. J's parents weren't around very much while he was growing up and he spent most of his time in the community. When Mr. J was

16, his father got laid off from his job. His dad was ashamed of his inability to provide for his family and began drinking in an attempt to dull his emotions. Mr. J's father never developed alcohol dependence disorder, but he did regularly drink in excess during the seven months he was unemployed. As an adult, Mr. J has had little contact with his immediate family and doesn't feel closely connected to anyone except his wife and children.

From these backgrounds, we learn important information about Mr. J's upbringing that likely influence his access to social resources and ability to choose appropriate coping mechanisms when faced with stress, or the fair opportunity component of responsibility. Youth and adolescence are critical periods during development that influence long term beliefs about social norms and appropriate behavior. In both biography A and B, Mr. J is exposed to actions and attitudes towards drinking that may either enhance or reduce his predisposition towards adopting unhealthy drinking behaviors. Given these two backgrounds, we can see that our answers to Edward's questions b-e spelled out above, may be different depending on whether biography A or B is true (2009).

In scenario A, Mr. J is given strong moral guidance during development with an acute focus on appropriate drinking and the dangers of overconsumption. He also receives social support at early and later time points during his life with ties to his family. Mr. J is provided with the knowledge and resources to recognize and adopt healthier coping mechanisms when faced with stress and therefore he has a fair opportunity to choose appropriate actions. In Edwardian terms, Mr. J's drinking is a result of him choosing to drink as he had proper exposure to moral values discouraging him to drink through childhood (Edwards, 2009). For these reasons, Mr. J has a high degree of responsibility in his initial decision to start increasing his drinking that leads to his chronic consumption.

In scenario B, Mr. J has limited social support throughout his life and is exposed to several risk factors within his household and community that likely increase his odds of developing unhealthy coping behaviors. In particular, his father's drinking in response to stress strengthen associations in Mr. J's mind between alcohol and difficult scenarios. Using Edward's terms, we may believe that although Mr. J's drinking could have been discouraged during childhood with "the inoculation of appropriate moral values," in order to reflect broader social values, it's best not to hold him responsible since he didn't have a chance to learn any other ways of coping during childhood and has never had social support (Edwards, 2009). For these reasons, Mr. J didn't lacked a fair opportunity to not start drinking in response to stress. These points all demonstrate that if background B is true, we may have reasons to believe that Mr. J didn't have the resources and information available to him to respond to his father's death in a healthier way and thus has a diminished degree of responsibility for adopting detrimental drinking behaviors.

When comparing Mr. J's backgrounds and consequent responsibility for starting to drink as a response to stress, it's clear that social resources and exposures influence his access and ability to choose healthier outlets and alternatives. This exercise emphasizes two things. One, that two individuals who receive the same diagnosis under DSM-5 criteria can have differential responsibility for their condition, and two, because at some point in time there

is an initial decision to engage in the behavior, it is possible to be responsible for developing a condition though at the time of diagnosis you may no longer be able to control your actions.

## **Responsibility and Blame**

In the case study above, Mr. J is assigned different levels of responsibility for his alcohol dependence depending on certain social and environmental factors and influences. However, it is critical to note that he is deemed responsible at some level for his actions. In previous chapters I discussed how moralized attitudes like blame and gratitude are only appropriate emotions when an actor is a proper candidate to assume responsibility (Brink and Nelkin, 2011). Given this, the logic then follows that since Mr. J is responsible, he is also a justified target for blame. But, in the clinical context concerned with both appropriately diagnosing conditions as well as treating the patient, I want to argue for an alternative model; choice without blame.

In this model crafted by Hannah Pickard, individuals can still be responsible for adopting certain behaviors and making decisions that ultimately can lead to different kinds of disorders. She believes that there are therapeutic reasons “for legitimizing the condition as a character trait that one can choose to develop or maintain” (Edwards, 2009). Namely that by applying responsibility, patients are given a sense of accountability and control over their actions that is not present when they are assumed to be passive victims to their condition (i.e. adopt the sick role) (Pickard, 2017). Even if the addict no longer has “insight into the condition’s effect upon” himself, cannot “take an active role in seeking treatment for himself,” and never had a fair opportunity to learn healthier ways of coping through “inoculation of appropriate moral values during childhood,” Pickard would argue that clinicians can and should still apply some level of responsibility for the initial decision to start drinking (Edwards, 2009). However, though the individuals are responsible, they are not blamed for their behavior (Pickard, 2017). Going back to ideas presented in chapter one, this means that an actor can be attributable responsible for an action without being morally responsible for an action. Given that moral responsibility is a subcategory of attributable responsibility, this is reality is achievable.

This framework has clear pragmatic, forward-looking benefits. Allowing patients to feel empowered to change their condition without fear of being emotionally punished or stigmatized for their initial poor choices is the best pragmatic response. It marries the moral and disease model of condition by “acknowledging the role of choice in addiction” and “mobilizing a sense of agency and empowerment... through acknowledging and working with their agency without adopting moralising attitudes or stigmatised attitudes and practices” (Pickard, 2017). This model entails sustained work and collaboration between providers and patients in recognizing and treating addiction that is best fostered through comprehensive analysis, honest dialogue, and compassion.

To demonstrate how this model would work in action, let’s apply it to the case of Mr. J. No matter if Mr. J had biography A or B, a clinician when discussing Mr. J’s condition with him

would acknowledge that Mr. J chose to start drinking more in response to his father's death and is responsible for this initial action. When doing this, the physician would make clear that Mr. J's behavior didn't reveal deeper insight into his moral character. The doctor would assure Mr. J that he is not a bad person and explicitly indicate to Mr. J and his loved ones that Mr. J shouldn't be blamed or stigmatized for his past actions. However, the physician would emphasize that Mr. J has the underlying capacity to control his condition. Though in the immediate situation Mr. J may need more outside assistance or medical intervention to help put him back into a position where he can exert full psychological and physical control, his ability to do so remains constant throughout his treatment. Ultimately, the physician would tell Mr. J that he can change his condition and that the success of his recovery is predicated on his desire and will to do so. By holding Mr. J accountable without blaming or shaming him, the clinician empowers Mr. J to overcome his addiction through compassion and without guilt or fear that he is morally inadequate.

### **Conclusion**

Psychiatric dysfunctions are conditions that have been historically identified in relationship to social norms for appropriate thoughts and behaviors, making mental illness categories both temporally and geographically relative. Different frameworks for consistently recognizing and defining what conditions constitute mental illness have been often ignored in the subjective component of psychiatric disorders in an attempt to minimize biases when defining mental illness. However, as discussed above, purely objective criteria are insufficient in capturing psychological dysfunction because normative considerations are embedded within the mental illness label. Because of these ethical implications, I argue that we should embrace and explore a more normative conception of psychiatric dysfunction and create a model of responsibility to help clarify when a condition should be labeled as mental illness. Ultimately, responsibility lies on a spectrum and thus the appropriateness of the sick role also comes in different degrees.

In clinical contexts, detailed and comprehensive case-by-case analysis is necessary to come to proper conclusions about responsibility and consequently of the appropriateness of the mental illness label as specifics are critical to this decision. Though not appropriate for all kinds of mental illness, in cases of dysfunction where there is potential for agency and responsibility, these two qualities should be recognized without blame. By acknowledging the capacity of patients to act on their own volition to change their condition and behaviors, clinicians better empower them to participate in their treatment and reintegrate with society. This approach and framework of choice without blame is a forward-looking model that offers the best pragmatic response for allowing patients to feel capable of fully recovering and changing their behaviors. Providers who treat and diagnose mental illness should be trained to approach patients with compassion and discuss conditions and treatment without stigmatization or implying that the patient has faults in personhood.

Although compassionate care may be the current standard in clinical practice, I suggest several radical changes to the way clinicians understand mental illness. I argue that physicians should abandon the traditional framework for understanding psychiatric

disease (i.e. the BPS model) as it is imprecise and fails to acknowledge the ethical implications of the mental illness label. When physicians fail to recognize that psychiatric disease is value-laden, they may inappropriately encourage patients to adopt the sick-role, which can be detrimental to their patients' perceived sense of self-efficacy and may weaken forward-looking goals therapy that includes patient empowerment and independence from the condition. It should be noted that in applications of the sick role, the ethical considerations used are informed by my account of responsibility but not decided by it. Though patients might not be entirely responsible for developing their condition, in some cases there may be pragmatic benefits to treating them this way so that they can become responsible. This approach is forward-looking and focused on long term recovery.

Try as we might, we cannot yet fully reduce therapeutic intervention to a purely scientific or technical process. The ambiguity and irreducibility of human emotional and psychological issues mandate an experiential and empathetic element in treatment. This necessity does not deny that there are biological causes for mental or emotional illness. Instead, it asserts that the manifestation of the problem may require both medical and social, interpersonal interventions. Neither approach is always sufficient by itself to obtain full wellness. In advocating for a more normative framework of mental illness, I am not suggesting that interpersonal approaches are more effective than more technical, drug-based treatments. However, what I am claiming is that case-by-case analysis is critical and that the consideration of individual factors captured through a more social, interactive approach is critical to the current diagnosis and treatment of mental illness.

Physicians should be mindful of the normative implications of the mental illness label and then decide on an individual basis whether the sick role is appropriate based on the patient's history, the type of psychiatric disease, and the patient's potential for recovery. Therapy must be specific to the patient as identical mental illness categories can be caused by a variety of factors and thus are best treated using different approaches and mixed-methodologies.

## Bibliography

American Psychiatric Association (APA). *Diagnostic And Statistical Manual of Mental Disorders: DSM-5*. Arlington, VA: American Psychiatric Publishing, 2013, Print.

Brink, David and Nelkin, Dana. "Fairness and the Architecture of Responsibility" 2011, pp. 1-22

Callicott, J. Baird. "The conceptual foundations of the land ethic." In Craig Hanks (ed.), *Technology and Values: Essential Readings*. Wiley-Blackwell. 2010.

Charland, Louis. "A Madness for Identity: Psychiatric Labels, Consumer Autonomy, and the Perils of the Internet." *Philosophy, Psychiatry, & Psychology*, vol. 11, no. 4, 2004, pp. 335-348.

Edwards, Craig. "Ethical Decisions in the Classification of Mental Conditions as Mental Illness." *Philosophy, Psychiatry, & Psychology*, vol. 16, no. 1, 2009, pp. 73-90.

Fischer, J.M. & Ravizza, M. *Responsibility and Control A Theory of Moral Responsibility*, 1998.

Ghaemi, Nassir. *Chapter 1: The Status Quo: Dogmatism, the Biopsychosocial Model, and Alternatives. The Concepts of Psychiatry*, Johns Hopkins University, 2003, pp. 1-23

Greenspan, Patricia. "Responsible Psychopaths." *Philosophical Psychology*, vol. 16, no. 3, 2003, pp. 417-427., doi:10.1080/0951508032000121797.

Horwitz, Allan V. *Chapter Three: The Emergence of Diagnostic Psychiatry. Creating Mental Illness*, University of Chicago, 2007, pp. 56-83.

Nelson, L. & Ramirez, E. 2017. Can suicide in the elderly be rational? in R. McCue & M. Balasubramaniam (Eds.) *Rational suicide in the elderly: Clinical, ethical, and sociocultural aspects*. 2017, pp. 1-21. Springer

Pickard, H. "Responsibility without Blame for Addiction." *Neuroethics*, 10, 169-180. 2017. doi:10.1007/s12152-016-9295-2

Wakefield, Jerome. "The Concept of Mental Disorder." *American Psychologist*, vol.47, no. 3, 1992, pp. 373-388

Wakefield. Jerome. "Spandrels, Vestigial Organs, and Such: Reply to Murphy and Woolfolk's 'The Harmful Dysfunction Analysis of Mental Disorder'" *Philosophy, Psychiatry, & Psychology*, vol. 7, no. 4, 2000, pp. 253-269.