THE SKILL OF MENTAL HEALTH

by

GARSON TITEFF LEDER

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(PHILOSOPHY)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2018

Abstract

Psychotherapy is effective. Since the 1970's, meta-analyses have consistently shown a significant effect size for psychotherapeutic interventions when compared to no treatment or placebo treatments. This effectiveness is normally taken as a sign of the scientific legitimization of clinical psychotherapy. A significant problem, however, is that most psychotherapies appear to be equally effective. This poses a problem for specific psychotherapies: they may work, but likely not for the reasons that ground their theoretical explanations for their effectiveness. This dissertation explains the common efficacy of psychotherapies by developing novel skill-based account of mental illness and healing. According to the view defended here, mental illness, and the success of mental healing, is best explained as an issue of the breakdown and development of skilled action. This skill view of mental health attempts to resolve a number of long-standing metaphysical questions about the roles of biological dysfunction, the environment, and values in the conception of mental disorder.

Lay Summary

This dissertation addresses two related questions: (1) What, if anything, is the difference between mental disorders and merely disvalued or norm-transgressing states? And (2), what explains the appearance of common efficacy of psychotherapy? This dissertation provides a skill-based solution to both.

Preface

This dissertation consists of three independent research articles (Chapters 2-4), a support chapter (Chapter 5), an introduction (Chapter 1), and a brief conclusion (Chapter 6). Chapters 2-4 are written as stand-alone articles and may contain some overlap in content. This dissertation consists of original and independent research by the author, G. Leder.

Chapter 2 contains material that was previously published: Leder, G. (2017). Know Thyself? Questioning the theoretical foundations of cognitive behavioral therapy. *Review of Philosophy and Psychology 8*(2), 391-410.

Table of Contents

Abstract ..... **ii**

Lay Summary ..... **iii**

Preface ..... **iv**

Table of Contents ..... **v**

Chapter 1: Introduction 1

1.1 Mental Health and Disorder 1

1.2 Naturalist Views 2

1.3 Hybrid Views 10

1.4 Normative Views 17

1.5 The Skill View of Mental Health 24

1.6 Overview 24

Chapter 2: Know Thyself? Questioning the Theoretical Foundations of Cognitive Behavioral Therapy. 27

2.1 Introduction 27

2.2 The CBT Model 28

2.3 Criticisms of the CBT Model 33

2.4 Thoughts, Beliefs, and Confabulation 36

2.5 CBT Controls for Introspective Accuracy 45

2.6 Conclusions 51

Chapter 3: What Does It Mean To Have A Meaning Problem? Meaning, Skill, and the Mechanisms of Change in Psychotherapy 55

3.1 Introduction 55
3.2 The Common Factors Theories 58
3.3 The Problems with Meaning 68
3.4 Alternative Hypothesis: Skill and Psychological Healing 81
3.5 The Skill Hypothesis and the Common Factors 85
3.6 Conclusion 87
Chapter 4: The Skill of Mental Health 88
4.1 Introduction 88
4.2 Naturalism and Mental Functions 90
4.3 The Skill of Mental Health 98
4.4 Pragmatics and the Skill View 119
4.5 Addendum 123
Chapter 5: The Metaphysics of Skill 126
5.1 Introduction 126
5.2 Skill 126
5.3 Intellectualism 127
5.4 Anti-Intellectualism. 145
Chapter 6: Conclusion 151
6.1 Conclusions 151
References 153

Chapter 1: Introduction

1.1 Mental Health and Disorder

While it is generally agreed that mental disorders are a serious problem, there is far less agreement about what the problem actually is. What exactly is *disordered* about mental disorders?[1] The answer is normally situated somewhere between the naturalist/normative divide. Naturalism, roughly, is the position that mental disorders are value-independent, and in principle scientifically identifiable and explainable, mental phenomena. Normative theories of mental health conceive of mental disorder as primarily a problem of the meanings attached to behavior and bodies; values, rather than natural facts, determine whether some way of being is disordered. Hybrid theories land somewhere between these poles. Naturalism, or hybrid theories with a necessarily naturalistic component, is the received view in academic psychology and psychiatry. Despite its wide acceptance, there have been significant problems in formulating the naturalist view. Notably, it has proven difficult to precisify what, exactly, the objective basis of mental illnesses is supposed to be.

The solution to this problem matters quite a bit. Many ways of being are, and have been, *labeled* as disordered. These include generally well-accepted disorders (such as depression, anxiety, bipolar disorder, and schizophrenia), controversial 'disorders' (such as sex addiction, asexuality, gender dysphoria, and oppositional defiant disorder), as well as diagnoses now generally considered to be mistaken (such as 'female hysteria', homosexuality, and 'sluggish

[1] This dissertation uses the words 'disorder', 'disease', and 'illness' interchangeably unless otherwise noted.

schizophrenia'). This nosological uncertainty raises a general problem: what, if anything, is the difference between normal problems of living and norm-transgressing behavior, on the one hand, and genuine mental disorders, on the other? The resolution to this question comes down, in large part, to what we think mental disorders are.

Tenable, but significantly flawed, theories of mental health and disorder have been constructed at most points along the naturalist/normative spectrum. The choice between them often comes down to which bullets one is comfortable biting. Sections 1.2 - 1.4 address the strengths and weakness of the most plausible existing approaches.

1.2 Naturalist Views

The primary challenge facing naturalist views of mental disorder is to specify what, exactly, the value-independent foundation of mental disorder is. The most common answer to this challenge has been to appeal to the concept of natural function. Mental disorders, then, are supposed to require dysfunctions of some part or process of the mind.

Consider, for example, how the two primary diagnostic manuals in the West define disorder. The American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5*, 2013) defines mental disorder as:

> 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 the 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 (italics added).

The World Health Organization's tenth edition of *International Classification of Diseases* (*ICD-10*, 1992), provides a similar function-based definition of disorder (though the focus here is on the function of 'persons', not just minds):

> 'Disorder' is not an exact term, but it is used here to imply the existence of a clinically recognizable set of symptoms or behaviour associated in most cases with distress and with interference with personal *functions*. Social deviance or conflict alone, without personal *dysfunction*, should not be included in mental disorder as defined here (italics added).

According to these function-based views, the identification of a dysfunction is supposed to be what differentiates merely norm-transgressing or disvalued behavior from genuine mental disorders. However, neither the DSM-5 nor the ICD-10 offer definitions of what a 'mental

dysfunction' is. This is a problem, given that any number of ways of being can by hypothesized to be 'dysfunctional'. Naturalist theories of mental health are meant to fill this lacuna.[2]

1.2.1: Naturalism Without Function

Note that while the naturalist attempt to identify the value-independent basis of mental illness is normally articulated in terms of proper function, this needn't be the case. Most notably, Scadding (1967) and Kendell (1975), advocate non-function based 'biological disadvantage' naturalist theories of disorder. According to Scadding (1967):

> A disease is the sum of the abnormal phenomena displayed by a group of living organisms in association with a specified common characteristic or set of characteristics by which they differ from the norm for their species in such a way as to place them at a biological disadvantage (p. 877).

So, according to this view, a disease is a statistically significant abnormality that (somehow) leaves one biologically worse off. The obvious problem here is vagueness. Scadding does not specify what a biological disadvantage *is*. We are told that:

[2] This does not imply that all naturalist theories must adhere to the DSM-5 or ICD-10 definition of disorder.

> The rather vague term "biological disadvantage" is as precise a statement as can be made about the criteria on which it is generally decided whether deviation from norm is to be regarded as associated with disease or not (p. 877).

Kendell (1975) precisifies this view by defining biological disadvantage as 'increased mortality and reduced fertility' (p. 310).[3] The biological disadvantage view, then, is supposed to give us a scientific foundation for the study and treatment of mental disorders; disorders are dysfunctions that lead to a reduction of fertility or an increase in mortality.[4] According to Kendell, Scadding's amended biological disadvantage concept is 'immune to idiosyncratic personal judgments of patients and doctors' (1975, p. 309). The concepts of health and disorder, then, are hypothesized to be scientifically grounded biological concepts, not mere reflections of values.

The problem with these views is that a reduction in fertility and longevity does not appear to be sufficient for disorder (Fulford, 1989; Wakefield, 1992; Boorse, 1998). As Cooper (2005) notes, choosing to be a mercenary or to engage in dangerous thrill-seeking sports may reduce one's life-expectancy, while choosing to not have children will definitely reduce one's fertility (as opposed to fecundity). None of these examples are plausible instances of mental disorder.

[3] Note that Kendell (1975) leaves open the possibility that there may be other standards of biological disadvantage *in addition* to fertility and longevity.
[4] Though both Scadding and Kendell note that the identification of *mental* disorders will be difficult if one accepts the biological disadvantage view.

It is, however, still possible that a statistically significant reduction in fertility and longevity are *necessary* for disorder. Boorse (1976, 1977, 1997, 2014), and his followers, adopt this view.[5] According to these revised disadvantage views, biological disadvantages are only disorders if they are based on a dysfunction of some somatic or mental part or process.

1.2.1 Naturalism and Natural Functions

Boorse's 'biostatistical theory' (BST) is the most influential 'pure' (i.e., non-hybrid) naturalist view of health and disorder. According to the BST, disorders are statistically significant deviations from the normal functioning of a part or process (with 'normal functioning' defined as its species-typical contribution to the goal of individual survival and reproduction), relativized to a reference class (i.e., age-group and sex), that negatively affect an individual's survival or reproductive success. Health is defined as the absence of disease.[6] The strength of the BST, and disadvantage views in general, is that, if successful, they give us a 'pure' naturalist conception of mental health; value-judgments are supposed to play no role in whether or not some mental mechanism contributes to the longevity or reproductive success of some organism (and are thus supposed to be irrelevant to our conception of health and disorder).

There have been a number of criticisms of specific aspects of this view. However, let's assume for the sake of argument that the BST can be made coherent. The question here is whether the

[5] For Boorse-inspired views, see: Hausman, 2012, 2014; Schwartz, 2014; Kraemer, 2013; Garson & Piccinini, 2013.
[6] Boorse has recently switched his terminology from 'disease' to 'pathological condition'. This does not mark a substantive change in his theory, but rather reflects what Boorse considers to be a 'more natural choice' (2014, p. 684).

BST, even in its strongest form, offers us a good theory of mental health.

The most significant problem facing the BST (and biological disadvantage views in general) is that it leads to a possibly massively revisionary conception of mental disorder. As Boorse notes (quoting Redlich & Freedman, [1966]), we lack a 'completely acceptable supertheory on which psychiatry…can rest its work' (1976, p. 61). The problem is that there is nothing like a consensus conception of the 'proper' functioning of the human mind from which to identify when any particular part or process is failing to function 'normally'. Thus, as Boorse states, '[a]part from a theory of the structure and functions of the human mind, virtually all assertions about mental health are either missuses of language or flatly conjectural' (1976, p. 81).[7] And accuracy matters here: whether or not we view some mental state as disordered, according to function-based views, will depend upon which specific theory of mental functioning we adopt.

Theories about the structure and functions of the human mind vary significantly. It is an open scientific question whether paradigmatic mental disorders such as depression, anxiety, panic disorder, and schizophrenia are in fact the result of dysfunctioning mental mechanisms rather than being adaptations or spandrels (e.g., Lilienfeld & Marino, 1995; Woolfolk, 1999; Murphy & Woolfolk, 2000; Bolton, 2001; Cooper, 2002). Depression, for example, is variously hypothesized to be the result of *dys*-functioning mental mechanisms that lead to 'interlocked' cognitive-affective cycles of mental processing (Teasdale & Barnard, 1993), the product of

[7] Boorse (1976, 1997), while cautious, suggests that Freudian psychoanalytic theory (and specifically the concepts of id, ego, and superego) provides the most plausible non-physiological model of mental functioning. Boorse (1997, p. 14) also notes that evolutionary psychology may also eventually provide us with a plausible structural model of the mind.

normally functioning mental mechanisms that lead to an increase in critical ruminative thought (Andrews & Thomson, 2009), the result of *dys*-functioning mental mechanisms that produce negatively valenced loss-based cognitions (Beck & Alford, 2009), or the result of *normally* functioning mental processes that socially signal the need for assistance (Allen & Badcock, 2003). Similarly, generalized anxiety disorder (GAD) is postulated to be the result of a *dys*-functioning evolved-for harm and danger avoidance system (Clark & Beck, 2011), the result of a potentially *normally* functioning mind that has learned to process information in an anxiety producing manner (Greenberg, 2010), the product of *dys*-functioning subconscious mechanisms that produce conflicts between forbidden wishes and defenses against these wishes (Busch et al., 1999), or the result of a possibly adaptive mental process to combat real or perceived threats to social exclusion (Baumeister & Tice, 1990).

Similarly, the acceptance of biological disadvantage theories makes it possible that seemingly benign mental states, such as atheism, may turn out to be mental disorders. As Davis (2017) notes, a central debate in the scientific study of religion is whether religious belief is an evolutionary adaptation or a byproduct of other adaptive traits (e.g., Sosis, 2009; Haidt, 2012; Wilson, 2002; Boyer, 2003; Bloom 2007). We needn't take a position on this debate to note that *if* it is the case that religious belief may be an adaptation (and thus a function of the mind), and *if* it is the case that a lack of religious belief may cause (or has in the past caused) a decrease in individual survival and reproduction (e.g., because it disrupts in-group cohesion), then atheistic belief would end up being (or having once been) a mental disorder. The take-away point here is not that this is the most likely theory (adaptationist views are in the minority). Rather, the point is that if we accept dysfunction and biological disadvantage as the grounding of our concept of

disorder, we must accept that it is an open question whether being an atheist is (or was) a mental disorder. This is a significant bullet to bite.

The biological disadvantage theories' focus on reproductive success also ends up pathologizing a number of lifestyles that are not normally considered pathological. For example, if we accept that fertility is one of two 'apex' goals of a healthy body and mind (Boorse, 1977, p. 556), then homosexuality, asexuality, and the lack of desire to reproduce, are all very likely to be mental disorders.[8] Disadvantage theories are led to this position because they are committed to the view that mental health is ultimately a question of the adherence to the biological norms of reproduction and survival, rather than a state experienced by individuals with their own varying goals and interests. A lowered lifespan and reduced fertility in comparison to one's peers may be considered a disadvantage when looked at from the perspective of competition amongst organisms, but it is less obviously a problem when looked at from the level of the individual (who may value neither procreation nor longevity). This is supposed to be a positive aspect of disadvantage views; health is a biological concept that is distinct from questions concerning agentive values and interests. The same standard of health is used for bacteria, plants, and human minds. The problem is that while this may offer a plausible theory of health for plants and simple organisms, it is less clear that this fits what we want out of a theory of a healthy *mind*.

[8] Note that for the BST, homosexuality's status as a disorder comes down to whether it serves an identifiable function (e.g., improving the fitness of one's genetic line through kin-selection) (Boorse, 1997, 2014). Kendell's (1975) biological disadvantage theory, on the other hand, is explicit in concluding that homosexuality is a mental disorder.

Disadvantage views, thus, make it the case that paradigmatic and harmful disorders such as depression, anxiety, and phobias may turn out *not* to be mental disorders (but rather products of *healthy* minds dealing with mere problems of living), while, on the other hand, behavior that is not inherently harmful and not commonly considered to be product of unhealthy minds, such as homosexuality and the lack of desire to have children, may in fact be mental disorders. This is not incoherent, but it is highly revisionary. Adherents to disadvantage views needn't consider this a problem. Boorse, for example, seems to believe that we should let our best theory of disorder drive our intuitions about what types of phenomena should count as disordered, not vice versa. This is a reasonable view. However, if our theory of mental health produces this kind of extensional output, it gives us reason to question whether this is really the best theory.

1.3 Hybrid Views

Hybrid views are meant to resolve (at least) some of the issues facing 'pure' naturalist theories. These views posit that the concept of mental disorder is *both* value-laden and grounded in value-independent facts. The most influential hybrid view has been Wakefield's (1992) 'harmful dysfunction' theory. According to Wakefield (1992):

> The concept of disorder must include a factual component so that disorders can be distinguished from a myriad of other disvalued conditions. On the other hand, facts alone are not enough; disorder requires harm, which involves values. Thus both values and facts are involved in the concept of disorder (p. 381).

The factual component of the harmful dysfunction view is, like the BTS, the dysfunction of some mental part or process.[9] But, according to Wakefield's hybrid theory, what makes some dysfunction a *disorder*, rather than merely a functional difference, is that the person is directly harmed by the failure of some naturally selected mechanism. This hybrid view is supposed to be an improvement upon 'pure' naturalist and normative theories: the dysfunction criterion is supposed to provide a principled metaphysical grounding for the difference between genuine disorder and mere value claims, while the normative 'harm' criterion is supposed to avoid the concerns about the pathologization of 'benign' dysfunctions faced by pure dysfunction views. Thus, according to Wakefield (1992):

> A condition is a mental disorder if and only if (a) the condition causes some harm or deprivation of benefit to the person as judged by the standards of the person's culture (the value criterion), and (b) the condition results from the inability of some mental mechanism to perform its natural function, wherein a natural function is an effect that is part of the evolutionary explanation of the existence and structure of the mental mechanism (the explanatory criterion) (p. 385).

The primary challenge facing the hybrid harmful dysfunction views is that it inherits many of the flaws from its constituent parts.[10] Most notably, it faces the same extensional problems as function-based views and the same problems with value-relativity as normative views. The

[9] Note that Wakefield and Boorse adopt different theories of natural function. Boorse (1977, 1997, 2014) advocates a goal-based theory of function. Wakefield (1992, 1999) adopts an etiological evolutionary theory of function.
[10] This is not the *only* challenge facing Wakefield's harmful dysfunction view.

problem is this: if what constitutes a harm is dependent upon the standards of one' culture, and if what constitutes a dysfunction is dependent upon conjectural (even if well-reasoned) evolutionary psychology, then depending upon the culture and the theory of mental functioning, norm-transgressing behaviors that can plausibly be conjectured to be the result of a 'dysfunction' (such as gender-alterations, homosexuality, and anti-social behavior) may be rightly considered mental disorders in one culture, but not in others.[11] So, despite the supposedly objective 'dysfunction' criterion, the harmful dysfunction view allows for a significant amount of relativism in the attribution of mental disorders. This does not entail that the theory is mistaken, but it does mean that the harmful dysfunction view faces the same challenges as other normative views (discussed in section 1.4).

1.3.1 Hybrid Views Without Natural Functions

The most prominent *non*-function based hybrid theory is Clouser, Culver, & Gert's (1981, 1997) and Culver, & Gert's (1982) (referred to collectively here as CCG) 'malady' theory of disease. CCG intend their neologism 'malady' to encompass what all 'disease' terms (e.g., 'disease', 'illness', 'injury', and 'disorder') have in common. According to CCG's hybrid malady theory, the concept of 'disease' (or 'malady' in their terminology) is based on the value-laden concept of 'harm', but it is nevertheless a matter of *natural fact* whether something is in fact a harm. According to CCG:

[11] This does not entail that hybrid theorists must also accept that all value systems are equally as reasonable, fair, or moral.

> Individuals have a malady if and only if they have a condition, other than their rational belief or desires, such that they are incurring, or are at a significantly increased risk of incurring, a harm or evil (death, pain, disability, loss of freedom, or loss of pleasure) in the absence of a distinct sustaining cause (Clouser, Culver, & Gert, 1997, p. 190).

There is much to unpack here. First, the 'rational belief or desire' clause is meant to differentiate between normal harm-inducing thoughts and desires (e.g., the thought that the stock market has collapsed or the desire to engage in rough sporting activities) from their supposedly pathological counterparts (e.g., delusional patterns of thinking or the uncontrollable desire to self-harm). Second, CCG define a 'distinct sustaining cause' as 'a cause whose effects come and go simultaneously, or nearly so, with the cause's respective presence or absence' (Clouser, Culver, & Gert, 1997, p. 189). This clause is intended to exclude from the extension of 'malady' harmful things that are somehow 'external' to an individual (e.g., being in a wrestler's headlock or being in a smoke-filled room), rather than being 'internal' and 'part of the individual' (Clouser, Culver, & Gert, 1997, p. 189). And, most importantly, maladies must cause, or put one at the increased risk of incurring, a harm or evil (defined as death, pain, disability, loss of freedom, and loss of pleasure).[12] Thus, according to CCG, maladies must be value-laden (because they necessarily involve harmful states that individuals normally disvalue and want to avoid). However, CCG also believe that whether something is a malady is a matter of natural fact because it is an

[12] Despite infelicitous name, 'evil' supposed to have no religious or moral connotations (CCG 1981, p. 31). CCG's original formulation of their 'malady' concept adopts the "evil" terminology, while their later work transitioned to the word "harm". This is not meant to be a substantive change, as they note that these terms can be used interchangeably (e.g., CCG 1981, p. 31).

objective fact that all rational humans disvalue harms. According to CCG:

> Although…values remain at the core of the concept of malady, the values are not only specified, but they are also objective and universal (Clouser, Culver, & Gert, 1997, p. 179)

CCG continue:

> [Death, pain, disability, loss of freedom, and loss of pleasure] are harms that every person acting rationally wants to avoid. This explains why and in what way malady (or disease) is a normative term. The concept involves values, certainly, but they are objective and universal values (Clouser, Culver, & Gert, 1997, p. 184).

There are some appealing aspects of this theory. Most notably, CCG's theory appears to avoid most of the epistemic difficulties facing traditional function-based views. According to CCG, we don't need an accurate supertheory of mental functioning in order to identify mental disorders; all we have to do is identify mental states that cause harms. So, if some mental state can plausibly be described as causing death, pain, disability, loss of freedom, or loss of pleasure, then it will be a genuine disorder. CCG (1997) acknowledge that there is some vagueness over what, exactly constitutes a disability, pain, loss of pleasure, and loss of freedom, but they take this vagueness to be more benign than the uncertainly facing function-based views. Thus, according to CCG, paradigmatic mental disorders such as depression, anxiety, and schizophrenia will be genuine disorders (or maladies), because they are harmful states that all rational humans want to

avoid. And, according to the malady theory, merely 'dysfunctional' states such as the lack of desire to reproduce, homosexuality, or asexuality will *not* be disorders (or maladies) because they do not (by themselves) cause harm (e.g., any harm would be due to a distinct sustaining cause such as cultural disapprobation). For the sake of argument, let's assume that this theory is coherent and tenable. As with the other views, the 'malady' theory, even in its strongest form, runs into difficulties. For example, according to CCG:

Pregnancy is a malady:

> [P]regnancy is a malady since it is clearly a condition of an individual, other than her rational belief or desire, such that the individual is suffering, especially in the last several months, some pain and disability. Also, throughout pregnancy she is at a significantly increased risk of incurring these harms (Clouser, Culver, & Gert, 1997, pp. 205-206).

Menopause is likely a malady:

> [Menopause] is a condition of the individual that necessary involves a disability" (Clouser, Culver, & Gert, 1997, p. 207).

And menstruation is *possibly* a malady:

> "to the extent that the [menstruation] does cause discomfort and pain, it seems as if it would be considered a malady (Clouser, Culver, & Gert, 1997, p. 208).

CCG accept that these examples are ‘troublesome borderline cases’, and acknowledge that they *may* require a modification of the otherwise explanatorily successful theory (Clouser, Culver, & Gert, 1997, p. 205). Given that no principled modifications are on offer, these ‘troublesome’ cases are indeed disconcerting. Martin (1985), for example, argues that CCG’s theory is sexist and pathologizes being a woman (or at least significant aspects of being a woman). Boorse (1997), in a similar vein, argues that ‘to call pregnancy *per se* unhealthy would strike at the very heart of medical thought; it is the analytic equivalent of the "Game Over" sign in a video game’ (p. 44). While it is very likely that CCG would consider many aspects of Boorse’s theory to also be the equivalent of a ‘Game Over’ sign, Boorse and Martin are correct in noting that the ‘troublesome’ aspects of CCG’s theory reflect a serious problem.

CCG’s view is also relativistic about the existence of mental disorders over time. The malady definition excludes harm caused by rational beliefs from being disorders. This is supposed to avoid the pathologization of *all* harmful beliefs. So, the (presumably rational) belief that one is bankrupt may be harmful, but not necessarily disordered, while the (presumably *irrational*) belief that one is therefore completely unlovable and will always be a failure likely will be both harmful and disordered (if it persists). The problem, here, is that the beliefs which count as ‘rational’ will change over time and between cultures. According to CCG, ‘[a] belief is irrational only if its falsity is obvious to almost everyone with similar knowledge and intelligence’ (Clouser, Culver, & Gert, 1997, p. 188). But, following this definition, an atheist who feels intense sadness at the realization that there is no afterlife would be irrational (and thus have a mental malady) if she lived in a time when most of her intellectual peers were devout believers

(even if she never told anyone about these beliefs).[13] This would likely not be the case now even if the beliefs caused the *same* exact amount of harm. The general problem here is that whether or not some way of mental functioning is a malady is not just determined by whether it is harmful, but is also contingent upon the cultural and religious standards of what is and is not 'obvious'. CCG's malady theory, then, may be right in claiming that there is an objective matter of fact about what we *now* consider to be an 'obvious' mental disorder, but this natural fact about humans must be relativized to our current, and evolving, standards.

According to CCG: '[o]ur account eliminates as much of the subjectivity as possible, allowing much less room for manipulation. Our account of maladies is both precise and systematic, thus enabling a more fruitful discussion of controversial cases' (Clouser, Culver, & Gert, 1997, p. 178). However, despite their goal of objectivity and clarity, the malady theory retains a great deal of subjectivity and introduces a number of controversial disorders (even if these can be fruitfully discussed). As with the function-based views, the malady theory is not incoherent, but it is seriously flawed.

1.4 Normative Views

It is possible that the main problem with precisifying the objective foundation of mental illness is that it does not exist. Normative theories of mental health take this view. These views are best

[13] Note that this example would not fit under the 'external sustaining cause' clause. It is not the case that the atheist is merely harmed by the (real or perceived) judgment of others, and that this harm would cease if her environment changed. The focus here is on private pain caused by the lack of religious belief.

defined negatively: what makes it the case that some mental phenomenon is disordered, rather than healthy or benign, is not a matter of natural fact. Normative theories can accept that there are differences in human functioning, and that some of these differences may have a biological or genetic basis, but they deny that these differences are necessarily disorders. While normative theories differ greatly in their specifics, the unifying idea behind these views is that the concept of a mental disorder is a value-laden construct used to label disvalued or norms-transgressing behavior.

1.4.1 Motivating Normativism

There are three primary motivations for this view. The first is the claim, made by most normativist theories, that the concepts of mental health and disorder are ineluctably value-laden.[14] The idea, here, is that designations of health and disorder are *necessarily* value claims because they reflect human interests (which are value-laden). Sedgwick (1973) gives us a clear statement of this view:

> Outside the significances that man voluntarily attaches to certain conditions, *there are no illnesses or diseases in nature*"…. What, [the naturalist] will protest, are there no diseases in nature? Are there not infectious and contagious bacilli? Are there not definite and objective lesions in the cellular structures of the human body? Are there not fractures of bones, the fatal ruptures of tissues, the malignant multiplications of tumorous growths?

[14] See, for example: Scheff, 1970; Sedgwick, 1973; Margolis, 1976, 1980; Fulford, 1989; Nordenfelt, 1995.

> Are not these, surely, events of nature? Yet these, as natural events, do not — prior to the human social meanings we attach to them — constitute illness, sickness, or diseases. The fracture of a septuagenarian's femur has, within the world of nature, no more significance than the snapping of an autumn leaf from its twig: and the invasion of a human organism by cholera-germs carries with it no more the stamp of "illness" than does the souring of mild by other forms of bacterial. Human beings, like all other naturally occurring structures are characterized by a variety of inbuilt limitations or liabilities, any of which may (given the presence of further stressful circumstances) lead to the weakening or the collapse of the organism….[O]ut of his anthropocentric self-interest, man has chosen to consider as "illnesses" or "diseases" those natural circumstances which precipitate the death (or the failure to function according to certain values) of a limited number of biological species: man himself, his pets and other cherished livestock, and the plant-varieties he cultivates for gain or pleasure (pp. 30-31, original italics).

So, according to this view, the search for biological dysfunctions (or any other objective foundation for attributions of disorder) will, by itself, tell us nothing about whether some way of being is healthy or disordered. The point is that even if a plausible function-based naturalist theory of health could be formulated, it would not be a theory of *health*, but would rather merely be a theory of proper mental function (which may or may not play a role in the concept of health); any putative 'dysfunction' (e.g., a snapped femur) can only be a disorder if it leads to some end that humans disvalue.

The second prominent motivation for normativism is based on the presumed *conceptual priority* of values judgments over any presumably 'objective' metaphysical theorizing about health and disorder.[15] For example, Fulford (1989), followed by Nordenfelt (1995, 2001), argues that the value-laden concept of illness is 'logically prior' to the putatively value-independent concept of disease (p. 141). The idea here is that individuals are first *identified* as being mentally ill (i.e., as having something that has gone *wrong* with them), and only then is this illness explained by appealing to some biological cause (i.e., a *disease* entity). Thus, the argument goes, 'the conceptual structure of medicine is essentially evaluative (rather than factual) in nature' because the concept of disease/disorder is conceptually dependent upon a value judgement (Fulford, 1989, p. 260). So, according to Fulford:

> 'illness', not 'dysfunction', is the conceptual root notion in medicine, conditions being first marked out as illnesses by the value judgement expressed by 'illness' (1989, p. 68)

Similarly, Nordenfelt (2001) argues:

> [The value-laden concept of] [i]llness must be characterizable first. If we are not able to say that a person is ill before a disease has been found, then we cannot get off the ground. We must have a coherent concept of illness first. Given that, we can look for diseases. This also means that "disease" is a concept derivable from the concept of illness (pp. 56-57).

[15] See, for example: Rezneck, 1987; Fulford, 1989; Nordenfelt, 1995, 2001.

Therefore, according to the 'conceptual priority' views of mental health and disorder, naturalism is a necessarily failed conceptual enterprise because it fails to recognize that we can only arrive at the putatively value-independent concept of disease by first appealing to the necessarily value-laden concept of illness.

The third primary motivating factor driving the normative position is the perceived failure of naturalist views to provide a plausible objective foundation for the concept of mental health. This is a the most nebulous of the three camps. The argument here is not that it is an ineluctable conceptual truth that the concepts of health and disease must be value-laden, but rather that naturalist theories of mental health have failed to be both internally coherent and to capture in their extension all the phenomena that the best theory of health should.[16] Naturalist views, then, are not claimed to be inherently conceptually flawed, they are just very likely wrong. There is no one paradigmatic argument here. Rather, this motivation is driven, in large part, by the concerns raised in the previous section regarding the challenges facing traditional naturalist and hybrid-naturalist views.

1.4.2 Problems with Normativism

Like traditional naturalist theories, the normative view faces challenges. Consider, first, the ineluctability argument. The strength of the ineluctability argument depends upon the highly

[16] See, for example: Cooper, 2002, 2005; Prinz, 2012

questionable assumption that our metaphysics of *x* will necessarily be value-based if we have value-based beliefs or assumptions about *x prior* to engaging in our attempt to carve nature at its joints. So, according to this view, a broken femur only constitutes a disorder because humans generally disvalue the harms (and possible death) that may accompany it. The problem with this argument is that, if generalized, it appears to prove too much. Namely, if our metaphysics must remain untainted by our prior values, then most (if not all) metaphysics will ultimately be value-based. To see this, consider the ontological questioning of the existence of some *x* (e.g., midsize objects, time, God, etc.). It is entirely plausible that humans, in general, value existence more than nonexistence. Most of us, for example, would presumably rather be than not be. But we need not therefore conclude that ontology is necessarily value-laden. Rather, the more reasonable response would be to make the (much) weaker claim that, given our prior values, the *practice* of doing ontology may be influenced by our values. The value-ladenness of our minds should give us good reason to be cautious in the strength of our metaphysical conclusions and should cause us to reflect on the motivations for our intuitions, but it does not appear to necessarily follow from this that we must be of pure, disinterested minds to conduct metaphysics. That we may have value-based interests in how the world is carved up does not obviously entail that our best metaphysical theories are merely reflections of these interests.

The 'conceptual priority' argument is also problematic. This argument moves from the presumed *epistemic* fact that we first recognize certain ways of being as harmful and disvalued before classifying them as disordered, to the *metaphysical* claim that our classifications of certain mental phenomena as 'disordered' is therefore necessarily value-laden (i.e., because our metaphysics must now be necessarily infected by the value-based epistemic route that got us to

our metaphysical theorizing). However, the conceptual priority argument runs into the same problem as the ineluctability argument. Namely, there does not appear to be any necessary relationship here between *how* we come to recognize some problem, and our best metaphysical theory of *what* that problem is. From the presumed fact that our epistemic route to some problem *x* may be value-laden, does not, by itself, necessarily entail that the metaphysics of *x* is a necessarily a value-based enterprise. As long as we think that we need not come from a completely pure and value-independent state of mind to do metaphysics, then there appears to be no special value-based problem concerning the metaphysics of mental health.

The primary challenge facing the 'failure of naturalism' argument, and normative views in general, is to account for the perceived differences in kind between paradigmatic instances of mental disorder (such as bipolar disorder, schizophrenia, and anxiety) and other harmful and disvalued mental states or ways of being (such as being poor, being bigoted, or just being unpleasant to be around). If mental disorders are just disvalued ways of being, then there should be no matter of natural fact that differentiates them and other disvalued states. Normative theories can of course just deny there *is* any difference. What counts as a mental disorder, then, must be relativized to some set of values. This view is not incoherent, but it is counterintuitive. Ultimately, the strength of the 'failure of naturalism' argument rests on whether or not a naturalist theory is on offer that can provide a more plausible motivation for the presumed distinction between 'genuine' disorders and merely disvalued ways of being. This dissertation attempts to provide such a theory.

1.5 The Skill View of Mental Health

The problem is this: existing theories of mental health and disorder are not incoherent or obviously untenable, but they are significantly flawed. This dissertation argues that a more plausible naturalist view of health can be constructed if we reorient the dialectic away from the concept of normal function and focus instead on the skills responsible for mental health. Skill, rather than function, can provide the objective basis for a naturalist theory of mental disorder.

The new naturalist theory of mental health outlined here argues that mental health is best conceived of as a skill. More specifically, Chapter 4 argues that mental health is skilled self-regulation; mental disorder is a failure or breakdown of this skill. The skill view of mental health provides a naturalist framework for the scientific study and treatment of mental disorders that avoids the explanatory and conceptual failings of traditional naturalist and normative views. The focus on skilled action, rather than the adherence to either biological norms or value norms, allows for a flexible theory of health that avoids pathologizing normal human difference while also providing a non-revisionary accounting of paradigmatic mental disorders.

1.6 Overview

The body of this dissertation consists of three independent articles (chapters 2-4), a support chapter (chapter 5), and a brief conclusion (chapter 6). The dissertation motivates the skill view of mental health by first focusing on the process of mental *healing.* Theories of the metaphysics of health are normally developed in isolation from theories of the process of mental healing. This

is understandable. There is no necessary relationship between the improvement of patients' symptoms (a common goal in mental healing), and identifying what makes any particular constellation of symptoms a mental disorder (the goal of a theory of the metaphysics of mental health). However, this dissertation attempts to show that we can get a better understanding of what mental disorders are by focusing on how and why individuals heal.

Chapter 2 is a case study in the process of mental healing. Cognitive Behavioral Therapy (CBT) has become the dominant form of psychotherapy in North America. The CBT model is based on the theoretical assumption that all external and internal stimuli are filtered through meaning-making, consciously accessible cognitive schemas (or core beliefs). The goal of CBT is to identify dysfunctional or maladaptive thoughts and beliefs, and replace them with more adaptive cognitive interpretations. While CBT is clearly effective as a treatment, there is good reason to be skeptical that its efficacy is due to the causal mechanisms posited by the CBT model. This chapter argues that the theory of psychological healing grounding cognitive behavioral therapy is likely mistaken. CBT may be effective, but its efficacy does not appear to be a result of the modification of the specific core beliefs identified in CBT interventions.

Chapter 3 focuses on the more general question of why any psychotherapy is effective. Psychotherapy *is* effective. Since the 1970's, meta-analyses have consistently shown a significant effect size for psychotherapeutic interventions when compared to no treatment or placebo treatments. This effectiveness is normally taken as a sign of the scientific legitimization of clinical psychotherapy. A significant problem, however, is that most psychotherapies appear to be equally effective. This poses a problem for specific psychotherapies: they may work, but

likely not for the reasons that ground their theoretical explanations for their effectiveness. A prominent explanation for the findings of common efficacy in psychotherapy is to postulate that all successful therapies work by altering maladaptive meanings and providing patients with new, more adaptive meanings. This chapter argued that the 'meaning view' of psychological change is likely mistaken; psychological problems are not normally problems of meaning nor are they directly ameliorated by changes in meaning. This chapter then outlined a skill-based explanation for the findings of common efficacy in psychotherapy.

Chapter 4 expands upon this skill-based theory of mental *healing* to develop a novel skill-based account of mental *health*. Ch. 5 then shows that the skill view of mental health is compatible with any plausible theory of metaphysics of skill.

Chapter 2: Know Thyself? Questioning the Theoretical Foundations of Cognitive Behavioral Therapy.

2.1 Introduction

In the latter half of the 20th century cognitive behavioral therapy (CBT) replaced psychodynamic and behavioral therapies as the dominant form of psychotherapy in North America (Westbrook, et al., 2011; Norcross & Karpiak, 2012). This was largely due to CBT's perceived superior testability and efficacy in comparison to other forms of treatment. Roughly, CBT is a combination of behavioral therapy and cognitive therapy that aims at identifying and replacing maladaptive or dysfunctional thoughts and beliefs and replacing them with more adaptive cognitive interpretations. The CBT model of psychological functioning posits an interconnected triad of thoughts, behavior, and emotions, with thoughts playing the primary role in the development and treatment of dysfunctional psychological states (A. Beck, 1979; J. Beck, 2011). Maladaptive or dysfunctional emotional or affective responses are modified by altering thoughts (either directly, or indirectly through behavioral interventions, or both). This model's theory is based on the idea that all external and internal stimuli are filtered through meaning-making, consciously accessible cognitive schemas, or core beliefs, that can represent the world in either adaptive or maladaptive ways. [17] The goal of CBT is to help patients to identify, challenge, and

[17] There is a good deal of ambiguity in the use of the word "schema". Many authors, such as A. Beck (1976), J. Beck (2011) and Clark (2004) us the terms schemas and core beliefs interchangeably. Others, such as Young et al. (2003) define schemas as any semantic cognitive filter. I will be following the latter usage.

replace the specific dysfunctional or maladaptive beliefs that are postulated to be the primary factor in their psychological disorder.

While CBT is clearly effective as a treatment, there is good reason to be skeptical that the efficacy is due to the causal mechanisms posited by the CBT model. This paper will argue that the specific cognitive schemas posited by the CBT model likely do not play a direct role in the development or treatment of psychological illness. Cognitive schemas, as identified in CBT interventions, are likely the result of patient confabulation and epistemically under-supported practitioner-based identification. CBT interventions appear to impose coherence on patients' psychological states, rather than actually identifying and modifying existent causally efficacious core beliefs.

This discussion will first outline Beck's CBT model of the affective disorders, then highlight the CBT model's reliance upon introspective and retrospective belief reports in identifying and challenging maladaptive cognitions. The discussion will then focus on problems with the CBT model's reliance upon direct introspective access to patients' cognitive processes, and conclude with suggestions for the construction of a more plausible cognitive theory.

2.2 The CBT Model

At its most basic, the CBT model, first posited by Albert Ellis (1962) and Aaron Beck (1967), is concerned with the relation between cognitions, emotions, and behavior. Cognitions (thoughts, beliefs, and assumptions) are posited as playing the primary role in the formation and treatment

of dysfunctional or maladaptive psychological states (Clark & Beck, 1999). How individuals interpret the world is supposed to influence, and be influenced by, their behavior and emotions. Maladaptive emotions (such as depressive states) are conceived of as subjective states caused by overly rigid and/or inaccurate cognitive appraisals or evaluations of internal or external stimuli (Clark & Beck, 1999). How a stimulus is interpreted by the informational processing system determines the valence, persistence, malleability, and intensity of emotional responses. The CBT model also maintains that behavior influences thoughts, and therefore also alters emotions. Changing maladaptive behavioral patterns is taken as a tool to indirectly change unhealthy cognitive patterns by way of challenging unhealthy cognitions (e.g., safely exposing a patient to an irrationally fear-inducing stimulus is used to challenge and alter negative thoughts and thus extinguish the negative emotional response). CBT interventions aim to provide patients with less dysfunctional or distorted, and more adaptive and realistic, meaning-making interpretations of the world.

The most influential, studied, and applied cognitive therapy is Aaron Beck's cognitive behavioral therapy (Beck, 1976, 1979).[18] The Beckian cognitive model (henceforth, CBT) posits three levels of cognitions that are supposed to filter all experience: automatic thoughts, intermediate beliefs/assumptions, and core beliefs/schemas (J. Beck, 2011; Leahy, 1996; Clark & Beck, 1999;

[18] The umbrella term 'CBT' has grown to include a number of related therapies that include both cognitive and behavioral components. A number of recent 'third wave' cognitive therapies (e.g., mindfulness-based cognitive therapy [MBCT], dialectical behavioral therapy [DBT], and meta-cognitive therapy [MCT]) are often categorized as 'cognitive behavioral therapies', despite differing in theory and practice from Beck's CBT. For context, Tolin's (2010) meta-analysis of the relative effectiveness of CBT included 26 studies labeled as delivering 'CBT'; 12 of the 26 were explicitly based on Beckian CBT (including 8 of the 10 studies on the treatment of depression). This paper focuses on the dominant Beckian model.

Westbrook et al., 2011). The most basic level of cognitive processing, core beliefs/schemas, are supposed to "enable individuals to make sense of their environment by breaking it down and organizing it into psychologically relevant facets...[and] direct all cognitive activity whether it be ruminations and automatic thoughts or cognitive processing of external events" (Clark & Beck, 1999, p. 52). If things are running well, one's schemas represent the world in ways that do not lead to psychological distress or maladaptive thoughts and beliefs. Things start to go poorly when one's cognitive processes represent the world in overly rigid, negative, or polarized ways.

Automatic thoughts are supposed to sit at the most salient end of the cognitive hierarchy. Automatic thoughts are defined as easily consciously accessible, context-specific beliefs about, attitudes towards, or semantic interpretations of external and internal stimuli. These are surface-level thoughts that superficially explain individuals' thoughts and behaviors. For example, a patient may report having the negative automatic thoughts "I will be picked last" or "I will embarrass myself if I try" when deciding not to join in a group sports activity. Importantly, these thoughts are usually not explicitly held or consciously entertained, but are supposed to be easily identifiable by introspection or elicited by practitioner-based questioning.

Intermediate beliefs are the middle level of the CBT cognitive model and the immediate platform from which automatic thoughts are formed (J. Beck, 2011). They are rules or patterns of association used to interpret and evaluate experiences. For example, the automatic thought "I am being boring" that may ground a patient's feeling of unease during a social situation may be grounded in the intermediate belief that "if I talk too much, then people will think I'm boring". These beliefs often take the form of conditional statements, such as "If I please my partner, then

he/she will treat me well" or "If I am criticized, then it means that I have failed" (Clark & Beck, 1999). Maladaptive intermediate beliefs are characterized by being overly rigid, based in thought errors such as catastrophizing or all-or-nothing thinking, and containing distorted world views. The CBT model posits that these beliefs are more difficult to identify than automatic thoughts and must normally be inferred by patient and practitioner from the patterns and content of automatic thoughts.

Core beliefs are the most basic and fundamental beliefs about oneself and the world and are the basis of all other consciously accessible cognitions. The beliefs are highly generalized, absolute, and difficult for patients to consciously access. Practitioners are trained to identify them by way of recognizing consistent patterns in patient belief and thought reports. Negative core beliefs may take the form of statements such as "I am a failure" and "I am unlovable", while positive core beliefs are expressed by thoughts such as "I am likable" and "I am worthwhile" (Clark & Beck, 1999). The CBT model postulates that all dysfunctional or maladaptive automatic and intermediate thoughts and beliefs are the result of dysfunctional or maladaptive core beliefs (Clark & Beck, 1999; J. Beck, 2011).[19]

Patients' intermediate and core beliefs are identified in CBT interventions via Socratic questioning and downward arrow interviewing (Neenan & Dryden, 2005; J. Beck, 2011). Socratic questioning (also called 'guided discovery') consists of persistent questioning of the

[19] The Beckian model has been modified to include non-consciously accessible cognitive processes (Beck, 1996; Beck & Clark, 1999). While cognitive therapy still focuses on identifying and challenging beliefs and and thoughts, the CBT model now postulates that clusters of interrelated schemas called 'modes' play a significant role in cognitive functioning.

patient's reasons and justifications for having specific automatic thoughts. The aim here is to aid the patient in searching for the (possibly distorted or unhealthy) thoughts or beliefs that explain her thoughts, feelings, and behavior (Neenan & Dryden, 2005). For instance, patients may be asked to identify patterns in their behavior, explain why they think they have certain thoughts, and explain what specific thoughts mean to them. Similarly, in downward arrow questioning, the practitioner attempts to identify core beliefs by asking the patient to identify what their previously identified automatic or intermediate beliefs *mean* to them (J. Beck, 2011). This process is repeated until the patient arrives at the lowest level of abstraction (core beliefs). Consider the following example of the downward arrow technique taken from a CBT training handbook (Harwood et al., 2010):

> Situation: At home on a Saturday afternoon.
>
> Emotions: Depressed (80%), anxious (60%).
>
> Automatic thought: "I should have a date on Saturday night."
>
> Therapist: What does it mean if you don't have a date on Saturday night?
>
> Patient: It means that I'll be home by myself on Saturday night.
>
> Therapist: What does being home alone on a Saturday night mean?
>
> Patient: It means that I'm not out having fun like everybody else.
>
> Therapist: And what does that mean to you?
>
> Patient: That I'm a loser, nobody loves me, and I'll always be alone.

In this example, the core belief "I'll always be alone" was elicited by the patient attempting to make sense of, or find the deeper meaning in, her higher level thoughts. Again, CBT theory

posits that thoughts and beliefs are based on more basic thoughts and beliefs, with core beliefs filtering all semantic interpretations of the world. Patients and practitioners attempt to make sense of maladaptive automatic thoughts by locating more general thoughts and beliefs that would explain why the automatic thoughts are held in the first place. From here, the thoughts can be challenged and modified. This process usually involves the patient keeping a thought record to identify her automatic thoughts, then challenging and weighing the accuracy of the thoughts both in session and through homework. Behavioral interventions, such as increasing pleasurable activities and benign exposure and habituation to perceived harmful or fear-inducing activities, may also be used. In both the approaches the aim is to change how the patient thinks about and interprets the world.[20] The patient is challenged (both in person and through homework assignments) to question and find reasons to undermine the distorted or dysfunctional core and intermediate beliefs and replace them with more accurate and adaptive beliefs.

2.3 Criticisms of the CBT Model

A number of randomized controlled trials, meta-analyses, and meta-analyses of meta-analyses have shown CBT to be an effective therapeutic treatment for a wide range of psychological problems (Leichsenring et al., 2006; Clark & Beck, 1999; Butler et al., 2006). Despite its efficacy, CBT is not without critics. Objections to CBT theory come in four main camps: (1)

[20] While CBT adopts many therapeutic methods from behavioral therapy, CBT theorists and behaviorists give differing explanations for the therapeutic change engendered by the use of behavioral therapeutic techniques. Traditional behaviorist theories focus on the alteration of conditioned, non-consciously accessible behavioral rules (rather than consciously accessible inner states) to explain psychological change (Skinner, 1977). According to CBT theorists, behavioral methods are successful only insofar as they help modify patients' maladaptive thoughts and beliefs (Beck, 1979).

criticisms of the primacy given to cognitions over other psychological processes (such as emotion or non-consciously accessible drives) (e.g., Teasdale & Barnard, 1993; Teasdale, 1997), (2) criticisms of cognitive theory being overly general or metaphorical (e.g., Coyne & Gotlib, 1983; Brewin, 1996), (3) criticisms of CBT's (and every other theory's) lack of causally efficacious specific effects (e.g., Wampold, 2015), and (4) criticisms of CBT's assumptions about the representational nature of cognition (e.g., McEachrane, 2009; Gipps, 2013). What is common to these objections is the idea that if CBT works, it is not because of the reasons given by the theory grounding the therapy.

Supporters of cognitive theories of psychopathology have countered group (1) type criticisms by appealing to the substantial literature on the ubiquity of, and central role for, maladaptive cognitions in cases of psychological distress (Clark & Beck, 1999; Beck, 2004). The first versions of cognitive therapy may have been vulnerable to group (2) type critiques. Early CBT theorists such as Beck and Ellis were mainly concerned with establishing a dominant role for thought in depression (in contrast to psychodynamic and behavioral learning models), rather than focusing on the specifics of how this might work. Later versions of the cognitive model have addressed this problem by becoming far more specific as to what meaning-making structures are, how they are structured in the informational processing system, and the roles they play in psychological disturbances (Clark & Beck, 1999; Beck, 2005). The standard CBT response to group (3) objections is to either challenge the accuracy of meta-analyses (Crits-Christoph, 1997; Butler et al, 2006) that purport to show an equivalence of effectiveness across different therapies or to claim that if other forms of therapy work it is only because they are changing cognitions- in effect, other therapies are actually doing a form of CBT (Alford & Beck, 1998). According to the

latter view, all therapies work by challenging specific cognitions- either directly (as in CBT interventions) or indirectly (as in other non-cognitive, but efficacious, treatments). There is some plausibility to this response. But, importantly, this reply rests on the assumption that there are specific beliefs and automatic thoughts that play a primary role in the production of psychopathological states.

The group (4) objections to the CBT model criticize its (and most of cognitive science's) assumptions about the representational nature of belief and thought. Representational theories of cognition take beliefs and thoughts to be causally efficacious mental representations of facts, states of affairs, or propositions. Critics of representationalism have argued that CBT confuses individuals' thought reports (which are represented as having imagistic or linguistic content), with their thoughts (which needn't have any distinct representational content at all) (McEachrane, 2009). Rather than respond directly to type (4) objections, CBT theorists take representationalism to be a foundational assumption of cognitive science and clinical cognitive theory (Clark & Beck, 1999). This paper will share CBT's assumptions about the representational theory of mind.

This paper raises a fifth set of concerns about the cognitive model. The CBT model is based on the assumption that our cognitions have a coherent, logical, and consciously accessible hierarchical structure; the content of all cognitions is based on more general cognitive content. If patients can introspectively identify their automatic thoughts, then the cognitive model predicts that they should be able to identify the more general schemas that ground these thoughts. However, there are serious problems with both the tenability of CBT's hierarchical model of

cognition and its assumptions about the accuracy of the cognitions being identified and challenged in therapy.

2.4 Thoughts, Beliefs, and Confabulation

2.4.1 Identifying Cognitions

There is an oddity to the CBT model. According to CBT, our meaning-making information processing systems are posited to be actively creating our reality, *unless we are introspecting.* CBT is based the idea that all “stimuli that impinge on the organism" are filtered through cognitive schemas that structure and give meaning to experience (A. Beck, 1967). Our informational processing system is supposed to "actively participate in the construction of reality”, and this construction "is not simply an act of representing, copying, or "coding" fixed objects but rather is a process that involves some degree of creativity" (Clark & Beck, 1999). Yet we are supposed to be accurate introspectors of our past thoughts and beliefs. CBT theory adopts a constructivist view of cognition, but a more-or-less realist view of introspection. Our cognitive processes are identified as “meaning-making structures" that can either represent the world in maladaptive or adaptive ways, but at the same time we are supposed to be able to accurately identify the cognitions underlying our behavior rather than simply "making sense" of our emotions and behaviors. This is likely not the case.

CBT assumes that patients have, or can be trained to have, direct and accurate access to the content of their own cognitions (Beck & Dozois, 2011). CBT practitioners are supposed to aid patients in identifying their own thoughts and beliefs by engaging in directed Socratic

questioning aimed at eliciting deeper cognitive schemas. Practitioners guide patients' introspection of their core schemas by identifying common themes in the patients' automatic and intermediate thoughts and directing patients to search for the underlying structure in their thoughts and beliefs. For example J. Beck states, "asking what a thought means to the patient often elicits an intermediate belief; asking what it means about the patient usually uncovers the core belief" (2011). This identification of cognitions often requires work on the part of both the patient and therapist. A critical part of cognitive therapy is to first *train* patients to recognize, attend to, and record this inner speech or automatic thoughts (A. Beck 1976, 1979; J. Beck, 2011). Patients often claim to be unaware of having an "internal communication system" or are unused to attending to the content of these thoughts or images (ibid). The training process involves explaining the cognitive model to patients and articulating the logical connection between core beliefs and automatic thoughts and the relation between thoughts and emotions. Beck (1976) states:

> The training in the observation and recording of cognitions makes the patient aware of the occurrence of images and self-verbalizations ("stream of thought"). The therapist trains the patient to identify distorted and dysfunctional cognitions. The patient may need to learn to discriminate between his own thoughts and the actual events. He will also need to understand the relationship between his cognitions, his affects, his behaviors, and environmental events. (p. 146)

Similarly, Beck and Alford (2009) state:

> At the beginning of therapy the patient is generally aware only of the following sequence: event or stimulus—>affect. He must be trained to fill in the link between the stimulus and the affect: stimulus—>cognition—>affect (p. 310).

There are reasons to be skeptical about the success of this training. Automatic thoughts are identified by simply asking patients to introspect what they were thinking at any given moment (e.g., what they were thinking while feeling sad staying home on a Saturday night). Most automatic thoughts are not explicitly entertained (insofar as they are not salient parts of a patient's inner monologue) and require introspection and practitioner-based prodding to identify. For example, the J. Beck CBT manual states: "Automatic thoughts are usually quite brief, and the patient is often more aware of the *emotion* she feels as a result of the thought than of the thought itself. Sitting in session, for example, a patient may be somewhat aware of feeling anxious, sad, irritated, or embarrassed but unaware of her automatic thoughts until her therapist questions her" (2011, original italics). In cases where the patient is unable to identify any thoughts or confuses thoughts with feelings, practitioners will use questions such as "what would you guess was running through your mind at that time?", "what did this situation mean to you (or about you)?", or even "might you have been thinking __ or __?" (J. Beck, 2011). According to the cognitive model, "the emotion the patient feels is logically connected to the content of the automatic thought" and it is the job of the practitioner to help the patient identify this logical connection (J. Beck, 2011). However, a serious problem with this process is that what a thought means to a patient is highly dependent upon the theory of cognition and psychological functioning being deployed by both patient and practitioner. This explicit search for meaning is,

according to the CBT model, itself based on meaning-making schemas that need not accurately represent anything (Clark & Beck, 1999).

It makes sense that automatic thoughts are based on more basic core and intermediate beliefs only if one adopts a theory of psychological functioning that posits a nested hierarchy of consciously accessible and causally efficacious thoughts. Importantly, many other explanations can also make sense of the same stimuli without reference to a hierarchy of cognitions. For example, most modern psychodynamic theories posit conflicts between subconscious and conscious feelings and drives (the sex drive, self-esteem, etc.) as the basis for psychological distress (Gabbard, 2000; Wolitzky & Eagle, 1992). For patients who adopt a theory of psychological functioning steeped in the Freudian-inspired psychodynamic model, thoughts such as "I should have a date on Saturday night" may mean that the patient has repressed subconscious-based anger towards the perceived loss of parental affection (or any number of other possible interpretations). In the middle of the 20th century Freudian-inspired drive based theories of psychological processing and object-relation theories (which based mental illness in the feeling of real or perceived loss of object(s) in early childhood) grounded how most of psychology—and much of the educated populace—made sense of their mental lives. The conflict between the Id, Ego, and Super Ego made sense to many people for a long time before falling out of fashion in favor of cognitive and behavioral theories. Just as we should be aware of the influence of theory upon a patient's Freudian interpretation of her lack of Saturday evening plans, we need to also be cautious in accepting at face value a CBT model inspired interpretation of the meaning of a patient's thoughts.

It is also important to be cautious in accepting appeals to common sense. CBT's theory of psychological functioning and assumed nested hierarchy of consciously accessible and causally efficacious thoughts is explicitly intended to be built upon a common sense notion of how the mind works (Beck 1976; Ellis 1962/1994). This is claimed to be another mark of its superiority to the supposedly more unintuitive theoretical assumptions of psychoanalytic and behavioral therapies. The problem, of course, is that common sense and armchair models of cognition do not necessarily map on to how the mind actually works; common sense and folk psychology are bound by culture and context. It may be common sense to some Freudian-inspired individuals to assume that most desires, including infantile and childhood desires for comfort and attention, are sexual in nature (Freud, 1905/2000). It is also common-sensical to adherents of the medical model of mental illness (including many psychiatrists) that negative or maladaptive automatic thoughts are the product of neurochemical imbalances in the brain, rather than of dysfunctional beliefs (Lebowitz, 2014; Pescosolido et al., 2010). Thoughts and beliefs, according to this model, are symptoms of neurochemical problems rather than the primary problem itself. The commonness, or intuitiveness, of the common sense assumption that consciously accessible, logically structured beliefs play a primary role in the development and treatment of the affective disorders is dependent upon the acceptance (be it implicit or explicit) of the cognitive model of cognition.

The theory-ladenness of the CBT processes is significant. In order for CBT to work as theorized, patients and practitioners must be able to accurately identify maladaptive automatic thoughts in order to then challenge and modify them (or the core and intermediate beliefs that ground them). This requires that patients identify the *actual* thoughts that explain and cause their feelings and

behaviors, not just identify thoughts that *describe* or *make sense of* these states. But the latter is what the cognitive model would predict. CBT is based on the assumption that our informational processing systems actively create and structure our subjective realities. Whether these representations are accurate or inaccurate should not be important; all that matters for healthy psychological functioning is that our schemas represent the world in adaptive rather than maladaptive ways.

The theory-ladenness of CBT patients' explanation for the meaning of their thoughts and emotions, and their subsequent belief identifications, does not entail that the theory itself is wrong. There are independent reasons to think this. Notably, there are serious flaws with the CBT model's assumptions about the logical connection between automatic thoughts and emotions, and less consciously accessible intermediate and core beliefs, as well as problems with CBT's assumptions about patients' introspective access to their own beliefs. This section will address these problems.

2.4.2 Problems with the Cognitive Model

CBT posits that automatic thoughts have logical connection to core beliefs; if you think something, you think it for an identifiable and (at least internally) coherent reason. But this is often not the case. Contrary to the CBT model, there is strong evidence that automatic thoughts are often not logically connected to, or derived from, stable and consciously accessible core and intermediate beliefs. Individuals have restricted introspective access to cognitive processes causally responsible for much of their behavior and thoughts; while we often know what we are

feeling, we do not often know why we are feeling it (e.g., Wilson, 2002; Kahneman, 2011, Carruthers, 2011). Individuals' self-reports of the causes of their cognitive states and behavior, rather than being based on direct introspective awareness, are often confabulations based on post-hoc rationalizations or a priori causal theories (Nisbett & Wilson, 1977; Wilson, 2002; Haidt, 2006). Individuals will often make sense of their emotions and behaviors regardless of whether the explanation accurately reflects the actual causal story. Confabulation, or spontaneous unintentionally fabricated or distorted memories, often occurs when individuals are put in a position to explain thoughts and behaviors. CBT training and therapy does just this.

We should be skeptical of CBT's assumption that consciously accessible core beliefs are the logical source of one's emotions and automatic thoughts. Non-conscious heuristics, environmental factors, and implicit cognitive biases, rather than core beliefs, often influence our behavior, thoughts, and judgments. Environmental factors play a significant role in influencing cognitions outside of conscious awareness. Priming effects (specific behavioral changes after being exposed to a stimulus) affect both behavior and cognition. For example, exposure to pleasant environments (such as pleasant smells) significantly increases helping behavior, while unpleasant environments (such as unpleasant smells or messy rooms) decrease such behavior (Isen & Levin, 1972). Being primed by negative or positively valenced words appears to make individuals more or less likely to act impolitely (Bargh et al., 1996). Neat or messy work environments appear to prime individuals' moral judgments (Schnall et al., 2008), their purchasing habits (Liu et al., 2012), and even their opinion of a therapist's competence (Nasar & Devlin, 2011). While behavioral priming studies mainly focus on responses to environmental factors, it is very unlikely that the true causal processes behind the behavioral responses are

noticed by individuals. One of the most interesting aspects behind the priming studies is that the subjects are normally unaware of the priming's effect on their behavior. For example, it is highly unlikely that people would explain their helping behavior as being caused by morally arbitrary factors such as standing in front of a bakery rather than by appealing to some explanation based on their character and personality. Instead of having direct introspective access to the cognitive processes that ground their behaviors (such as the positioning of a garment or cleanliness of a room), individuals often appear to be in the position of cognitive interpreters of their own past feelings and behaviors. By asking for the meaning and cause of thoughts, CBT practitioners are asking patients to identify the causal processes responsible for their cognitions; this is something people are famously bad at.

By employing a post-hoc introspective lens on their cognitive processes, patients are likely often identifying thoughts that *describe* how they feel rather than uncovering their actual thought processes. This is an important difference. The CBT cognitive model maintains that CBT interventions work by accurately identifying and challenging dysfunctional or maladaptive thoughts, not merely by helping patients find ways to conceptualize, then alter, their psychological illness (e.g., "cognitive techniques are aimed at delineating and testing the patient's *specific* misconceptions and maladaptive thoughts" (Beck 1979, italics added) and "[w]hen you [the therapist] ask for patients' automatic thoughts, you are seeking the *actual* words or images that have gone through their mind" (J. Beck, 2011, original italics)). If the CBT model was to consider the beliefs identified and challenged in therapy as just *one* of many equally effective ways for patients conceptualize their psychological problems, then CBT would be on similar theoretical footing with other successful therapies (such as psychodynamic or

Freudian psychosexual therapies) with distinct and often incommensurate theoretical rationales. Nevertheless, CBT interventions only require that patients identify thoughts that it would make sense to have given their feelings/behaviors, or that offer coherent explanations for their feelings/behaviors. The CBT model has no method for testing whether a patient actually *had* the non-consciously entertained underlying thought "I should be out on a Saturday night" when tasked to explain what she was thinking while feeling sad and lonely, rather than it being the case that the thought was a post-hoc confabulation given by the patient to *explain* to the practitioner and herself why she felt sad and lonely.

It is important not to overgeneralize. It is certainly not the case that we have no idea about the content of our thoughts and beliefs. Many thoughts are explicitly held, repetitive, and easily identifiable (as is often the case with obsessive-compulsive disorders). The important point here is that most core beliefs and automatic thoughts are not consciously entertained. CBT interventions are based primarily on post-hoc identifications of normally non-salient beliefs and the focus of most interventions is for both the patient and practitioner to become aware of the patient's previously implicitly held thoughts. Cognitive therapy requires that patients try to identify what thoughts would make sense of their actions and feelings at some particular time. It is this theory-laden post-hoc act of thought and belief identification that we should be skeptical of.

In support of the CBT model, there do appear to be strong relationships between negative and overly rigid cognitions or thinking styles with depression (Solomon & Haaga, 2004), overly rigid, ruminative, and irrational cognitions with anxiety (Clark & Beck, 2011), and overly rigid,

repetitive, and intrusive cognitions with obsessive disorders (Clark, 2004). However, these findings only show a correlation between styles of thinking and reported thought content, on the one hand, and the diagnosis of affective and personality disorders, on the other. These studies do not directly support the hypothesis that individuals have stable, consciously accessible core and intermediate beliefs, nor do they directly support the hypothesis that there is an introspectively accessible logical connection between one's automatic thoughts and one's intermediate and core beliefs.

It seems likely, then, that the CBT model has a significant problem. If CBT is to work as theorized, patients must be able to accurately identify their automatic thoughts and the core and intermediate beliefs that ground them. However, rather than identifying actually held thoughts, it is likely that Socratic questioning and downward arrow meaning-questioning produce confabulated post-hoc explanations for the patients' emotions and behaviors. And, crucially, this is exactly what we should expect given CBT's own assumption of world-constructing information processing. CBT interventions appear to be imposing coherence on patients' illness by giving them a way to conceptualize their emotions and behaviors rather than identifying and challenging specific thoughts and core beliefs.

2.5 CBT Controls for Introspective Accuracy

CBT theorists have been largely unconcerned about the issue of patient confabulation. Therapists are cautioned to be careful about the possibility of patients misidentifying their own cognitions and warned to avoid influencing patients' belief reports, but these suggestions are brief and

optimistic. For example, in regards to patient's belief reports, A. Beck's (1976) manual suggests that:

> [T]he therapist should be on guard against accepting facile explanations and should check the reliability of the patient's reports of his introspections. The therapist can acquire confidence that he understands the totality of a particular experience by entering into the patient's "phenomenal world" (p. 30).

The idea here is that by carefully listening to patients' descriptions of their thoughts and beliefs, the therapist should be able to "step into the patient's world" and help identify which beliefs and cognitive patterns are playing the primary roles in patients' psychological distress. At the same time therapists are also prompted to be on guard against leading patients' narratives of their cognitions:

> Since the therapist's questions and other verbal techniques are derived from his own theory, he must be especially vigilant regarding "putting ideas in the patient's head." The therapist should be aware of his leading questions, the patient's suggestibility, and his desire to please the therapist by providing the answers he believes the therapist is seeking (p. 142).

J. Beck's (2011) CBT manual also warns practitioners to avoid "leading" the patient, while at the same time requiring that the therapist train subjects to accurately identify their thoughts and

beliefs and teach patients about the causal relation between thoughts and emotions. J. Beck states that:

> Whenever you [the therapist] present your interpretations, you will do so tentatively and label them as hypothesis, asking patients whether they "ring true." Correct hypothesis generally resonate with the patient (p. 198).

> You should regard your hypothesis as tentative until confirmed by the patient…Some patients are intellectually and emotionally ready to see the larger picture early on in therapy; you should wait to present it to others (especially those with whom you do not have a sound therapeutic relationship, or who do not really believe the cognitive model). As mentioned previously, whenever you present your conceptualization, ask the patient for confirmation, disconfirmation, or modification on each part (p. 205).

While therapists are briefly cautioned to be careful about thought insertion and confabulation, CBT theorists seem confident that the process of practitioner-guided discovery allows patients to gain direct introspective access to the logical relation of their thoughts and emotions. There appears to be a number of possible sources for this confidence.

First, CBT's reliance on veridical belief reports may be thought to be supported by the use of empirically supported questionnaires such as the Cognitive Bias Questionnaire (CBQ, Krantz & Hammen, 1979), the Beck Depression Inventory (BDI, Beck & Steer, 1991), and the Young Schema Questionnaire (YSQ, Young, 2003) that attempt to measure the accuracy and emotional

valence of patients' thoughts. These questionnaires all have statistically significant, though sometimes modest, test-retest reliability (which assesses patients' scores on the same test taken at different times) (Beevers et al., 2007). For example, the Cognitive Bias Questionnaire requires that patients read vignettes involving interpersonal situations then imagine they are in the situation in question and answer a series of multiple choice questions about what they would think and how they would feel. The multiple choice options include obvious over generalizations (e.g., "nobody wants to work with me"), signs of depressive thinking (e.g., "I don't deserve the raise because I'm worthless"), or healthy responses (e.g., "I probably didn't get the job because someone else was more qualified") (1979). The questionnaire is then scored to identify possible thought errors (e.g., catastrophizing, over generalizations, or all or nothing thinking), distorted world views (e.g., the belief that the world is entirely unsafe), or distortedly valenced thoughts (e.g., overly negative interpretations of events). The CBQ, and questionnaires like it, do seem successful in identifying biases towards distorted or erroneous thought patterns and depressive thinking. However, these questionnaires do nothing to test the veridicality of patients' own belief reports about the specific contents of their automatic thoughts and beliefs. The CBQ, and questionnaires like it, measure whether patients *identify* with certain maladaptive thoughts, not whether they are trustworthy interpreters of their own cognitions. The problem with the CBT model is not that it fails in identifying whether individuals are prone to certain maladaptive psychological states, but rather that it fails in accurately identifying the specific cognitions that are the putative focus of CBT interventions.

Similarly, another common test, the Young Schema Questionnaire, asks patients to evaluate statements such as "I believe that other people can take care of me better than I can take care of

myself" on a 1-6 Likert-scale (a score of 1 being "completely untrue of me" and 6 being "describes me perfectly") (2003). The answers are then scored to identify patients' underlying core schemas. The YSA has been shown to have a statistically significant test-retest reliability of .5 to .8 (Young et al. 2008). This may be taken as evidence that the questionnaire is measuring persistent, stable thoughts. However, the problem again is that the YSQ, and tests like it, only measure whether the patients' beliefs are consistent with the possession of certain core beliefs, not whether the patient actually has the beliefs in question. And while the test-retest correlation rate may be statistically significant, a 30-50% difference in answers between tests is also evidence that the questionnaire is identifying general themes (e.g., concern about loss and self-esteem) rather than specific core beliefs (e.g., "my life is out of balance"). Given that patients can behave and feel in ways that are consistent with a number of theoretically distinct psychological explanations, the use of questionnaires is only successful at addressing intra- (rather than inter-) theory issues.

CBT theorists may also take patient and practitioner identification of cognitions at face value based on the proven efficacy of CBT interventions. The CBT model posits that specific psychological dysfunctions are caused by specific maladaptive thoughts and beliefs and prescribes a uniform treatment plan for each unique dysfunction. This uniformity makes CBT easier to study than less rigid forms of psychological intervention such as psychodynamic approaches which focus heavily on the patient-practitioner relationship and the uniqueness of each patient. CBT's superior testability has led to it becoming the most tested, and most empirically supported, form of psychological intervention. One serious problem, however, is that other forms of psychotherapy, with distinct theoretical foundations, also seem to work. While

there is significant debate over whether other therapies work as well or better, there is little doubt that a number of therapies with seemingly disparate theoretical groundings (most notably psychodynamic approaches), are also effective psychological treatments. For example, Grissom's (1996) meta-analysis of 32 meta-analyses and Luborsky et al.'s (2002) meta-analysis of 17 meta-analyses found statistically insignificant differences between effect-sizes between all theory-based treatments. These findings are consistent with meta-analyses by Wampold et al. (1997) and Assay and Lambert (1999). The accuracy of these meta-analyses is also supported by a number of direct comparisons between CBT and psychodynamic approaches that claim no statistically significant differences between the two approaches (Cuijpers et al., 2010; Wampold et al., 2002). CBT may work, but given that other psychological approaches work as well, CBT's efficacy is not strong evidence for its distinct theoretical model.

In response to the apparent lack of statistically significant differences in success rates between CBT and other theory-based psychological treatments, a number of theorists have attempted to identify common factors that underlie the seeming disparate treatments (Frank, 1991, Wampold, 2015, Messer & Wampold, 2006). CBT theorists, and most notably A. Beck (Alford & A. Beck, 1998; Clark & Beck 1999; A. Beck, 2004), have argued that the process of cognitive change identified by the cognitive model is the primary causally efficacious common factor found in effective psychological treatments. According to A. Beck, "a common denominator of the various systems is the ascription of cognitive mechanisms to the process of therapeutic change…[I]mprovement in the clinical condition is associated with changes in cognitive structuring of experience irrespective of the type of therapy" (2000). The idea here is that any therapy that works does so insofar as it changes how we think about the world. However, even if

it is the case that cognitions play a primary role in the efficacy of therapeutic treatments, this does not mean that the CBT model is accurate. The CBT model maintains that identifying and challenging the *specific* thoughts and core beliefs that are the primary causes of the patient's symptoms is the agent of change in psychological interventions (J. Beck, 2011; A. Beck, 1979). While it is possible that other treatments such as psychodynamic interventions work by way of indirectly challenging specific core beliefs, this paper has argued that this is likely not the case. Rather, it seems that challenging a patient's thoughts helps give her new, adaptive ways to conceptualize her mental life regardless of what specific thoughts or beliefs she previously held.

2.6 Conclusions

CBT works, but likely not for the reasons given by the CBT model. CBT is based on the cognitive model of psychological functioning which postulated a nested hierarchy of consciously accessible cognitions consisting of automatic thoughts, intermediate beliefs, and core beliefs. However, there is good reason to be skeptical that core and intermediate beliefs are accurately and reliably consciously accessible or that they exist in the form postulated; they may serve as useful descriptions or ways to conceptualize psychological illness, but patients are likely not accurately identifying causally efficacious cognitive structures. Furthermore, it is likely that the Socratic method and downward arrow techniques proscribed by the CBT model lead to confabulation rather than accurate identification of dysfunctional or maladaptive automatic thoughts, and thus identification of deeper logically connected cognitions. While it may be the case that changes in cognitive processing are the basis of successful therapeutic treatments, the specific model posited by CBT theorists is likely false.

There remains the question that if CBT does not work by accurately identifying and challenging beliefs and thoughts, why does it work? There are a number of possible answers. First, CBT may work, not by accurately challenging specific cognitive content, but by challenging maladaptive cognitive processes. Recent cognitive theories have argued for a change of therapeutic focus from the cognitive content of automatic thoughts and schemas, to thoughts about thinking. Mindfulness-based cognitive therapy (Segal et al, 2004, 2012) maintains that affective change is not just about changing content of depressive thinking, but also about changing one's relationship to one's thoughts. Mindfulness-based therapies posit that it is the change in one's perspective about one's negative thoughts, rather than challenging the thoughts themselves, which leads to direct and lasting change in psychotherapy. Related views can be found in Dialectical Behavioral Therapy (DBT) and Acceptance and Commitment Therapy (ACT), which both focus implicitly on "decentering" one's relationship to one's cognitions (Segal et al., 2012). Similarly, Meta-Cognitive Therapy (MCT), developed by Wells (2009), focuses on metacognitions (or "beliefs about thinking") rather than on specific cognitive content. MCT "proposes that disturbances in thinking and emotion emerge from metacognitions that are separate from these other thoughts and beliefs emphasized in CBT" (Wells, 2009). Instead of challenging the content of specific core beliefs or automatic thoughts, MCT aims to challenge the beliefs about thinking (e.g., "if I worry about my symptoms, I won't miss anything important") from which these other cognitions are supposedly derived (Wells, 2009).

While these meta-cognitive and decentering approaches are offered as rivals to CBT, the differences may be merely superficial. Both MBCT and MCT share many of the same theoretical

commitments about cognitive primacy and therapeutic focus on consciously accessible cognitions. MCT, like CBT, assumes that consciously accessible thoughts or beliefs play a primary role in therapy and focuses on accurately identifying and challenging beliefs about beliefs (rather than CBT's focus on beliefs about the world, self, and future). And MBCT, like CBT and MCT, aims at altering patients' perspectives on their negative cognitions; MBCT focuses on decentering and detaching oneself from one's thoughts, while CBT aims to challenge patients' views about the rationality or validity of their thoughts. Both MBCT and CBT work by patients identifying specific cognitions; they differ only in how the patient is instructed to treat these beliefs. It is unclear, then, whether the new theoretical and therapeutic focus on thoughts about thoughts offers a genuine theoretical challenge to the CBT model.

Another plausible explanation is that CBT may work for the same reasons that other effective therapies work; CBT fosters a challenging and caring therapeutic allegiance between patient and practitioner and offers a plausible explanation and method of treatment for the patient's problems. The "common factors" theory postulates that non-specific (to any given theory) common factors (such as a healing setting, a coherent theory/rationale, a healing ritual, and an emotionally charged confiding relationship) play a dominant role in psychological change (as opposed to the specific factors postulated by distinct theories) (Frank & Frank, 1991; Anderson et al., 2010; Messer & Wampold 2006). This response, while plausible, is underdeveloped. It still must be explained why these factors are necessary for successful therapy and what these common factors have in common. Most common factors theorists take as their model Jerome Frank's idea that therapy is best understood as a form of rhetoric (Frank, 1961; Frank & Frank, 1991). According to Frank, psychological healing is a matter of persuasion with the common

factors being necessary components. What is left unexplained, and what the cognitive model purports to answer, is why persuasion is the mechanism of change in psychotherapy. Rather than being a rival to the CBT model, the common factors approach implicitly assumes something like a cognitive model of psychopathology; therapy, like rhetoric, is just supposed to be a matter of convincing the patient to accept more adaptive beliefs. Therapy may work by imposing coherence upon a person's mental life, but it still must be explained why and how this might work.

Chapter 3: What Does It Mean To Have A Meaning Problem? Meaning, Skill, and the Mechanisms of Change in Psychotherapy

3.1 Introduction

Psychotherapy is effective. Since the 1970's meta-analyses, and meta-analyses of meta-analyses, have consistently shown a significant effect size for psychotherapeutic interventions when compared to no treatment or placebo treatments (e.g. Smith & Glass, 1977; Luborsky et al., 2002; Wampold et al., 1997). This effectiveness is normally taken as a sign of the scientific legitimization of clinical psychotherapy. A significant problem, however, is that psychotherapies with distinct, and often incommensurate, theoretical foundations appear to be equally effective. While individual studies directly comparing therapies, or comparing therapies to placebos, often show the superiority (if often only minor) of one particular therapy over another, meta-analyses of clinical studies consistently show a general lack of statistically significant differences between the outcomes of most forms of standardized psychological interventions (e.g., Bergin & Garfield, 1994; Hubble, Duncan, & Miller, 1999; Lambert, 2013; Wampold & Imel, 2015). This poses a problem for specific psychotherapies: they may work, but likely not for the reasons that ground their theoretical explanations for their effectiveness.

The two prominent types of explanation for the finding of common therapeutic efficacy have been to either (1) challenge the accuracy and/or methodology of meta-analyses that purport to show an equivalence of effectiveness across different therapies (e.g., Crits-Christoph, 1997;

Butler et al., 2005; Marcus et al., 2014), or (2) to attempt to identify underlying common factors that would explain the common efficacy of seemingly disparate therapeutic techniques and theories (e.g., Rosenzweig, 1936; Frank & Frank, 1991; Wampold & Imel, 2015). While most meta-analyses support the common efficacy findings, not all do. The type (1) explanations normally accept that different therapies may be efficacious, but argue that specific therapies appear to be superior in the treatment of one or more disorders. For example, a meta-analysis by Cuijpers et al. (2008) found interpersonal therapy for depression to be 'somewhat more efficacious than other treatments' (p. 917), though it found no statistically significant difference between most of the other forms of treatment (including CBT, psychodynamic therapy, and problem-solving therapy). Tolin's (2010) meta-analysis found CBT for depression and anxiety to be superior to psychodynamic treatments, but found no significant difference between CBT, interpersonal therapy, and supportive therapy. Similarly, a recent meta-analysis by Marcus et al. (2014) found CBT to be 'slightly more effective' (p. 519) than other therapies in treating primary symptoms, though it found no differences between therapies when assessing secondary outcomes (e.g., quality of life).[21] The dispute over the accuracy of type (1) explanations shows no sign of a resolution, but even if it is the case that some therapies are found to be more successful in treating some psychological maladies, it still must be explained why most treatments are effective (even if not equally effective) and why many forms of psychological disturbances respond equally as well to different, and often theoretically incompatible psychotherapies.

[21] Though, see Wampold et al., (2017) for a criticism of these findings.

The most common type (2) explanation has been to postulate that non-specific (to any given theory) common factors (such as an emotionally charged confiding relationship, a healing setting, and a coherent theory/rationale) explain psychological change (as opposed to the specific factors postulated by distinct theories) (e.g., Frank & Frank, 1991; Messer & Wampold, 2002; Miller et al., 2005). According to this view, theoretically and functionally distinct therapies such as cognitive behavioral therapy (CBT) and psychodynamic therapies are supposed to be efficacious only because they share particular therapeutic ingredients common to all efficacious therapies. The empirical or theoretical 'truth' of the particular delivery method is taken to be irrelevant; all that matters is that the therapy succeeds in delivering the common factors that lead to psychological healing (e.g., Frank, 1995; Wampold, 2001).

The so-called 'common factors' theories, while plausible, are also in need of a model of the relation between the common factors and therapeutic change; they need to explain why the common factors are supposed to enable psychological healing. According to influential common factors models proposed by psychologists Jerome Frank (Frank, 1961; Frank & Frank, 1991), Bruce Wampold (Wampold, 2001; Wampold & Imel, 2015), and Aaron Beck (Beck, 1987, 1991, 2004; Alford & Beck, 1998), psychotherapies are supposed to work by altering maladaptive meanings and providing patients with new, salubrious, and more adaptive meanings.[22] These models share the assumptions that the alteration of meanings is the primary mechanism of

[22] For similar common factor views, see: Miller, Duncan, & Hubble (2005); Lambert (1992); Orlinsky & Howard (1986); Goldfried (1980); Anderson, Lunnen, & Ogles (2010).

change in psychotherapy and that the problem being addressed in psychological interventions is primarily a problem of maladaptive meanings.

This essay will address three interrelated philosophical and theoretical questions concerning the 'meaning theory' of psychological change. First, what does meaning have to do with psychopathology? Is psychopathology a problem of meaning, or is it merely ameliorated in part by a meaning-based solution? And finally, what is supposed to be maladaptive about the meanings being altered in psychotherapy (and what is adaptive about the meanings that replace them)? This essay argues that the meaning theory of psychological change is likely mistaken; psychological disorders are not normally problems of meaning nor are they directly ameliorated by changes in meaning. Rather, psychotherapeutic change is best explained by the development of the patient's self-regulatory skills. According to the skill view outlined here, the therapeutic common factors are effective only insofar as they help enable skilled action.

3.2 The Common Factors Theories

The findings of common psychotherapeutic efficacy has been christened the 'Dodo bird' effect, after the psychiatrist Saul Rosenzweig's (1936) reference to the Dodo's pronouncement in *Alice in Wonderland* that: "Everybody has won and all must have prizes". Rosenzweig's Dodo bird claim was directed toward the apparent lack of differences in outcomes of rival therapies despite the proliferation of theories and inter/intra-theoretical disputes. Rosenzweig also offered the first attempt at an explanation for the Dodo bird findings: effective therapies are likely effective because "there are inevitably certain *unrecognized* factors in any therapeutic situation— factors

that may be even more important than those being purposefully employed" (1936, p. 412). In the recent psychological literature, there have been three significant attempts to identify these 'unrecognized factors' and explain the common efficacy of psychotherapies: Frank's persuasion theory, Wampold's contextualist model, and Beck's integrative cognitive theory.

The most influential starting point for explaining the Dodo bird effect is Jerome Frank's claim that therapy is best understood as a form of rhetoric (Frank, 1961; Frank & Frank, 1991). According to Frank, all successful psychological healing systems share four non-specific (to any given theory) common factors: an emotionally charged confiding relationship, a healing setting, a healing ritual, and a coherent theory/rationale. Psychological healing is conceived of as a matter of persuasion with the common factors being necessary enabling components. According to Frank (1991):

> the aim of psychotherapy is to help people feel and function better by encouraging appropriate modifications in their assumptive worlds, thereby transforming the meanings of experience to more favorable ones (p. 30).

Frank's 'assumptive worlds' are supposed to be meaning-making interpretations of all external and internal stimuli based on "assumptions about what is dangerous, safe, important, unimportant, good, bad, and so on" (1991, p. 24). These assumptions (both conscious and unconscious) form highly structured attitudes, beliefs, and behavioral dispositions. Frank claims that individuals seek psychological help in order to combat 'demoralization', which results from unhealthy, unfavorable, or somehow maladaptive assumptive worlds. To be demoralized, in this

context, is "to deprive a person of spirit, courage, to dishearten, bewilder, to throw a person into disorder or confusion" (1991, p. 35). The four hypothesized common factors are supposed to promote healing by restructuring the patient's 'assumptive world' to a more favorable and remoralized state.

Frank's persuasion theory is clear about the source of psychological problems: psychopathology is primarily caused by maladaptive meanings (or demoralized assumptive worlds), and psychotherapy, when effective, is effective because it remoralizes patients by challenging their maladaptive meanings and replacing them with more adaptive ones. According to Frank (1991):

> effective psychotherapies combat demoralization by persuading patients to transform these pathogenic meanings to ones that rekindle hope, enhance mastery, heighten self-esteem, and reintegrate patients with their groups (p. 52).

Wampold's contextual model is derived from, and meant to be largely consistent with, Frank's common factor theory.[23] The "basic premise" of Wampold's model is that "the benefits of psychotherapy accrue through social processes and that the [clinician/patient] relationship, broadly defined, is the bedrock of psychotherapy effectiveness" (2015, p. 50). The idea here is that psychological healing is an inherently interpersonal process, with the common factors

[23] The difference between Frank's and Wampold's theories is minimal. The primary differences between the theories are the number of common factors (i.e., Frank's four compared to Wampold's three), and Wampold's greater emphasis on the holistic nature of the common factors. of the therapeutic 'real relationship'. According to Wampold, Frank's common factors model '[contains] a set of common factors, each of which makes an independent contribution to outcome…in a contextual conceptualization of common factors, specific therapeutic actions…cannot be isolated and studied independently' (2001, p. 26).

playing a necessary enabling role in establishing and maintaining the patient/clinician relationship. The effectiveness of psychotherapy is supposed to be explained by three necessary common factors that influence the healing relationship: the 'real relationship', client expectations, and specific ingredients (i.e., specific psychotherapies). The real relationship is defined as an intimate personal relationship between the therapist and the patient, marked by empathy and caring on behalf of the healer and trust and emotional openness by the patient. The alteration of the patient's expectations is considered necessary for psychological change insofar as it instills hopefulness (or remoralization) and, most importantly, because it challenges the patient's folk-explanation for her psychological distress. According to this model, successful psychotherapy "provides an explanation for the client's difficulties that is adaptive in the sense that it provides a means to overcome or cope with the difficulties" (2015, p. 58). The specific ingredient (i.e., the particular therapy) is claimed to be a necessary ingredient in successful therapy because it provides the patient with a coherent explanation for her problem and a cogent rationale for the expected healing process. As with Frank's common factors theory, Wampold's contextualist model posits that the truth or empirical status of the particular therapeutic ingredient is irrelevant to the healing; all that is necessary for psychological healing is that the patient accept the theoretical rationale, emotionally connect with the therapist, and adopt a more salubrious explanation for her psychological problem.

Wampold's contextual model, like Frank's persuasion theory, conceives of psychotherapy as a process of challenging and modifying patient-meanings. The ameliorative effects of psychotherapy are supposed to be explained by modifications of the meaning of patients' folk-psychological beliefs about their psychological functioning. The therapeutic 'real' relationship,

patient expectations, and a novel theoretical rationale are claimed to be essential aspects of psychotherapy because they enable the patient to construct or adopt new, more adaptive explanations of the meaning of his or her psychological problem. According to Wampold:

> The essential aspect of psychotherapy is that a new, more adaptive explanation is acquired by the patient…what is critical to psychotherapy is understanding the patient's explanation (i.e., the patient's folk psychology) and modifying it to be more adaptive (2007, p. 862-3).

This new, 'adaptive' explanation is supposed to be salubrious because:

> the contextual model states that the treatment procedures used are beneficial to the client because of the meaning attributed to those procedures rather than because of their specific effects. (2001, p. 27)

Wampold's contextualist model is less explicit than Frank's about identifying the primary causes of psychological maladies (as opposed to the causes of psychological healing), but is similarly based on the assumption that the psychological problems being addressed by therapy are problems of meaning. The contextual model frames psychological healing as process of altering patients' 'explanations' of their mental functioning from the (somehow) maladaptive to the adaptive. According to Wampold:

> it is my contention that the patient's idiosyncratic explanations of mental functions are deeply involved in creating the patient's problems, that psychotherapy is intimately involved in altering these explanations (2007, p. 862)

Similarly, in reference to the 'real relationship', Wampold claims that:

> a critical component of how [a good patient/therapist relationship] leads to change is involved in replacing a maladaptive explanation with an adaptive one. The maladaptive explanation is discouraging because the client cannot see how any action will lead to progress: Put simply, they are stuck (2010, p. 70).

The contextualist model, like the persuasion model from which it was built, assumes that the alteration of maladaptive meanings are the primary mechanism of change in psychopathology. Therapy is supposed to help patients become 'unstuck' by challenging and replacing their maladaptive meanings and explanations with more salubrious interpretations.

The third major common factors theory is derived from Beck's cognitive-behavioral therapy (CBT), the dominant form of psychotherapy in North America. According to this view, if other forms of therapy are effective, it is only because they are (either implicitly or explicitly) doing cognitive therapy (Beck, 1987; Alford & Beck, 1998; Beck, 2004). Note that the conditional in the last statement should be taken as truly conditional; Beck and colleagues routinely publish studies that purport to show the clinical and theoretical superiority of CBT over other forms of psychotherapy (e.g., Beck & Dozois, 2011; Butler et al., 2006; Beck, 2005). At its most basic,

the cognitive model of psychotherapy posits that cognitions (i.e., thoughts, beliefs, and assumptions) play the primary role in the development and treatment of dysfunctional psychological states (Beck, 1967, 1979; Clark & Beck, 1999; J. Beck, 2011). Cognitive therapy aims at identifying and challenging maladaptive thoughts and beliefs (either directly through introspection and talk therapy or indirectly through behavioral change) and replacing them with more adaptive interpretations.

The alteration of patient-meanings plays a central role in Beck's cognitive theory of psychopathology and psychological change. According to the cognitive model:

> Psychopathology results from maladaptive meanings constructed regarding the self, the environmental context (experience), and the future (goals), which together are termed the *cognitive triad.* Each clinical syndrome has characteristic maladaptive meanings associated with the components of the cognitive triad (Alford & Beck, 1997, p. 17).

Cognitive theory is grounded on the assumption that 'information processing', or the transformation of endogenous and exogenous stimuli into meaningful representations, is the primary function of the human mind (Clark & Beck, 1999). Clark & Beck (1999) state that, "the central tenet of the cognitive model is that human information processing or meaning construction influences all emotional and behavioral experiences" (p. 55) and that the "modification of meaning-assignment structures is central to the human change process" (p. 70). According to the cognitive model, psychological disorders are the result of maladaptive schemas

(i.e., the "basic structures that integrate and attach meaning to events") and modes (i.e., interconnected clusters of schemas) (Alford & Beck, 1997, p. 36).[24]

The cognitive model posits that how an individual's 'core schemas' (used interchangeably with 'core beliefs') organize and process incoming stimuli determines how that individual organizes and conceptualizes his or her "personal construction of reality" (Clark & Beck, 1999, p. 60). Therapy is supposed to be successful because it challenges the patient's specific maladaptive core beliefs (e.g., "I am unlovable" or "the world is unsafe"), either directly (as in CBT interventions) or indirectly (as in other non-cognitive, but efficacious, treatments), and provides patients with more adaptive and salubrious interpretations of the world

According to Beck's integrative cognitive theory, the modification of maladaptive schemas is supposed to be *the* common factor found in effective psychological treatments. Alford and Beck (1997) state that:

> "the analysis of the therapeutic components and procedure of psychoanalysis, behavior therapy, and other systems of psychotherapy suggests one common factor— the modification of core beliefs or schemas' (p. 99).

[24] Beck's use of the term "schema" is often ambiguous. Beck normally uses the terms "schema" and "core belief" interchangeably. See: Beck et al., 1979, p. 4; Beck, 1991, p. 195; Beck, 1997, p. 58; Clark & Beck, 1999, pp. 82-83; Wenzel, Chapman, Newman, Beck, & Brown, 2006, pp. 504-505; DeRubeis, Webb, Tang, & Beck, 2010, pp. 280-281; Clark & Beck, 2010, p.36). Beck also sometimes uses the term belief to mean the content of schema (e.g., Beck, Davis, & Freeman, 2015, pp. 30-33).

Of course, most theories accept that at some level therapy changes the way we think; cognitive theory is making the much more substantive claim that groupings of specific, consciously accessible maladaptive beliefs, thoughts, and assumptions (e.g., 'If I'm not a success, I'm a failure') are the primary mechanisms of both psychological dysfunction and change (Beck, 1991; Alford & Beck, 1997). Cognitive theory posits that "individuals can become aware of the content and processes of their thinking", and that psychotherapy is effective when it enables patients to identify and challenge their maladaptive cognitive content and processes and "shift their cognitive appraisals from one's that are unhealthy and maladaptive to ones that are evidence-based and adaptive" (Beck & Dozois, 2011, p. 400).

Despite Beck's claim that cognitive modification is the sole common factor, there is significant overlap with both Frank and Wampold's theories. Beck's integrative cognitive theory acknowledges that remoralization (or in Beck's terms, "expectations for improvement") is typically an important constituent of effective therapy (Alford & Beck, 1997, p. 45), and it also accepts that the therapeutic relationship is a "major vehicle for improvement" (Alford & Beck, 1997, p. 48). It differs from the other two theories, however, in claiming "cognitive primacy". According to Beck's cognitive theory, "all other psychological processes are explained by means of cognitive concepts…cognition alone provides meaning (or coherence) to the various other basic psychological processes" (Alford & Beck, 1997, p. 45). The other putative common factors, then, may *influence* healing, but they are supposed to be salubrious only insofar as they help enable the alteration of patients' maladaptive core beliefs or schemas.

3.2.1 Meanings in Psychotherapy

What the three common factors theories have in common is a shared assumption that the alteration of patient-meanings is the primary mechanism of change in psychotherapy. The common factors (whatever they may be) are supposed to be necessary constituents of successful psychotherapy only insofar as they enable patients to alter their maladaptive thoughts, beliefs, or attitudes. The theories differ (if only slightly) in identifying the specific common factors that are supposed to enable this change in meaning.

The three models also share a belief-based conception of 'meaning'; they all assume that the meanings being altered in therapy are, or are determined by, doxastic attitudes. According to Frank, "meanings are determined by an organized set of assumptions, attitudes, or beliefs…that we have termed the assumptive world" (1991, p. 50). According to Wampold, patients' meaning-making explanations for their psychological problems are either adaptive or maladaptive mental states "formed from their own psychological beliefs" (Wampold & Imel, 2015, p. 58). And, according to Beck, meaning-making schemas are "attitudes, beliefs, and assumptions which influence the way an individual orients himself to a situation, recognizes and labels the salient features, and conceptualizes the experience" (1964, p. 561). Successful therapy, then, is just supposed to be a matter of convincing the patient to accept more adaptive doxastic states.

So, according to the three common factors theories, the alteration of maladaptive patient-meanings is supposed to explain why the common factors lead to salubrious psychological change, and, thus explain the cause of the Dodo bird findings of the general efficacy of

psychotherapy. The second half of this paper argues that this view is likely mistaken, and outlines a skill-based alternative to the meaning-based explanation of psychological change.

3.3 The Problems with Meaning

The main common factors theories assume that meanings are the primary mechanisms of change in psychotherapy. What is left unclear from this explanation is why certain meanings or explanations are harmful or maladaptive. What exactly is maladaptive or harmful about the meanings being challenged and replaced? And what is it about 'adaptive' meanings that is supposed to be ameliorative? The four most plausible explanations given by the common factors theories are that the content, valence, hopefulness, and dominance of the meanings explains whether they are adaptive or maladaptive, and thus why therapy is effective. This section considers, and rejects, all four answers. Alterations in patient-meanings clearly play *a* role in explaining psychological healing, but a further variable is needed to explain why changes in meaning (be they changes in content, valence, hopefulness, or dominance) may be salubrious.

3.3.1 The Meaning Content Hypothesis

The common factors theories all focus to some extent on the *content* of patient-meanings as a target of psychotherapeutic interventions. Beck's CBT aims to accurately identify, then challenge, specific maladaptive core beliefs (e.g., 'my value as a person depends on what others think of me' or 'I should always be at peak efficiency') (Beck, 1979), while both Frank's and

Wampold's models claim that successful therapy replaces patients' maladaptive 'idiosyncratic explanations' or 'assumptive worlds'. Call this the 'Meaning Content Hypothesis':

> **Meaning Content Hypothesis**: specific meaning-content is maladaptive. Effective psychotherapies challenge and alter the maladaptive content of patients' meaning attributions.

This does not work. The common factors theories cannot consistently claim that it is the particular content of the meanings that is maladaptive, while also maintaining that the specific content of the therapy is irrelevant to its success. Different therapies target different meanings. Consider the plurality of meaning contents that different psychotherapies attribute to pathological anxiety. The content being addressed in Freudian treatments range from the patients' perceived loss of some object to subconscious castration anxiety (Freud, 1936). Modern psychodynamic psychotherapy targets patients' unacknowledged rage and subconscious fantasies (Busch et al., 1999). Radical acceptance therapy targets patients' self-judgments and unworthiness-centered belief content (Brach, 2003). While cognitive therapy for anxiety focuses on directly challenging threat-based content (Clark & Beck, 2011). According to the Dodo bird findings, these therapies are all supposed to be effective in treating anxiety; everyone wins and all get prizes. And according to the common factors theories, this common efficacy is explained by a change in meaning. But if specific meaning-contents are the variables of control in psychotherapy, and different therapies target different meanings, then everyone should not be winning.

According to the common factors theories, the content of a specific therapy is supposed to be irrelevant to its success. For example:

> The criterion of the 'truth' a psychotherapeutic interpretation, as of a religious text, is its plausibility. The 'truest' interpretation would be one that is most satisfying or that makes the most sense to the particular person or interpretative community (Frank, & Frank, 1991, p. 73).

> The truth of the explanation is unimportant to the outcome of psychotherapy. The power of the treatment rests on the patient accepting the explanation rather than whether the explanation is 'scientifically' correct (Wampold, 2007, p. 863).

> Techniques from diverse systems of psychotherapy (cognitive, behavioral, psychodynamic, humanistic, and experiential) enable patients to disconfirm the basic dysfunctional beliefs embodied in the dysfunctional schemas…Regardless of the approach to cognitive modification (direct or indirect), the dysfunctional beliefs that are activated during acute episodes of a disorder are no longer found when the episode is over (Alford & Beck, 1997, p. 99).

Both the Freudian focus on sexuality and subconscious wishes and the CBT-based focus on consciously accessible thoughts and beliefs are supposed to be effective because *despite their differences in content,* they are hypothesized to share some underlying common factor(s). Similarly, the specific content of the maladaptive patient-meanings being treated in any

particular therapy, as well as the content of the successful patient's new, and more adaptive meanings, should be irrelevant to explaining the ameliorating powers of the common factors (and successful therapies in general). The common factors theories are committed to the claim that it is immaterial to the success of therapy, and thus to patients' mental health, whether patients' specific meaning attributions are focused on existential terror, perceived loss, thoughts of unworthiness, learned behavioral rules, spiritual closeness to a creator, or whatever. Whether one views the world through a Freudian inspired worldview (i.e., one constituted in part by subconscious, non-cognitive drives and desires) or a CBT inspired folk-psychological theory of mind (i.e., one that assumes a theory of cognitive functioning that includes consciously accessible, hierarchically structured, and logically connected thoughts and beliefs) is supposed to have no direct influence on one's psychological health. Similarly, whether a patient attributes her psychological distress to repressed subconscious memories of childhood trauma or to consciously accessible maladaptive core beliefs is, according to the common factors models, irrelevant to the patient's mental health. The common factors theories are committed to the claim that particular theories of psychological functioning are only relevant to psychological health insofar as they offer the patient *a* coherent explanation for her problem; the truth of the explanation, and its particular content, is supposed to have no necessary relation to mental health.

One option available to an advocate of the content hypothesis is to claim that while different therapies may target different thoughts, beliefs, and assumptions, they are only effective because they are either directly or indirectly altering specific patient-meanings. Beck's integrative cognitive theory takes this view. So, for example, while a psychodynamic treatment of panic

disorder may focus on uncovering patients' unconscious "compromise between angry feelings and fantasies and fears of abandonment" (Busch et al., 1999, p. 235), and CBT for panic disorder may focus on uncovering and challenging patients' "catastrophic misinterpretation of bodily sensations" (Clark & Beck, 2011, p. 292), both forms of psychotherapy are supposed to be effective because both are challenging specific maladaptive core beliefs such as "heart palpitations are dangerous" and "I could suffocate to death". CBT is supposed to do this directly, while psychodynamic (and all other efficacious) therapies challenge these specific beliefs indirectly. However, this approach doesn't work either. First, this option is available to all theories of psychotherapy; *any* theory can claim that all other efficacious therapies are effective only because the content of these therapies is translatable (however tortuously) to the content of the therapy in question. Second, and most seriously, the translation claim doesn't explain why the direct or indirect alteration of specific patient-meanings appears to make *no difference* to the success of therapy. If the alteration of specific patient-meanings is the mechanism of change in psychotherapy, then it is mystery why the success of psychotherapy should have no necessary relation to the actual thoughts, beliefs, and assumptions being explicitly altered in therapy.

3.3.2 The Meaning Valence Hypothesis

The common factors theories, then, also need an explanation for why the disparate meanings being changed in therapy are health-promoting. One option is to claim that the *valence* of a patient's meanings is the locus of both treatment and development in psychopathology. Call this the 'Meaning Valence Hypothesis':

> **Meaning Valence Hypothesis**: the alteration of the valence of patient-meanings (e.g., from negative to positive) explains therapy's common efficacy.

This, on the surface, may seem plausible and seems to be assumed by all three theories. Both Frank (1991) and Wampold (2007) claim that patients come to therapy because they are demoralized, while Beck states that mental disorders are marked by positively or negatively polarized core beliefs (Clark & Beck, 1999). Indeed, it is well established that positively or negatively valenced thoughts are associated with emotions valenced in corresponding directions. Negative thoughts are correlated with negative feelings and low mood while positive thinking in correlated with an increase in positive mood and the experience of well-being (Clark & Beck, 1999; Fredrickson et al., 2008). It may be thought, then, that valence of one's meanings (whatever their specific content) is the primary cause of psychopathology and the primary control variable in psychological healing. Negatively valenced meanings (e.g., 'I am unlovable') may be the primary cause of maladaptive psychological states such as depression, while positively (or neutrally) valenced meanings may be the primary cause of adaptive (or at least non-maladaptive) psychological states. According to this view, then, the alteration of the valence of patient-meanings (e.g., from negative to positive) explains therapy's common efficacy. However, while plausible, this is likely not the case; the valence of meanings may affect mood, but it does not explain why low or high mood becomes pathological.

The problem here is that the valence of a person's meanings does not explain why some meanings are supposed to be pathological and others healthy. Negative or pessimistically valenced meanings do not necessarily lead to depression, despair, or demoralization, and

positively valenced meanings are not necessarily salubrious. For example, a number of studies show Asian-Americans to be significantly more pessimistic (though not less optimistic) than European-Americans (e.g., Chang, 1996, 2002; Hardin & Leong, 2005). However, this significant difference in experienced pessimistic meaning-attributions does not correlate with an increase in maladaptive psychological states (such as depression or anxiety); Asian-Americans are just as mentally healthy (or ill) as the national average (e.g., Chang, 2002). Similarly, nihilistic philosophers (or any adherent to world views that deny the existence of inherent meaning in life) and futurists who forecast coming centuries to be constituted by an inevitable human destruction of the planet, are, as far as anyone can tell, not necessarily mentally ill despite spending much of their mental lives thinking about, and endorsing, negatively valenced ideas. Pessimism and negatively valenced thoughts may be correlated with pathological low mood, but the valence itself does not explain the pathology. Negatively valenced meaning-attributions about the world, self, or future may affect one's happiness or mood, but they needn't have any effect on one's mental health.

The converse is also true: positive meaning-attributions needn't lead to adaptive psychological states. Manic episodes are diagnosed in the DSM-5 in part by the symptoms of persistent elevated mood, elevated self-esteem, and extreme goal-directed behavior (APA, 2013). Individuals experiencing manic states often describe them as exhilarating, hopeful periods of optimism and high self-regard, and these symptoms can range from weeks to months in duration. However, a person experiencing a manic episode is not normally considered psychologically well-functioning just because she is experiencing positively valenced thoughts about herself, her

world, and her future. On the contrary, in many cases (such as in the case of bipolar disorders) positive-valenced meaning content is not salubrious, and is in fact maladaptive.

It is certainly true that the valence of a person's meanings is often strongly correlated with the adaptiveness of her or his psychological state. It would be surprising if a person diagnosed with major depression had frequent positively valenced thoughts about herself and her world. But it is a mistake to conflate the valence of a person's meanings with the primary cause for her dysfunction or to consider it the primary variable of control in therapy. Believing that the future is doomed does not make you pathologically depressed, just as believing that you are wonderful and the future is full of opportunity (as one might think while in a manic state) does not make you non-pathological. The valence of an individual's meanings may lead to mood changes, but not, by itself, to pathological mood or psychological health. Something besides the valence or the content of meaning attributions is needed to explain the Dodo bird effect.

3.3.3 The Hope Hypothesis

One possible explanation, similar to the valence hypothesis, is to claim that meaning change is salubrious, at least in part, because it raises patients' expectations and instills hopefulness. Call this the 'Hope Hypothesis':

> **The Hope Hypothesis**: meaning change is salubrious, at least in part, because it raises patients' expectations and hopefulness.

Frank's and Wampold's common factors theories both adopt this view.[25] According to Frank, this new hope helps to remoralize the patient, while Wampold claims that raised expectations influence patients to accept new, more adaptive explanations for their psychological problems. Both Wampold and Frank are likely correct in arguing that hope and expectations of improvement are important parts of successful therapies. An important aspect of any therapy is motivating the patient to participate, and hope often is a powerful motivating force. But there is a significant disanalogy between hopefulness and mental health: expecting, hoping, or believing that one is psychologically healthy does not make one healthy. Individuals dealing with common maladaptive psychological issues such as mania or narcissistic personality disorder often think they are fine, or even great, and often have very high expectations about their future. Mania, as noted above, is marked by what we can call *super*-moralization (as opposed to demoralization), while pathological narcissism is marked by grandiose views of one's self, future, and place in the world (APA, 2013). Hopeful and super-moralized cognitions are often part of the problem being treated in psychotherapy (Gruber et al., 2011; Gruber, 2011; Greenhouse et al., 1999). A significant difficulty in treating these psychological problems is convincing the patients that they have a problem that needs treating in the first place. If positive and hopeful expectations is the explanation for improved mental health, then we should expect hopeful individuals to be psychologically well functioning; but this is often not the case. Successful psychological interventions may, and likely very often do, raise the expectations of patients and heighten their perceived sense of self-efficacy, but this change in expectations cannot be the primary control

[25] Beck's cognitive theory acknowledges that increased expectations play a significant role in successful therapy, but claims that these expectations are salubrious only insofar as they help engender changes in core schemas (e.g., Beck et al., 1979; Alford & Beck, 1997).

variable in psychological healing given that that hopefulness and positively valenced expectations are not always a good thing for mental health.

3.3.4 The Meaning Dominance Hypothesis

A final meaning-based explanation for the efficacy of the common factors, found to varying degrees in all three theories, is that maladaptive meanings (somehow) exclude other, more adaptive, meanings and come to *dominate* the mental life of individuals experiencing mental disorders. The idea, here, is that successful therapy enables a more balanced mental life. Call this the 'Meaning Dominance Hypothesis':

> **The Meaning Dominance Hypothesis:** Maladaptive patient-meanings (somehow) dominate an individual's mental life during instances of mental disorder. Meaning change is salubrious, at least in part, because it enables the inclusion of other, more adaptive, patient-meanings.

Frank clearly endorses this view. He claims that maladaptive assumptive worlds are 'resistant to change' and biased towards 'confirmatory experiences', and argues that successful therapy helps patients modify these meaning-attributions to be more harmonious and adaptable to changes in circumstance (1991, p. 32, pp. 50-51). Wampold's contextualist theory is consistent with the dominance hypothesis, but does not explicitly endorse this claim. As noted, the contextualist theory argues that individuals seeking therapy have idiosyncratic folk-explanations for their distress that leave them 'stuck' (Wampold, 2010, p. 70). It is consistent with this view to hold

that these maladaptive explanations are intransigent because they dominate or exclude other, more adaptive, explanations. Beck explicitly endorses the dominance hypothesis. According the Beckian cognitive theory, mental disorders are marked 'hyperactive idiosyncratic schemas' that '[because] of their greater strength…displace…more appropriate schemas' (Beck, 1967, p. 286). In the case of depression, 'specific idiosyncratic schemas assume a dominant role in directing the thought processes' (Beck, 1964, p. 564), and these 'depressive schema [are] so potent that the patients are unable to energize other schemas sufficiently to offset its dominance' (Beck, 1967, p. 286). Successful therapy is supposed to identify, then challenge these dominant maladaptive schemas, thus weakening the strength of maladaptive modes of thinking and allowing for more 'appropriate' core beliefs to structure one's interpretations of incoming stimuli.

There is some initial plausibility to this hypothesis. Indeed, an intuitive way of describing mental disorders such as depression, anxiety, or obsessive disorders is as mental states that are (somehow) dominated by specific doxastic states (e.g., depressive beliefs, anxious or worried thoughts, or obsessions). Successful therapy, then, would just be a process of weakening the strength of certain patient-meanings to enable more adaptive meanings to structure how individuals interpret the world. The problem, however, is that, if tenable, the dominance hypothesis cannot be a claim about dominant *meanings*. If this hypothesis is supposed to be an explanation for why the doxastic states being altered in psychotherapy are maladaptive, and why the new doxastic states arrived at through therapy are salubrious, then the dominance hypothesis runs into the same problems as the content, valence, and hope hypotheses. Namely, there is nothing necessarily maladaptive about the 'dominant' beliefs, assumptions, or thoughts being

altered in therapy, and nothing necessarily salubrious about the new meanings arrived at through therapy.

The same meanings may be present and dominant in both disordered and non-disordered states. For example, Beck hypothesizes that the minds of depressed individuals are often dominated by purportedly maladaptive beliefs such as 'If I'm not on top, I'm a flop', 'In order to be happy, I have to be successful in whatever I undertake', and 'It's wonderful to be popular, famous, wealthy; it's terrible to be unpopular, mediocre' (Beck, 1976, p. 255). However, the minds of non-pathological optimists and extremely goal-directed individuals may be described as being dominated by these exact same beliefs. Similarly, both depressive and philosophically nihilistic world views may be dominated by the same 'maladaptive' core beliefs such as 'the future is pointless' or 'I don't see any point to living' (Beck, 1967, p. 12, p. 84). The thoughts of both achievement-oriented optimists and philosophical nihilists can be driven by rigid, change-resistant beliefs that dominate their work and personal lives, yet rather than being pathologized, these mindsets are (at least sometimes) lauded for their single-mindedness and stubbornly-held doxastic states. The problem for the meaning-based dominance hypothesis is that while the predominance of certain beliefs may be associated with specific disorders (e.g., the predominance of negatively valenced beliefs may be associated with depression), the predominance or absence of specific patient-meanings has no necessary relationship with mental health. The dominance of any doxastic state (or states), then, does not explain why some belief states are adaptive, while others are maladaptive.

The common factors theories, then, are missing a tenable explanation as to why the presence of the common factors in psychological treatments is salubrious. Beck, Frank, and Wampold's theories assume that the success of the therapeutic common factors is explained by the alteration of patient-meanings. This section has argued that this view is likely mistaken. While therapy may often succeed by changing how patients attribute meaning to, and conceive of, the world, it still needs to be explained why providing patients with new meanings can be health-promoting.[26] The meaning dominance hypothesis, however, does point us in the direction of a more plausible explanation for the common efficacy of psychotherapy. The final section of this paper argues that dominance hypothesis is right in claiming that mental illness is marked by a 'dominated' mind, the mistake is in positing that this dominance is explained by the strength of particular doxastic states. The dominance hypothesis should not be a claim about meaning, but rather should be a claim about what individuals are *able* to do with these meanings (and mental phenomena, more generally). A more plausible dominance hypothesis will be a claim about skilled action, not meaning. According to this alternative view, the efficacy the common factors, and thus psychotherapy, is best explained not by the alteration of meanings, but by the enabling and development of the patients' skill of regulating how they respond to these meanings.

[26] Note that this paper is not making the (fallacious) argument from (1) 'some people meeting conditions C are not ill', to (2) 'interventions on conditions C cannot, per se, be therapeutic. Rather, the argument here moves from (1) some people meeting conditions C are not ill, to (2) explanations for the efficacy of therapy can't appeal to ameliorating C as the full story as to why folks heal (given that C, by itself, is not a problem). So, interventions on C can be therapeutic, but the explanation for why they are therapeutic (e.g., that they are altering maladaptive meanings), needs an explanation that does not just appeal to the badness of C.

3.4 Alternative Hypothesis: Skill and Psychological Healing

The last section of this paper outlines the beginning of a skill-based explanation for the success of the therapeutic common factors that is compatible with versions of the three major common factors theories. The 'skill hypothesis' argues that the primary mechanism of change in psychotherapy is the patient's skill of self-regulation. According to this view, psychotherapy is best understood as *themed* skill training. What the common therapeutic factors have in common is that they provide patients with the skill to regulate their responses to their thoughts, emotions, and behavior. The explanatory focus here is on the modification of skilled action (such as the alteration of how responds to one's doxastic states), rather than the alteration of the content, valence, hopefulness, or dominance of patient-meanings. If changes in patient-meanings are salubrious it is only because the alteration of meaning allows patients to construct a coherent conceptual framework from which to develop regulatory skill. Psychological healing, according to this view, is skilled action that psychotherapy helps cultivate.

This paper adopts theory-neutral conceptions of skill and self-regulation.[27] Self-regulation is defined here as the process of altering or controlling how one responds to stimuli.[28] This includes stimuli that is created both exogenously (e.g., the words and actions of others) and endogenously (e.g., one's own thoughts, feelings, beliefs, and inclinations). Note that the focus here is on how

[27] See chapters 4 and 5 for a more through discussion of both skill and self-regulation.
[28] Psychological theories of self-regulation normally focus on the ability of individuals to regulate their behavior to some 'ultimate' distal goal or to their conception of an 'idealized self' (e.g., Carver & Sheier, 1981; Baumeister et al., 1994; Fujita, 2011). This paper is adopting a much narrower conception of self-regulation. The focus here is on whether folks are able to alter, modify, or control their responses to their mental phenomena *regardless* of the standard that they are attempting to regulate to. See Chapter 4 for a more thorough discussion of this point.

one *responds* to mental content, not on the nature of the content itself. Self-regulation does not require the (likely impossible) ability to completely control the generation of all of one's thoughts, emotions, and behaviors, nor does it require that one only feel or think what one wishes. Rather, the focus of self-regulation is on one's ability to alter, override, or control how one responds to one's mental content, rather than on the specific content, valence, or hopefulness of one's doxastic states. Skill, here, is defined as the ability to act intelligently.[29] The 'intelligence' constraint is meant to distinguish skill from merely reflexive, lucky, or successful action. Theories of skill differ over how to specify intelligence, but, minimally, intelligent action requires the ability to adapt, learn, and intentionally modify one's behavior in response to new information. Unskilled behavior, in contrast, is marked by an inability to control or intelligently modify how one performs some act. *Skilled* self-regulation, then, is the exercise of the ability to intelligently alter or control how one responds to one's thoughts, emotions, and environment.

The focus on the improvement of patients' skill of self-regulation, rather than the alteration of patients' meanings, explains why the psychotherapeutic common factors are salubrious. Successful psychotherapy necessarily provides patients with the skill to alter or control how they respond to their mental phenomena. Consider, for example, the DSM-5's list of the symptoms of common disorders: generalized anxiety disorder is marked by excessive anxiety and worried thought, and problems controlling these thoughts; major depression is marked by persistent sadness or lack of pleasure and negatively valenced moods; bipolar disorders are marked by the

[29] This paper is not committed to the metaphysical claim that skill just *is* ability. See chapter 5 for a discussion of the metaphysics of skill.

periods of both extreme and unregulated negative and positively valenced emotions and cognitions; while schizophrenia is marked by a combination of delusions, hallucinations, disorganized speech, and diminished emotion or motivation (APA, 2013).[30] Different theories of psychological healing offer differing, and often disparate, explanations for the causes and best treatment of these symptoms, but, if successful, no therapy leaves individuals *unable* to regulate how they respond to these symptoms.[31] For example, regardless of the theoretical rational, no successful therapy for depression will leave patients unable to alter or control how they respond to their negatively valenced thoughts or dysphoric mood, while no successful therapy for anxiety will leave individuals unable to regulate how they respond to their worry and anxious feelings. Regardless of whether a psychotherapy focuses on challenging maladaptive core beliefs or on providing insight to unconscious conflicts, all successful therapies will provide patients with the tools to alter or control how they respond to the symptoms of their disorder.

Psychotherapy need not, and often does not, completely excise these symptoms from patients' mental lives. The successful treatment of anxiety and depression, for example, does not require that individuals no longer feel anxious or dysphoric, nor does it require that they no longer experience intense worry or negative thoughts (APA, 2013). Similarly, obsessive thoughts, compulsions, hallucinations, and extremely elevated mood can all be present without an individual fitting the diagnostic criteria of any mental disorder (APA, 2013). The skill

[30] These are not exhaustive definitions.
[31] 'Success' here is judged either by symptom reduction or by a patient no longer fitting a standardized diagnostic criterion. Of course, not all therapies consider symptom reduction the ultimate end-goal of therapeutic interventions. But, if the efficacy of therapies is to be plausibly statistically compared, the comparison needs to be based on controlled studies using similar diagnostic criteria.

hypothesis' focus on skilled action, rather than meaning, explains why these same doxastic attitudes can be present, and even dominant, in both healthy and disordered mental states. The difference, for example, between a philosopher whose mental life is dominated by thoughts about the meaninglessness of human existence, and an individual experiencing major depression is that the philosopher, presumably, is able to regulate how she responds to the negative content, valence, and hopelessness of her thoughts, assumptions, and beliefs, while the individual experiencing a depressive state cannot. The nihilist philosopher may even feel intense sadness and angst due to the content and valence of her doxastic attitudes, but unlike the individual experiencing a mental disorder, the nihilist is able to override, alter, or otherwise modify her negatively valenced thoughts, beliefs, and assumptions (even if she chooses not to). Similarly, while the same doxastic attitudes can be predominant in both extreme optimism and mania (e.g., inflated self-esteem, high-risk behavior, intense goal-directed behavior), extreme optimists are able to regulate their responses these 'symptoms', while individuals experiencing manic episodes cannot. Achievement-focused optimists may orient much of their mental and emotional lives towards the achievement of some (possibly unrealistic) goal, and this goal-dominated mindset may be harmful, but it is not necessarily disordered (and, at least in the case of business and the arts, it is sometimes admired). The problem being ameliorated in psychotherapy is not the presence, valence, or strength of any particular doxastic attitude, but is rather a problem of individuals' skill in regulating these mental states.

The skill hypothesis also explains what is right about the meaning dominance view. Mental disorders are marked by dominated minds, but this dominance is best understood as the inability

to skillfully self-regulate. Consider, for example, Beck's (1967) description of a mind dominated by depression:

> The vulnerability of the depression-prone person is attributable to the constellation of enduring negative attitudes about self, world, and future. Even though these attitudes (or concepts) may not be prominent or even discernible at a given time, they persist in a latent state like an explosive charge ready to be detonated by an appropriate set of conditions. Once activated, these concepts dominate the person's thinking and lead to the typical depressive symptomology (p. 277)

Beck is likely correct in claiming that depression is marked by a dominated mind, but as we've seen, there is nothing necessarily maladaptive about the content, valence, or hopefulness of the attitudes or concepts that are predominant in depression. Rather, in the case of depression (and mental disorder full stop), individuals are no longer able to intelligently alter or control how they respond to these predominant attitudes or concepts. The minds of individuals seeking psychotherapy may be 'stuck' or 'dominated', but they are only dominated insofar as individuals are unable to flexibly regulate how they respond to their mental content. Psychotherapy, if successful, enables patients to engage in skilled action to offset this dominance.

3.5 The Skill Hypothesis and the Common Factors

The skill hypothesis, like the meaning view, is meant to be a general claim explaining the efficacy of the therapeutic common factors and is compatible with versions of all three common

factors theories. The theories of Frank, Wampold, and Beck attempt to answer two questions: (1) *what* are the common therapeutic techniques or processes that underlie all effective psychotherapies, and (2) *why* do these techniques or processes engender psychological healing. In response to the 'what' question, Frank postulated four common factors (a healing relationship, a healing setting, a healing ritual, and a coherent rationale), Wampold three (the real relationship, the client's expectations, and a coherent rationale), and Beck one (the modification of core beliefs). In response to the 'why' question, all three theories postulate that changes in patient-meanings are the primary mechanism of change in psychotherapy, and thus explain why the presence of the common-factors are salubrious. This paper challenges only the later claim.

The skill hypothesis is not committed to the truth of any particular answer to the 'what' question. It is possible, for example, that one of Frank, Wampold, or Beck is right about which common factors are required for successful therapy. This is still an open question. This paper only disputes their explanations for *why* these common factors are salubrious (whatever they end up being). Beck, then, may be correct in claiming that *the* common factor in effective psychotherapy is the modification of core beliefs or schemas. His mistake, however, is in arguing that the modification of patient-meanings explains *why* the presence of this common factor leads to psychological healing. The claim, here, is that the success of therapeutic interventions (regardless of the specific common factors) is best explained by the improvement of individuals' regulatory skill, not their personal meanings. The alteration of the content, valence, hopefulness, or dominance of patient meanings, along with other factors, may play a role in psychological healing, but the primary control variable is the patient's skill. The problem being ameliorated in psychotherapy is not necessarily *what* patients thinks, believes, or feels, but rather *how.*

3.6 Conclusion

The Dodo bird findings create a problem for psychotherapeutic theories of psychological change. If all therapies are winning and getting prizes, regardless of theoretical and practical orientations, what explains the shared efficacy? Common factors theories maintain that the efficacy of psychotherapy is explained by shared mechanisms of change that enable adaptive alterations of the patient's maladaptive meanings. This paper has argued that the meaning view cannot fully explain the efficacy of therapy and is itself in need of a theoretical explanation. The skill view is meant to provide the beginning of such an explanation.

Chapter 4: The Skill of Mental Health

4.1 Introduction

While it is generally agreed that mental disorders are a serious problem, there is far less agreement about what the problem actually is. What exactly is *disordered* about mental disorders? The answer is normally situated somewhere between the naturalist/normative divide. Naturalism, roughly, is the position that mental disorders are value-independent, scientifically identifiable and explainable mental phenomena, akin to somatic disorders. Normative theories of health conceive of mental disorder as primarily a problem of the meanings attached to behavior and bodies; what makes something a mental disorder is primarily a question of value judgments about norm-transgressing behavior. Hybrid theories land somewhere between these poles. Naturalism, or hybrid theories with a necessarily naturalistic component, is the received view in academic psychology and psychiatry. Despite its wide acceptance, there have been significant problems in formulating the naturalist view. Notably, it has proven difficult to precisify what, exactly, the objective basis of mental illnesses is supposed to be.

The most influential naturalist approach has been to postulate that mental disorders are the result of objective mental *dysfunctions*.[32] Function-based naturalist views posit that mental disorders necessarily involve the failure of some mental mechanism, or mechanisms, to function

[32] For example, see: Wakefield (1992); Boorse (1975, 1997, 2014); Spitzer & Endicott (1978); APA (1987; 2013); WHO, 1992; Murphy, 2006; Garson & Piccinini, 2010; Hausman, 2012.

'normally'. The most serious difficulty facing these views has been to give an account of normal psychological functioning, and an explanation of what, if anything, is *the* function of the mind (or of any particular hypothesized mental mechanism). The perceived failure of function-based naturalist views to offer a plausible theory of normal mental function has been taken by many to be evidence for the failure of naturalism, full stop, and motivation for the acceptance of normative views.[33] In contrast, this paper argues that mental functions are the wrong point of emphasis in explaining the objective nature of mental disorder and that a more plausible naturalist theory can be constructed by grounding the value-independent basis of mental disorder in the concept of skilled action, rather than proper mental function. According to the novel skill view, outlined here, mental health is a skilled action of individuals, not the normal functioning of mental mechanisms. Mental disorder is the failure or breakdown of this skilled.

This paper is an argument to the best explanation. There are at least two desiderata that a theory of mental health should address: (1) it should offer a plausible theory of the boundary between health and disorder, and (2) it should explain how we might know whether any particular mental state falls on one side or the other of this boundary. The motive for the first desideratum is clear: a naturalist theory of mental disorder needs to explain what differentiates disorder from health. The second desideratum is also crucial: a good theory of mental health and disorder should be *useful* and provide us with the means to differentiate between the two. The skill view of mental health provides a naturalist answer to both (1) and (2), while avoiding the more serious metaphysical and epistemic problems of traditional function-based naturalist views.

[33] See for example: Cooper, 2005; Fulford, 1989; Prinz, 2012.

The structure of the paper is as follows. Section 4.2 addresses the flaws of traditional function-based naturalist views. Section 4.3 introduces the alternative skill-based naturalist theory of mental health. Sections 4.4 and 4.5 argue for the conceptual and explanatory superiority of the skill view over both function-based naturalist theories and normative conceptions of mental health and disorder.

4.2 Naturalism and Mental Functions

The identification of mental disorders is traditionally conceptualized in functional terms. For example, the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5*, American Psychiatric Association, 2013), the primary diagnostic manual in North America, defines mental disorders as follows:

> 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*. (italics added)

The DSM's international counterpart, the World Health Organization's *International Classification of Diseases* (*ICD-10*, 1992), provides a similar function-based definition (though it focuses on the functions of persons rather than just minds):

> ‘Disorder’ is not an exact term, but it is used here to imply the existence of a clinically recognizable set of symptoms or behaviour associated in most cases with distress and with interference with personal *functions*. Social deviance or conflict alone, without personal *dysfunction*, should not be included in mental disorder as defined here.
> (italics added)

Both the DSM-5 and ICD-10 are explicitly atheoretical. They are intended to serve as theory-neutral diagnostic tools for the identification of symptoms that are jointly sufficient for the diagnosis of a disorder, not as theories of what disorders are. Theories of mental disorder, on the other hand, are attempts to explain what it is about certain modes of functioning that makes them healthy, disordered, or benign. And, according to the most influential naturalist views, what differentiates ‘genuine’ disorders from normal human variation is that the former, but not the latter, are marked by objective failures of mental functions.

The primary challenge facing function-based naturalist views has been to provide a value-independent account of ‘normal’ mental functioning. Things can perform multiple functions. Hearts pump blood, eyes enable vision, noses enable olfaction, and so on. But hearts also make beating sounds, eyes also attract potential mates, and noses hold up reading glasses. Function-based naturalist views of health need to explain what makes it the case that any given function is *the* normal function of any particular hypothesized internal part or process (such that a *dys*-function of that part or process distinguished mental disorders from non-disordered states). Naturalist solutions to this problem differ in their specifics, but share the common strategy of

defining normal mental function as the contribution of a mental part or process to an ultimate biological norm. The most influential candidates for this 'ultimate' norm have been evolutionary design or individual survival and reproduction.[34] According to these 'biological norms' views, the normal function of the heart is to pump blood, and not make pounding sounds, because the pumping of blood, but not the production of sound, conforms to some ultimate biological norm (e.g., it is what the heart was designed by natural selection to do, it contributes to the ultimate end-goals of individual survival and reproduction, or it conforms to some other biological norm). The literature on the metaphysics of mental health now consists largely of attempts to articulate, criticize, or defend different versions of this biological norm theories of mental function.[35] For the purposes of this paper, let's assume that function-based naturalist views can be made coherent. The focus of this section is to examine whether, even in their strongest forms, function-based views offer us a good explanation of the desiderata.

There are two general, and well known, problems shared by dysfunction-based conceptions of disorder that should give us pause. The first problem is epistemic: even if we accept that something like 'dysfunction' is the right way to conceive of disorder, there is nothing close to a consensus theory of mental functions from which we can identify whether or not some behavior is a deviation from normal function. In the case of certain somatic functions, the identification of their proper functions may seem clear enough: presumably, the function of the heart is to pump blood, the function of the eye is to enable vision, and the function of the stomach is to digest

[34] The touchstones for the evolutionary and longevity/reproduction views are Wakefield (1992, 1999) and Boorse (1977, 1997, 2014), respectively. See Garson (2008) for a good overview.
[35] The literature here is extensive. For a good overview, see: Cooper, 2005; Schroeder, 2013; Kingma, 2014; Boorse (1997, 2014).

food. Things are far less clear concerning the function of the hypothesized parts and processes of the mind.

In the case of somatic functions, there is normally agreement about the existence of the parts or processes whose function we are trying to identify (e.g., we all agree that hearts exist, even if we disagree about their proper function). In contrast, the existence of the mental mechanisms whose normal functioning we are trying to identify vary significantly between theories of psychopathology. The 'mechanisms' included in a psychodynamic model of psychopathology may include subconscious defense mechanisms (e.g., repression or sublimation), the ego, and noncognitive psychological drives; the 'mechanisms' involved in a cognitive-behavioral based theory may include 'core beliefs' that guide all cognitive activity (e.g., 'I am flawed and therefore unacceptable' or 'the world is unsafe') and 'primal modes' (i.e., evolutionarily designed clusters of similarly themed beliefs) which are triggered during disorders; while strict behavioral theories postulate 'mechanisms' associated with learned behavioral rules, punishment, and rewards to be the primary cause of psychopathology (rather than either consciously accessible core beliefs or subconscious conflicts). And, problematically, the efficacy of psychotherapeutic interventions appears to be unrelated to which of these hypothesized mental parts and processes are being explicitly treated in therapy (Smith et al., 1980; Luborsky, 2002; Wampold & Imel, 2015). The problem for function-based theories, then, is that for whatever biological norm they postulate to ground the concept of 'normal' functioning, they cannot justifiably identify *failures* of normal mental functioning unless they have an accurate accounting of the parts and processes of the human mind. So far, this is not on offer.

It is also an open empirical question whether behaviors and mental phenomena normally considered to be 'disordered', such as the phenomena associated with depression, anxiety, and schizophrenia, are in fact the result of dysfunctioning mental mechanisms rather than being the consequences of mechanisms functioning as they normally should.[36] Depression, for example, is variously hypothesized to be the result of *dys*-functioning mental mechanisms that lead to 'interlocked' cognitive-affective cycles of mental processing (Teasdale & Barnard, 1993), the product of *normally* functioning mental mechanisms that lead to an increase in critical ruminative thought (Andrews & Thompson, 2009), the result of *dys*-functioning mental mechanisms that produce negatively valenced loss-based cognitions (Beck & Alford, 2009), or the result of *normally* functioning mental processes that socially signal the need for assistance (Allen & Badcock, 2003). Similarly, generalized anxiety disorder (GAD) is postulated to be the result of a *dys*-functioning evolved-for harm and danger avoidance system (Clark & Beck, 2011), the result of a potentially *normally* functioning mind that has learned to process information in an anxiety producing manner (Greenberg, 2010), the product of *dys*-functioning subconscious mechanisms that produce conflicts between forbidden wishes and defenses against these wishes (Busch et al., 1999), or the result of a possibly adaptive mental process to combat real or perceived threats to social exclusion (Baumeister & Tice, 1990). And accuracy matters here: whether or not we view some mental state as disordered, according to function-based views, will depend upon which specific theory of mental functioning we adopt.

[36] This concern has been raised by a number of authors (including advocates of function-based naturalism). See: Boorse, 1976; Lilienfeld & Marino, 1995; Woolfolk, 1999; Murphy & Woolfolk, 2000; Bolton, 2001; Cooper, 2002.

Without taking a position on these theories, we can see the general problem. Without an accurate theory of mental functioning we have no non-speculative way of differentiating between genuine instances of mental disorder and mere human difference. It may be possible that this epistemic problem has a plausible empirical solution and that the lack of clarity concerning mental functions, cognitive architecture, and the etiology of mental disorders is just a reflection of the relative youth of the field of cognitive science. This may turn out to be the case. But, again without taking a position on the likelihood of the future theoretical unification of cognitive science, psychology, and psychiatry, the problem remains that function-based theories are *now* reliant upon epistemically uncertain foundational assumptions about mental functioning. This is a significant flaw. The acceptance of function-based views, then, may give us (the possibly false) confidence that there is a difference between genuine mental disorders and norms-transgressing behaviors, but leave us without the means to accurately tell when this might be the case.

The second general problem with function-based views is conceptual. Function-based theories commit us to the view that mental health necessarily involves the adherence of mental mechanisms to certain ultimate biological norms. This focus on biological norms (e.g., evolutionary design or individual survival and reproduction), while potentially establishing a value-free foundation for the concept of disorder, also leads to a potentially massively revisionary conception of mental health.[37] Dysfunction views are committed to the claim that it is an open (though possibly unanswerable) empirical question whether any postulated 'mental

[37] Murphy (2006) makes a similar point, but denies that this poses a problem for function-based views. He argues that we should revise our intuitions in light of our best theories of mental functioning.

disorder' is in fact the result of a dysfunctioning mental mechanism, and thus a genuine mental disorder. It is possible that paradigmatic disorders such as depression, anxiety, and phobias may turn out *not* to be mental disorders, but rather products of *healthy* minds dealing with mere problems of living. While, on the other hand, behavior not commonly considered to be product of unhealthy minds, such as homosexuality, a lack of desire to reproduce, and the engagement in dangerous 'thrill-seeking' activities, may in fact be mental disorders if it turns out that they are the result of dysfunctioning mental mechanisms.[38] This is not incoherent, but it is highly counterintuitive.

Being revisionary is not necessarily a flaw in a theory of health. Many behaviors and ways of being have at one time been considered to be pathological but are now commonly perceived to be normal and healthy; new additions or subtractions should not be overly surprising. What is surprising is that function-based conceptions of health make it an open question whether *paradigmatic* instances of disorder are in fact problems of mental health. The behavior and mental phenomena associated with depression, anxiety, and schizophrenia are amongst the core explanandum that a naturalist theory of disorder is supposed to account for. We should expect a naturalist theory of disorder to explain *why* paradigmatic 'disordered' states are genuine mental disorders, and to explain what differentiates them from normal problems of living or norm-

[38] Note that while function-based views agree that mental dysfunctions are necessary for disorders, they differ over whether they are also sufficient. 'Pure' function-based views (e.g., Boorse, 1977) posit that the presence of a mental dysfunction is both necessary and sufficient for disorder. So, for example, homosexuality is a mental disorder if it is the result of a mental dysfunction (e.g., if it interferes with whatever ultimate biological norm is used to ground 'normal' function). 'Hybrid' views (e.g., Wakefield, 1992) posit that dysfunction is necessary, but not sufficient for disorder. Genuine disorders must also be judged to be 'harmful' (with the 'harm' criterion being a value-based claim that varies between cultures and value systems). So, the 'pathological' status of homosexuality will depend both upon its status as a dysfunction and whether it is judged to be harmful in a given value-system.

transgressing behavior. It is, of course, possible that the best theory of disorder will end up revising the original set of paradigmatic phenomena, but we need good reason to do this. Being revisionary isn't necessarily a problem if the theory's explanatory utility outweighs its costs, but given the epistemic weakness of function-based views, that does not appear to be the case here.

The use of dysfunction, then, to ground the objective basis of mental disorders gives us a theory that is *both* explanatorily weak and possibly massively revisionary. If function-based theories of health make it possible that depression and anxiety are not mental disorders, but homosexuality is, then this gives us good reason to think that the presence or absence of mental dysfunctions is not necessarily related to the concept of mental health. That function-based views may be so revisionary and contrary to common folk conceptions of disorder is *prima facie* evidence that something has gone wrong with the focus on dysfunctions to explain the distinction between mental health and disorder, and gives us reason to search for a more plausible naturalist conception of mental health.

4.2.1 Summary

The problem is this: despite their prominence, there have been serious epistemic and metaphysical problems in formulating a naturalist conception of mental disorder. Dysfunction-based naturalist views are not incoherent or untenable, but they are seriously flawed. These flaws can either be accepted as part of the cost of formulating a naturalist theory, or taken as evidence for the failure of naturalism. In contrast, this paper argues that a more plausible naturalist view of health can be constructed if we reorient the dialectic away from the concept of normal function

and focus instead on the skills responsible for mental health. Skill, rather than function, can provide the objective basis for a naturalist theory of mental disorder.

The new naturalist theory of mental health outlined here argues that mental health is best conceived of as a skill. More specifically, this paper argues that mental health is skilled self-regulation; mental disorder is a failure or breakdown of this skill. The skill view of mental health provides a naturalist framework for the scientific study and treatment of mental disorders that avoids the epistemological and conceptual failings of traditional function-based views. The focus on skilled action, rather than the adherence to biological norms, allows for a flexible theory of health that avoids pathologizing normal human difference while also providing a non-revisionary accounting of paradigmatic mental disorders.

4.3 The Skill of Mental Health

The skill view of mental health makes the following two claims:

1: Mental health is skilled self-regulation.

2: Mental disorder is a failure or breakdown of this skill.

Sections 4.3.1-4.3.2.2 outline theory-neutral naturalist accounts of skilled action and self-regulation. Sections 4.3.3-4.4 will address the skill view's claims in order.

4.3.1 Skill, Self-regulation, and Values

The term 'skill' has number of common uses. 'Skill' can refer to a particular type of possible action (e.g., the skill of playing tennis), to a way of acting (e.g., being a skilled, as opposed to unskilled, tennis server), and to an expertise in action (e.g., being a skilled, as opposed to novice, tennis player). This paper is concerned only with the first two conceptions of skill. The focus here is on what differentiates skilled from unskilled action, not what differentiates low skill from high skill, or skill from expertise.

This paper adopts a theory-neutral conception of skill: minimally, to be skilled at ϕ-ing requires that one possesses the ability to ϕ intelligently.[39] Both 'ability' and 'intelligence' can be disambiguated. Following Mele (2003), we can distinguish between two senses of 'ability':[40]

> (General) ability$_1$: An agent is able to ϕ, given appropriate background conditions.
>
> (Specific) ability$_2$: An agent is able to ϕ *and* has the appropriate background conditions satisfied to successfully ϕ.

The distinction is straight-forward. There are many things an agent may be generally able$_1$ to do (such as play tennis), but that she is at some point in time unable$_2$ to do because of something that masks this ability$_1$ (such as a sprained ankle). Skill at ϕ-ing necessarily requires the ability$_1$

[39] Note that this paper is not committed to the claim that skill just *is* an ability. This minimal definition is consistent with both Intellectualist (e.g., Stanley & Williamson, 2001, 2017) and Non-Intellectualist (e.g., Ryle, 1949) theories of skill. See Chapter 5 for a more thorough discussion of the metaphysics of skill.

[40] Mele (2003) does not use this exact formulation. Berofsky (2003) and Noë (2005) make a similar distinction.

to ϕ. There is no scenario in which an agent is skilled at performing some action, yet there is no possible circumstance in which she is able$_1$ to do so. For example, whatever the specifics are of a theory of skill, a person cannot be a skilled tennis player if she is not now, nor ever was, able$_1$ to play tennis. However, possessing the ability$_1$ to ϕ does not entail that one is always *able*$_2$ to ϕ. A skilled tennis server is able$_1$ to strike tennis balls into the service area, *only* given that certain background conditions are met (e.g., that she has access to a racket, ball, is free from injury, and so on). The lack of a racket or a sprained ankle might render her unable$_2$ to play tennis on any given day, without necessarily affecting her ability$_1$, or skill, to play. Skills may diminish after long periods of being unable$_2$ to perform, but one's ability$_1$ to ϕ is not necessarily dependent upon whether one is able$_2$ to ϕ at any given time. A skilled *action*, on the other hand, does require that one is able$_2$ to ϕ. For example, skilled tennis service requires that one is both able$_1$ to strike balls into the opposing service box, *and* that the enabling conditions are met for her to do so successfully.

Skill may require ability$_1$, but ability$_1$ is not sufficient for the presence of skill. A simple tennis-ball machine may be able$_1$ to regularly propel balls into the opposing service area, and a lucky maladroit may be able$_1$ to successfully, though accidentally, strike tennis balls, yet neither activity is a skilled act. The 'intelligence' constraint is meant to distinguish skill from reflexive, lucky, or merely successful action. Theories of skill differ greatly over how to specify the intelligence clause, but there is broad agreement that, minimally, 'intelligence' requires the ability$_1$ to adapt, learn, and intentionally modify one's behavior in response to novel stimuli. The tennis machine and lucky maladroit lack the ability$_1$ (at least at that moment) to intentionally modify their behaviors to adapt to new problems such as a change in wind speed, a damaged gear

or muscle, or a change in altitude. A skilled tennis player, on the other hand, would be able$_1$ to learn from, and adapt to, different surfaces, opponents, and physical constraints. Theories of skill differ over what else, if anything, is needed in addition to intelligence and ability$_1$ to explain skill. The primary point of disagreement in metaphysical accounts of skill is how to precisify the intelligence constraint. 'Intellectualist' theories of skill claim that skill at ϕ-ing requires being guided by knowledge of propositional attitudes about how to ϕ. According to these views, one can only ϕ intelligently if one knows facts about how to ϕ. 'Anti-intellectualist' theories deny this claim. As with our analysis of function-based views, we need not take a position here. The aim of this paper is not to argue for a particular theory of skill, but rather to examine whether the concept of skill (given the above constraints) can ground a naturalist theory of mental health.

Importantly, the concept of skilled action can be formulated in value-independent terms. How *well* one performs some task and how well one modifies one's behavior clearly *are* value-judgements. The concept of skill as *expertise* also appears to be value-laden (insofar as whether or not some action is done expertly will vary depending upon the standards set). In contrast, whether an agent is able$_2$ to intelligently ϕ is not, or at least need not be, a value judgment. For example, metaphysical theories of ability commonly take the conditional form: an agent is able$_2$ to ϕ if she would ϕ if she tried (given the appropriate background enabling conditions). Theories differ over how to best specify these enabling conditions, but these theoretical differences needn't be differences in *value* (just as metaphysical disputes over the existence of midsize objects are not necessarily disputes over values). Similarly, theories of the metaphysics of intelligence generally are attempts to precisify some cognitive faculty that things like humans have, and things like modern computers currently lack. There is certainly room for values to

enter into the answer to this question, but there is no necessary reason that this need be the case. More generally, insofar as a value-independent metaphysics is possible, the metaphysics of skill, like the metaphysics of function, poses no unique problem that necessitates a value-based theory.

There are, of course, *epistemic* challenges in identifying when any particular action is skilled. Section 4.3.4 will address these epistemic constraints, and compare them favorably to those facing function-based views. Sections 4.3.2 and 4.3.3 will outline theoretical commitments of the skill-based view of health.

4.3.2 Skill and Self-regulation

This paper assumes that self-regulation is a skill. This is trivially true. Clearly, self-regulation is *a* skill, just as playing tennis or negotiating are skills. To be a skill is just to be a possible mode of intentional action. One can be skilled at playing games such as chess or tennis, at interpersonal interactions such as negotiating or manipulating, or at seemingly trivial tasks such as balancing things on one's finger or differentiated between types of wine. That something is a skill is not a controversial claim; *any* possible mode of intentional action can be a skill. Claim (1) makes the stronger claim, defended in the next section, that mental health is skilled self-regulation.

'Self-regulation' refers to the process of altering or controlling one's responses to align with one's goals or standards (Fujita, 2011; Baumeister et al., 1993). The absence of self-regulation is automatic cognition, emotion, and behavior based on 'learning, habit, inclination, or even innate tendencies (Baumeister et al, 1994). Everyday examples of self-regulation include regulating

emotional responses to perceived losses or gains (e.g., inhibiting the experience of intense emotions), regulating cognitions (e.g., avoiding ruminative patterns of thought, or conversely, focusing one's attention despite distractors), and regulating impulses (e.g., in regard to food, drink, or self-expression). Psychological accounts of self-regulation are traditionally focused on the ability$_2$ to regulate one's responses towards some ultimate distal goal or idealized conception of self (Carver & Sheier, 1981). This paper adopts a more restricted conception of self-regulation; we are interested in the ability$_1$ to regulate one's responses to one's *immediate* willings.[41] We are not concerned here with the notion of the idealized self or in identifying individuals' ultimate goals. Rather, the focus is on whether one is *able*$_2$, at any given time, to skillfully regulate one's responses to whatever standard one wills.

There is also an important difference between regulating one's *responses* and completely *controlling* one's mental life.[42] Self-regulating involves the 'overriding, stopping, modifying, or otherwise changing' of one's responses to exogenously and endogenously generated stimuli (Rawn & Vohs, 2010). This skill can, but need not, involve controlling the generation mental content. Completely controlling the generation of mental content would require the ability$_1$ to only think, feel, and desire what one wills (e.g., only thinking pleasant thoughts and experiencing positive emotions). Self-regulation does not require the (possibly superhuman) ability$_1$ to completely control the generation of one's cognitions and emotions. Regulatory strategies may *include* the attempt to control the generation of mental content (e.g., by avoiding situational

[41] Section 4.4 addresses potential concerns with this restricted definition.
[42] The terms 'regulation' and 'control' are often used interchangeably (e.g. Baumeister et al., 1998; 1996), though they are sometimes used to mark a means-end distinction (with control being one of many methods to self-regulate) (e.g., Fujita, 2011).

triggers), but attempts to control content can also be part of the *problem* that requires regulation (as can be the case with obsessive disorders). Automatic and unwanted thoughts, emotions, and urges are a normal part of mental functioning, and the failure to completely control the content of one's mental life is not necessarily a failure of self-regulation. Failures of self-regulation involve inability$_2$ to regulate how one *responds* to these thoughts, emotions, and life-events.

4.3.3 Skill and Mental Health

Claim (2) states that mental health is skilled self-regulation. The term 'skilled' in claim (2) refers to the *exercise* of the skill of self-regulation. Mental health requires not just the possession of the skill to self-regulate, but also that the appropriate enabling conditions are met in order to be able$_2$ to exercise this skill. According to this conception of health, mental health is a skilled way of acting, rather than the proper functioning of mental parts.

In support of (2), we should first notice that mental health necessarily *involves* self-regulation. This is made clear if we consider the process of mental healing. Regardless of whether disorders are conceived of as a problem of natural function or of value, one of the primary end-goals of psychotherapy is improved self-regulation. In the case of normative theories of health, successful psychological treatment will just be whatever causes the patient to regulate her responses to fit within the established norms of 'healthy' functioning. In the case of naturalist function-based theories of health, whatever the mental dysfunction is that is hypothesized to be responsible for a particular mental disorder, the *proper* function of this hypothesized mechanism will involve the ability$_2$ to self-regulate. This is not to say that function-based theories are committed to the claim

that self-regulation is the only function of the mind or that self-regulation is the ultimate function of all mental mechanisms. However, if, after successful therapy, the proper functioning of some hypothesized mental mechanism involves persons being unable$_2$ to self-regulate, then we should take this as evidence that we are mistaken about our standard of mental health.

There is nothing normally recognizable as a *mental* disorder (as opposed to a brain disorder) that does not involve a problem of self-regulation.[43] For example, if a naturalist-inclined therapist claims that a patient seeking treatment for major depression is 'healed' (i.e., her hypothesized dysfunctional mechanism(s) have returned to normal functioning), yet reports that the patient is still unable$_2$ to regulate her negative emotions and cognitions, we should conclude that the therapist is confused about what depression is. Similarly, obsessive compulsive disorder (OCD) involves the experience of reoccurring intrusive and unwanted thoughts, urges, or impulses, as well as compulsive repetitive behaviors or mental acts to suppress or ignore these obsessions; any psychological intervention that successfully treats OCD will necessarily require that patients become able$_2$ to regulate their responses to their obsessions and compulsions. And we see the same patterns with other paradigmatic disorders. Different mental disorders involve different regulatory problems: anxiety disorders are marked by excessive anxiety and worry, and difficulty in controlling these worries; schizophrenia is marked by delusional beliefs or hallucinations as well as disorganized speech and behavior; bi-polar disorders involve both serious depressive symptoms and manic behaviors and patterns of thought (APA, 2013).[44] It is an open empirical

[43] The distinction between brain and mental disorders in addressed later in this section.
[44] These descriptions come from the DSM-5. They are not exhaustive definitions.

question as to how best treat these disorders, but the end-goal of psychotherapeutic treatment is the same: no successful treatment will leave patients unable$_2$ to regulate their responses to these symptoms

Theories of psychotherapy differ over the *causes* of the symptoms that are being regulated, and over how to best ameliorate these problems. But, regardless of whether the therapy focuses on identifying and challenging core beliefs, providing insight into subconscious processes, extinguishing learned behavioral responses, or ameliorating any other hypothesized cause of psychopathology, successful psychotherapeutic interventions will provide patients with the skill to self-regulate. If successful treatment does not involve the patient regaining some of her ability$_2$ to intelligently control how she responds to thoughts, emotions, and events, then it is unclear what is supposed to be 'successful' about the therapy. For example, no effective psychotherapy for anxiety (measured by the patient no longer fitting the diagnostic criteria for a disorder) will leave the patient unable$_2$ to regulate how she responds to anxious feelings, thoughts, or (formerly) anxiety-producing life events. And no effective therapy for schizophrenia will leave a patient without the ability$_2$ to regulate her mental phenomena. This is not to say that improved self-regulation is the *only* goal of all psychotherapies, or that it is even necessarily an explicit goal. But, if psychotherapy is effective, it is effective insofar as it enables improved self-regulation.

If there is no problem of self-regulation, then there should be no problem of mental health.[45] However, some state not being a *mental* disorder does not mean that it is not a disorder. The focus on skilled self-regulation helps explain the common distinction between mental disorders and brain disorders.[46] Brain disorders, such as amnesia disorders, epilepsy, dementia, and Alzheimer's disease, may result in failures of self-regulation, but they are not best treated as regulatory problems. Treatment for neurological disorders such as Alzheimer's disease may include skill training (e.g., learning to cope with the fear and frustration sometimes associated with memory loss), but even if a patient were to (somehow) be able$_2$ to skillfully self-regulate, this would not make it the case that she no longer has a disorder (though it may be less subjectively distressing). In contrast, if an individual seeking treatment for a mental disorder such as depression or anxiety no longer has a problem skillfully self-regulating, then there should no longer be any disorder. What makes a mental disorder *mental*, rather than somatic, is that mental disorders are best explained and treated as problems of skilled self-regulation while somatic disorders are not.[47]

---

[45] A possible exception to this claim are some of the Cluster B personality disorders. There is debate about whether personality disorders such as antisocial personality disorder (commonly referred to as sociopathy in lay-terms), histrionic personality disorder, and narcissistic personality disorder should be considered mental health problems or moral problems (e.g., Charland, 2004; Kendell, 2002; Pickard, 2009). The resolution of this debate will come down to what theory of disorder we adopt. The skill view entails that at least some personality disorders are not mental disorders *if* it is the case that they do not involve inabilities$_2$ to skillfully self-regulate. And this is just what we should expect. If a putative nosological entity such as antisocial personality disorder does not involve any inability$_2$, it seems clear that the behavior being pathologized better reflects disvalued behavior and character, not poor health.

[46] Not everyone agrees that this distinction is justified. The worry is that this distinction implies a form of Cartesian dualism between the body and mind. This concern is not justified here. We can accept that mental functions are physical while also accepting that there is a different standard of health for somatic and mental disorders. (See section 4.4)

[47] Graham (2013) makes a similar point. According to Graham, we should use the effectiveness of 'mind-centered' therapies (i.e., talk-therapies such as CBT or psychoanalysis) in treating specific conditions as a method to distinguish whether the condition is best classified as a mental disorder or a brain disorder.

Both normative and function-based naturalist theories should agree with this, they just disagree about whether individuals with mental disorders are regulating their mental lives to an objective standard of health, and what, if anything we need to posit in *addition* to self-regulation to explain mental health. For normative views, the additional explanatory piece is a value-judgment. According to these views, while self-regulation plays a role in mental *healing*, whether the healing is successful, and, more generally, whether an individual is judged to be mentally *healthy*, will depend on a value judgment of some kind. In contrast, function-based naturalist theories posit that we need the extra concepts of normal function and dysfunction in order to establish a value-independent basis for the distinction between health and disorder (and between healthy and disordered self-regulation). The problem, as we've seen, is that the use of function to draw this boundary leads to serious epistemic and conceptual problems. The concept of skill can do better.

4.3.4 Skill, Mental Disorder, and the Boundary Problem

There are two related, but distinct, concerns about boundaries that theories of mental disorder must address: the first is a metaphysical issue ('*what* is the difference'), the second epistemic ('how might we *know* the difference'). The epistemic boundary problem will be addressed in section 2.3.1. This section is focused on the metaphysical question.

Claim (3) states that mental disorder is unskilled self-regulation. According to the skill view, the boundary between mental health and mental disorder is a question of ability$_2$ to intelligently self-regulate. This does not entail that any regulatory failure is necessarily a sign of mental disorder.

Healthy individuals can and do lose control of their thoughts and emotions, and act in ways contrary to their explicit intentions. This is normal and often benign. The difference between normal regulatory lapses and breakdowns of skilled self-regulation (and, thus, the difference between mental health and disorder), is that disorders involve an *inability*$_2$ to self-regulate, not just the poor exercise of regulatory skill. Just as a skilled athlete does not become unskilled when she misstrikes a ball or makes a tactical mistake, mentally healthy individuals do not become disordered merely by engaging in dysregulated behavior. The boundary between mental health and disorder is a question of what individuals are able$_2$ to do, not just what they do.

This solution to the boundary problem provides a principled, value-independent distinction between mental disorders and merely socially disvalued behaviors. Mental health requires being *able*$_2$ to regulate one's responses to emotions and cognitions, but not necessarily that one regulates the content of these mental states to any particular end. The difference between paradigmatic mental disorders (such as GAD or OCD) and instances of normal human variation (such as fringe religious or political beliefs) is that the latter involves the *ability*$_2$ to self-regulate, while the former does not.[48] This is not to say that individuals with a mental disorder do not have the capacity to self-regulate (presumably they do), but only that at the time in question they do

[48] Frankfurt (1982) seems to dispute this claim. According to Frankfurt, people with deeply held religious, moral, or political beliefs are at least sometimes genuinely unable$_2$ to do otherwise than to follow them (or, in Frankfurt's terms, they are 'constrained by volitional necessity') while not experiencing a mental disorder (1982, p. 87). Frankfurt states that these individuals are unlike addicts (and, therefore, not experiencing a mental disorder), because they are compelled by forces that they are 'unwilling to oppose' (e.g., they are unwilling to oppose the supposed truth of their ethical beliefs) (p. 87). However, it seems clear that Frankfurt is using the concept of 'ability' metaphorically here. For example, it is possible that some addicts are also unwilling to oppose their addiction; this does not impact their current inability$_2$. Many choices are hard, and many choices can metaphorically be described as being necessitated, but if one is genuinely unable$_2$ to alter or modify one's choice (e.g., about one's perceived ethical duties or about one's addictive behavior), this much more closely resembles a pathological compulsion or obsession than a deeply held belief.

not have the ability$_2$ to do so. A mentally healthy political radical who suffers as a result of her beliefs could (in the sense of having the ability$_2$) choose not to focus her mental life exclusively on her political views, while an individual suffering from GAD is not, at that time, able$_2$ to regulate her anxious emotions and cognitions. Similarly, (3) entails that while some sexual thoughts or behaviors may be symptoms of a mental disorder (e.g., unwanted and intrusive obsessive sexual thoughts and images), there is nothing necessarily disordered about any particular norms-transgressing sexual behavior or thought. Mental disorders (such as OCD) can be differentiated from non-disordered states (such as homosexuality) because the former, but not the latter, involves the inability$_2$ to successfully regulate inflexible cognitive and emotional patterns. The content or object of one's sexual desires, as opposed to the ability$_2$ to regulate these sexual thoughts or compulsions, have no necessary relation to mental health.[49] Being attracted to some object *O*, is not necessarily a problem; being unable$_2$ to regulate one's *O*-directed obsessions and compulsions will be. Similarly, having fringe religious views need not be a symptom of a mental disorder, but having the genuine inability$_2$ to regulate these views will be.

The focus on skill, rather than function, also explains the difference between disorder and normal human difference. People differ greatly in the content of their mental states. There is nothing necessarily disordered about atypical or norms-transgressing beliefs or emotional reactions, and a theory of mental health should reflect this. This normal human variation reflects a disorder only

[49] Note that it is possible that individuals with norm-transgressing sexual orientations or preferences *may* be distressed by their sexual thoughts and desires (e.g., because they run counter to social norms) and experience anxiety or depression based on their inability$_2$ to alter or control the content of these thoughts and desires. This felt need to control the object of one's sexual desires and thoughts may lead to mental health problems, but this does not entail that there is anything pathological about the content of these thoughts or the object of these desires. These problems would be mood disorders (influenced by cultural norms).

when persons are no longer able$_2$ to regulate these mental states. For instance, both philosophers and individuals suffering from psychotic episodes may claim that the moon does not exist.[50] Both persons may present arguments in defense of this claim, and both may also become emotionally agitated when others disagree with them. The difference between a philosopher denying the reality of the moon (and, say, all other non-living composite objects) and a person suffering from delusional beliefs is that, presumably, the philosopher has the ability$_2$ to challenge and regulate her thoughts and beliefs about the moon while the person suffering from delusions cannot. Similarly, most people report experiencing intrusive thoughts (e.g., about unwanted sex acts, violence, contamination, etc.), while only a small minority meet the diagnostic criteria for obsession-related disorders (Rachman & de Silva, 1978; Radomsky et al., 2014). The difference between normal (if still distressing) intrusive thoughts, and disorders (such as OCD), is demarcated by individuals' ability$_2$ to intelligently regulate their cognitive and emotional responses to them. The difference between mental health and disorder, in general, is drawn by individuals' ability$_2$ to regulate their mental lives, not by the content or objects of their cognitions and emotions.

It is important to note what the skill view is *not* claiming. The skill view is not claiming that mental content plays no role in the etiology or experience of mental disorders- this is clearly not the case. For example, negatively valenced thoughts and emotions are constitutive of depression; a necessary condition of what it is to be depressed is the presence of depressed mood or diminished pleasure. But the presence and prevalence of negatively valenced mental content is

[50] The example is adapted from Van Inwagen's (1990) argument concerning the reality of composite objects.

not sufficient for mental disorder. Nihilists and existential philosophers may regularly think about the meaninglessness of life, and may even feel intense and persistent angst or sadness in response to these thoughts, but they needn't be suffering from clinical depression (or any other mental disorder). A predominance of negatively valenced thoughts and emotions may play a causal role in the development of depression, but this need not be the case. One needn't be disordered to believe that life is meaningless and painful, to devote much of one's mental life to these beliefs, and to even be profoundly saddened by this. What makes a state marked by negatively valenced mental content a *disorder* is the failure of skilled self-regulation of this content, not merely the presence of any specific thoughts or emotions. The difference between normal (even if intense) sadness, dysphoria, or hopelessness, on the one hand, and depression on the other, is that depression (and mental disorders in general) involve an inability_2 to skillfully regulate how one responds to mental content. Mental content, then, plays a role in mental health insofar as responses to content is often what is being regulated, but the presence or absence of any particular mental content does not explain why certain states are disordered and others not. If there is a mental health problem, it is because of a failure of skilled self-regulation, not because of content being regulated.

The skill view is also not claiming that *any* problem of self-regulation is a problem of mental health. There are numerous instances of poor self-regulation that are clearly not disorders.[51]

[51] Note that the skill view does not claim that a lack of the *capacity* (as opposed to ability_2) to skillfully self-regulate entails that an individual will have a mental disorder. Infants and small children may be unable_2 to skillfully self-regulate, yet they are clearly not mentally disordered. The skill view argues that mental disorders are failures to skillfully self-regulate; this assumes that one has the capacity to accomplish this act. In the case of adult mental disorders, the disorders usually involve an inability_2, rather than an inability_1, to self-regulate (e.g., an adult experiencing clinical depression is likely able_1 to self-regulate, but is temporarily unable_2 to exercise this skill). In

These problems can range from the seemingly trivial (e.g., one can have difficulty resisting a bowl of sweets), to the potentially more serious (e.g., sticking to a diet, or keeping one's temper while watching a sporting event). The skill view predicts that these cases are, or at least can be, normal problems of living. Normal problems of living may be harmful and distressing, but they only become mental disorders if they involve unskilled self-regulation (which requires a genuine *inability*$_2$ to intelligently modify, override, or change one's responses). Having difficulty regulating one's diet is not a mental disorder, being *unable*$_2$ to skillfully regulate one's food intake will be. And this is just what we should expect: mental disorders (such as eating disorders) involve genuine inabilites$_2$ to self-regulate, normal problems of living (such as having a problem limiting carbohydrates in one's diet) do not.

Finally, the skill view is not claiming that the act self-regulation, by itself, explains mental health. The act of self-regulation may be necessary for mental health, but it is clearly not sufficient; mental health, according to the view defended here, is *skilled* self-regulation. For example, an individual under the influence of intoxicants or psychotropic drugs may will to engage in harmful activities and regulate their responses towards this end. Similarly, a person experiencing a manic episode (which is often marked by grandiosity, flight of ideas, and increased goal-directed behavior) may successfully regulate her responses towards some mania-inspired standard.[52] Neither case of successful self-regulation is normally considered healthy

---

contrast, infants and small children are poor self-regulators not because of a breakdown or failure of ability$_{(1 \text{ and } 2)}$, but because of the lack of the *capacity* to skillfully self-regulate. Children do not have a problem of ability, but a (normal) lack of capacity.

[52] As noted in Section 4.2., this paper adopts a restricted definition of self-regulation that focuses is on whether individuals are able$_2$ to intelligently alter, modify, or control their responses to their mental phenomena, regardless

mental states. The difference between mere successful self-regulation to a potentially harmful standard, on one hand, and mental health, on the other, is that mental health requires the ability$_2$ to intelligently and flexibly modify one's behavior. This skill is often masked or absent in while individuals are experiencing manic states or are under the influence of psychotropic drugs.

4.3.4.1 Skill and Function

The skill view offers us a conceptually superior naturalist definition of mental disorders than traditional function-based views. The skill view, in contrast to function-based views, captures what is disordered about mental disorders (i.e., a failure of skilled self-regulation) without revising our conception of paradigmatic disorders and without pathologizing normal human difference. Unlike dysfunction views, the skill view predicts that paradigmatic mental disorders such as depression, anxiety, and schizophrenia are clear instances of disorder insofar as they involve failures of skilled self-regulation. And, unlike dysfunction views, the skill view predicts that mental states that merely involve norms-transgressing mental content and behavior (or even dysfunctional mental mechanisms), but do not involve problems of skilled self-regulation (such as homosexuality, asexuality, and norms-transgressing gender-identification), are clearly not

---

of the standard that they are attempting to regulate to. This is not to claim that the standard is irrelevant to health. It is clearly the case that individuals' goals or standard may be part of problem that therapy is attempting to ameliorate (e.g., overly high or 'perfectionist' standards). In the case of perfectionist or unrealistic goals, therapy may attempt to modify a patient's high standards, but the high standards will only be disordered if the patient is genuinely unable$_2$ to intelligently alter or control how she responds to stimuli. Impossibly high standards needn't be maladaptive, and may often serve as useful motivators. The standards that one regulates to has no necessary relationship to mental health. The skill view argues that mental health is determined by one's ability$_2$ to intelligently modify, alter, control their responses *towards* the achievement of some end, it does not matter (at least in the context of mental health) how well one does this or how achievable the goal is.

mental disorders. Mental disorders are best conceived of as problems of mental regulation, not mental content or the proper functioning of mental mechanisms.

The skill view of health, in contrast to function-based views, is also ecumenical to divergent theories of psychopathology and psychological healing. The causes of breakdowns in skilled self-regulation are going to various, and may include physiochemical, genetic, cognitive, or environmental factors (and likely some combination of multiple variables). A significant advantage of conceiving of mental disorders as problems of skill, rather than as the problem of the functioning of specific mental mechanisms, is that the identification of mental disorders does not depend upon the truth of any particular theory of psychopathology and psychological functioning. According to function-based views, accuracy about the etiology of hypothesized disorders is crucial, given that whether or not some way of being (such as anxiety or homosexuality) counts as a disorder necessarily depends on whether it is the result of a mental dysfunction. The skill view of health, on the other hand, is compatible with the etiology of mental disorders being an open empirical and conceptual question. Mental disorders are failures or breakdowns of skillful self-regulation, *whatever the cause.* Regardless of whether the etiology of a disorder is a best described by a cognitive-behavioral theory, psychodynamic theory, or any other theory of psychopathology, if some state is a disorder, it will be marked by an inability$_2$ to skillfully self-regulate.

4.3.4.2 The Epistemic Boundary Problem

The skill view draws the metaphysical boundary between health and disorder at individuals'

ability$_2$ to intelligently self-regulate. There remains the question of how we might know when this is the case. There are two types of epistemic challenges in identifying mental disorders: one is nosological, the other diagnostic. The nosological question is concerned with identifying what types of mental phenomena are and are not mental disorders (i.e., is the mental state normally associated with depression a disorder or a mere problem of living). The diagnostic question is concerned with how clinicians might identify disorders, given the answers to the nosological question (i.e., how might we know if an individual meets the criteria for any given mental state that we postulate to be disordered?). The section argues that there will be a degree of vagueness in the answer to both question, but, importantly, the epistemic uncertainty facing the skill view is more benign than the vagueness facing function-based views.

The skill view identifies mental disorders by identifying inabilities$_2$ to intelligently self-regulate. This requires identifying when individuals cannot flexibly alter or control how they respond to their thoughts, emotions, and behaviors. So, identifying specific disorders (such as depression, anxiety, and OCD) is just a matter of identifying inabilities to self-regulate (e.g., to regulate depressive thoughts and emotions, anxious feelings and worry, or obsessions and compulsions). Psychometric and diagnostic tools like the DSM-5 and ICD-10 are heuristics that are meant to do just this. While no one should think that these diagnostic tools are carving nature at its joints, they do a good, but imperfect job of identifying when individuals are unable$_2$ to intelligently self-regulate and are in need of help. Consider, for example, part of the DSM-5 diagnostic criterial for a major depressive episode (APA, 2013):

> Five (or more) of the following symptoms have been present during the same 2-week period

and represent a change from previous functioning; at least one of the symptoms is either (1) depressed mood or (2) loss of interest or pleasure.

1 Depressed mood most of the day, nearly every day, as indicated by either subjective report (e.g., feels sad, empty, hopeless) or observation made by others (e.g., appears tearful).
2 Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated by either subjective account or observation).
3 Significant weight loss when not dieting or weight gain (e.g., a change of more than 5% of body weight in a month), or decrease or increase in appetite nearly every day.
4 Insomnia or hypersomnia nearly every day.
5 Fatigue or loss of energy nearly every day.
6 Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick).
7 Diminished ability to think or concentrate, or indecisiveness, nearly every day (either by subjective account or as observed by others).
8 Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.

This is a plausible sketch of an individual unable[2] to regulate his or her behavior, emotions, and cognitions. Diagnostic tools will be imprecise. For instance, there is no metaphysical significance to the DSM-5's requirement of a minimum of two-weeks of symptom expression for the diagnosis of major depressive disorder, as opposed to three weeks or ten days. And

diagnostic tools may be based on mistaken nosological assumptions (e.g., homosexuality was considered a mental disorder in the DSM-II and was not fully removed until 1973). However, while the diagnostic tools' lack of precision clearly has practical implications (insofar diagnoses often entail improved access to resources and aid), this epistemic problem does not reflect a metaphysical shortcoming. For the skill view, the nosological question is a *practical* issue of identifying genuine inabilities$_2$ (e.g., we know that *if* someone is genuinely unable$_2$ to intelligently regulate their depressive thoughts and emotions, then they are experiences a mental disorder). In contrast, for function-based views, the epistemic uncertainty is nosology reflects a serious metaphysical failing (because we don't even know what kind of mental phenomena is and is not disordered because of our lack of knowledge about mental functions). Function-based theories posit *that* there is a difference between mental health and disorder, but they can only give us more or less plausible conjectures and inferences about what this might be and about the specific functions that any hypothesized mental mechanism is supposed to perform. The skill view, in contrast, provides us with a plausible framework to differentiate between accurate and mistaken nosological claims.

The skill view faces no unique diagnostic challenges. While the skill view and function-based views may differ over their nosological commitments, both views have to rely on the same diagnostic and psychometric tools (such as the DSM-5 and ICD-10, and the Beck Depression Inventory and the Hospital Anxiety and Depression Scale, respectively) in order to differentiate between health and disorder for any hypothesized nosological entity. For instance, even if one were to adopt a function-based conception of mental health, identifying the precise border between any particular individual's *poor* functioning (say, normal, but intense, anxiety) and *dys-*

functioning (say, GAD) will be still come down to the arbitrary drawing of lines. Similarly, identifying whether or not an individual is unable$_2$ to intelligently self-regulate (as opposed to just being a poorly skilled self-regulator) will not always be clear. But in both cases, this is an issue of psychometrics, not metaphysics. Providing a specific answer to the diagnostic question is going to be a problem for *any* naturalist view, and the existence of fuzzy boundaries between health and disorder for any particular nosological entity is only problematic if we take diagnosis to be accurately reflecting metaphysical distinctions rather than serving as imperfect heuristics to identify breakdowns of skilled action.

Both dysfunction views and the skill view accept that diagnostic boundaries will be fuzzy. This is a serious *practical* problem, but not a metaphysical one. What is crucial for a naturalist theory of mental disorder is that it provides a plausible objective metaphysical grounding from which we can base the epistemic differentiation of genuine disorders from mere problems of living and value-based 'disorders'. The skill view does this.

4.4 Pragmatics and the Skill View

The previous section has shown that the skill view offers us a more plausible answer to the boundary problems than either dysfunction or normative theories. This section will address some potential objections to the skill view.

A significant difference between the skill view and function-based theories is that the skill view is a theory of *mental* health, not health, full stop. Traditional function-based views adopt the

same conception of health for both somatic and mental health. This uniformity may be considered an advantage in their favor. However, the cost of this uniformity in our conception of somatic and mental disorders is a theory of disorder that is both potentially massively revisionary and epistemically weak. This is a high price. The focus on skill, rather than function, also more closely aligns with how we normally conceive of mental health and healing. Mental health is best conceived of as a quality of *persons*, not of impersonal mechanisms. Dysfunction views assume that the same standard of health can be used for *any* living thing (be it a plant, a bacterium, or a mind). In contrast, the skill view posits different standards of health for minds and bodies. It is far from clear why would we think otherwise? Why think that the questions concerning the health of a mind and the health of a plant or a bone are after the same information? Two concerns may be that adopting different standards of health assumes some sort of substance dualism about minds and bodies and an anti-scientific conception of mental disorders. But neither worry is legitimate. We needn't claim that minds are non-material to also posit that their health is best explained by skill rather than the functioning of traits in relation to evolved-for purposes or the non-agentive goals of reproduction and longevity. Minds are more complex than bones and plants, and it is not surprising that their health is best judged by different standards.

A second objection may be that the concept of normal function is necessarily connected to our understanding of health. The two most influential naturalist theories of health, Boorse's (1975) 'pure' naturalist theory, and Wakefield's (1992) 'hybrid' view, both appear to share this view. For example, Boorse states that:

> there is clearly some plausibility in the claim that the history of medical theory is nothing

> but a record of progressive investigation of normal functioning" (1977, p. 560)

And, Wakefield:

> the notions of function and dysfunction are central to the factual-scientific component of disorder (1992, p. 381)

> [it is a] virtual universal tendency to fall back on dysfunction to explain disorder (1992, p. 381).

Boorse and Wakefield may both be correct in claiming that health and disorder (or at least naturalist views of health and disorder) are normally conceptualized in functional terms. But, given the conceptual and epistemic problems facing function-based views, we have little reason to continue to do so. The strength of the intuitive pull of the presumed connection between normal function and health, if felt at all, should be dependent upon whether or not function-based conceptions of mental disorder can offer us explanatorily successful theories of health and disorder. The concept of natural function hasn't been able to do this. Skill can.

A final concern with the skill view is practical. It might be objected that the adoption of the skill view may increase negative stigma sometimes associated with mental disorders. The concern may be that the conception of mental disorders as failures or breakdowns of skilled action could encourage the idea that mental disorders are the results of personal weakness. If you want to become more skilled at some mode of functioning, the thought might go, you just need to

practice and put in the work to develop it. While it is certainly true that avoiding and recovering from mental illness is to some degree a matter of effort (this is something that most theories of psychological healing accept), conceiving of mental health as a skill does not entail that mental health is merely a question of willpower or effort.

There are constraints to the development of *any* skill that have nothing to do with strength of character or volition. Individuals will differ in their potentials to develop specific skills due to biological, cognitive, and environmental differences. The skill of mental health, like skill in sports and intellectual activities, will depend in large part on one's physical and cognitive capacities. While it is likely that most individuals can improve their ability to perform skilled actions through practice and study, there are many factors that will constrain one's development of skills. In the case of a sport such as tennis, enabling factors such as joint and bone structure, bodily coordination, height, vision, and the like will all contribute to, and limit, the development of one's playing skills. Practice and diligence will not overcome basic anatomical or physical dispositions or constraints, such as visual impairments or the loss of a leg. Similar constraints exist in developing cognitive and interpersonal social skills. For example, individuals with theory of mind deficits may have varying degrees of trouble developing life skills that require easily recognizing and understanding the intentions and thoughts of others. Individuals with these deficits can usually work to *improve* their interpersonal skills, but their ability to function highly in certain social tasks will often be limited (when compared to the general population). More generally, we should expect individuals to normally be able to improve their skill of self-regulation, but there will be constraints on their peak level of functioning that will vary across persons. The causes of this variance are likely very complex (including, but in no way limited to,

a person's prenatal environment, genetic endowment, parenting, and socio-economic standing), but clearly is not just a matter of will-power. The skill of mental health, like the other athletic or intellectual skills, is something that must be developed and maintained, and will be significantly constrained or enabled by physiological and environmental factors.

4.5 Addendum

This chapter has defended a value-neutral skill-based naturalist theory of mental health. It may be objected, however, that the skill view smuggles in normative concepts. Specifically, it may be objected that (some or all of) the concepts of skill, intelligence, ability, and self-regulation are value-laden and that the skill view of mental health, thus, is either a normative or hybrid view. This section explains why this concern is misplaced.

It is certainly true that all of these terms *can* be applied in a normative context. Displaying skill or intelligence, for example, is often met with praise and positive valuations, while a lack of ability or self-regulation is often (though not always) considered to be undesirable. But that we often value these concepts does not entail that they are ineluctably value-laden. As argued in sections 1.4.2 and 4.3.1, a theory of the metaphysics of some *x* need not be influenced by whether or not we value *x*. It is likely that values can (and often do) influence metaphysical theorizing, but this dissertation takes this as a reason for caution and intellectual humility, not as reason for the abandonment of value-neutral metaphysics. The naturalist skill view of mental health rests on the assumption that the project of value-neutral metaphysics is possible. It is beyond the scope of this dissertation to defend this view, but it is important to again note that

insofar as value-independent metaphysics is possible, there are no unique theoretical problems facing the metaphysics of health and disorder.[53]

The skill view claims that mental health is skilled self-regulation. The minimal definition of skill presented in 4.3.1 is composed of two parts (i.e., ability and intelligence), neither of which are defined in value-laden terms.[54] Self-regulation is defined in section 4.3.2 as the $ability_1$ to regulate one's responses to one's immediate willings. This definition is also value-neutral. An individual's values may play a role in the setting of the standard that she is attempting to regulate towards, but whether she is $able_2$ to self-regulate is a question of ability, not values. Put differently, what is being measured in self-regulation is whether an individual is $able_2$ to regulate her responses to *whatever* standard she sets; the standard itself is irrelevant to the identification of the $ability_2$ to self-regulate.[55] Values, then, play no necessary role in establishing whether one is able to skillfully self-regulate. There is, of course, room for values to enter into the process of *identifying* when one has this ability, but as argued in sections 4.3.2.1 - 4.3.2.2, this is a practical psychometric concern, not a conceptual one.

It is important to note, as Boorse (1997) does, that the choice to *treat* a disorder likely *is* a normative judgment. Choosing to promote mental health over mental disorder, and choosing to attempt to effect a change from disorder to health, reflects a value judgment preferring mental health to disorder. But the value-ladenness of the choice to treat mental disorders does not have

---

[53] See sections 1.4.2 and 4.3.1.
[54] See pp. 101-102.
[55] Though the standard is clearly relevant to whether one is in fact $able_2$ to skillfully self-regulate. For example, perfectionist standards may hinder one's ability to skillfully self-regulate. See pp. 113-114, footnote 52.

any necessary bearing on what a mental disorder is. The belief that mental health is preferable to mental disorder is a value-judgement. What it means to be healthy and disordered need not be.

Chapter 5: The Metaphysics of Skill

5.1 Introduction

Ch. 3 argues that mental healing is explained by improvement in skilled action while Ch. 4 argues that mental health is a skill. Both chapters adopt a theory-neutral conception of skill. The aim of this chapter is to show that the skill view of mental health (and the theory-neutral conception of skill from which it is based) is compatible with any plausible theory of skill.

5.2 Skill

Following Ryle (1946, 1949), skills have been traditionally associated with a kind of knowledge: knowledge-how (or know-how). The assumption here is that a person is skilled at some act, then, in some sense, she knows how to perform it. This conception of skill as a type of knowledge is supposed to account for the intuition that skilled action is *intelligent* action (as opposed to being merely reflexive or successful). Most theories of skilled action agree on this. The primary point of disagreement in metaphysical accounts of skill, however, is how to describe the nature of this knowledge. Theories of skilled action diverge into two camps: so-called intellectualists and anti-intellectualists. Roughly, intellectualist theories of skill postulate that know-how requires the involvement of factive propositional attitudes; to be skilled is to know facts about how to perform some action. Anti-intellectualists deny this claim.

The aim of this chapter is not to take a side in this debate, but to show that the skill view of

mental health is compatible with any plausible version of intellectualist and anti-intellectualist theories of skill.

5.3 Intellectualism

According to the intellectualist conception of skill, to be skilled is to know that for some action ϕ, and some way *w*, *w* is a way to ϕ (Stanley & Williamson, 2001, 2017; Stanley, 2011; Stanley & Krakauer, 2013). Put differently, the know-how involved in standard intellectualist accounts of skilled action is supposed to be *factive.* According to this view, an "action is skilled only if it is guided by knowledge of facts about ways of performing it" (Stanley 2011, p. 175). The factive-knowledge requirement for skill is supposed to account for the intelligence of skilled action in comparison with mere reflexive, automatic, or successful actions. For example, Stanley (2011) states that, 'what makes an action an exercise of skill, rather than mere reflex, is the fact that it is guided by the intellectual apprehension of *truths'* (p. 174). The idea here is that skilled, rather than reflexive, actions require the possessor to be able to intelligently navigate between relevant counterfactual options rather than mindlessly and automatically act. Being a skilled chess player is being guided by the knowledge of certain propositions about chess rules and strategies, and being a skilled tennis player is knowing that certain movements and strategies will normally lead to successful outcomes. And, according to intellectualists, these relevant options are propositionally represented facts about ways to perform actions. For example, Stanley and Krakauer (2013) state:

> In the case of virtually any activity ϕ, having skill at ϕ-ing requires knowing what to do to

> initiate actions of ϕ-ing. Such knowledge is propositional knowledge…Knowing what to do to initiate action is *clearly factual knowledge*; it is the knowledge that activities $x_1 \ldots x_n$ could initiate that action. It is a kind of factual knowledge required by skill possession (p.4).

Note that the intellectualist position is not claiming that the knowledge of facts required for skilled action need be articulated, or even articulable. We can know more than we can verbally communicate, and neither the development nor the exercise of a skill necessitates that one be able to explicitly entertain all the relevant counterfactual propositions. So, for example, a tennis player is supposed to know how to hit a ball only if she is guided by the knowledge of propositions such as 'if I start moving my hips early, they will pull my arms and impact my swing like so', where complex movements represented by demonstrative expressions like 'like so' needn't be articulable in detail). According to intellectualists, skillful action is skilled only insofar as it is produced *from* a knowledge base of propositionally encoded facts (regardless of whether one can verbalize these facts).

Moreover, *merely* being able to articulate facts about how to perform some act is not sufficient for intellectualist conceptions of skill. Memorizing an instructional book about how to play tennis (and thus knowing numerous facts about ways to successfully play the game) does not necessarily make one skilled a skilled player. Skill, under this view, requires that individuals know facts about how to perform some action under a 'practical mode of presentation' (Stanley & Williamson, 2001, 2017; Stanley, 2011). Thus, only a tennis player that knows that way *w* is a way for *her* to play tennis (presumably by learning this fact through practice), knows how to play

tennis.

Why think this? There are two primary arguments in support of intellectualism: (1) the linguistic argument, and (2) the intelligence argument. The linguistic argument states that the best semantic analysis of 'know-how' show that it is a species of 'knowledge-that'. Jason Stanley and Timothy Williamson have been the standard-bearers for this view (2001, 2017). Stanley and Williamson argue that formal semantic interpretations of know-wh constructions (such as knowledge where, knowledge why, and including knowledge how) all agree that to know-wh is to know the answer to a question. For example, we are told that:

> One knows when to raise one's mitt when a fly ball approaches in virtue of knowing a proposition that answers the question "when to raise one's mitt when a fly ball approaches?"...One knows where to move one's racket in returning a serve if and only if one knows a place that one's racket can occupy that will generally lead to returning a serve successfully. One knows where to go for a drink, if one knows of a place that remains open (say late at night). One knows where to go for a drink, in a different sense, if one knows where the good places to go for a drink are (even when every place is open and available) (Stanley & Williamson, 2017, p. 715)

And, clearly, one cannot *know* an answer to a question without knowing facts (e.g., about *where* to place one's mitt to catch a ball, or *when* to raise one's mitt to catch ball, or *how* to raise one's mitt to catch a ball). So, Stanley and Williamson conclude, know-how (and thus, skill) involves the knowledge of facts.

Assume for the sake of argument that Stanley and Williamson are correct in claiming that the formal semantic interpretation of knowledge-wh statements entail the knowledge of propositions.[56] This alone does not get us to the claim that skill requires propositional knowledge. The linguistic argument moves from a claim about how to formalize the way we use language to a conclusion about how the world is.[57]

However, the semantics of knowledge ascriptions has no necessary relation to the metaphysics of skill. For example, if our best formalization of know-how statements leads us to wildly implausible metaphysical claims about skill, then this should give us good reason to either question our linguistic analysis or our commitment to the claim that the type of know-how involved in skilled action requires the knowledge of facts. The linguistic argument, if sound, may offer *prima facie* evidence that skill is guided by propositional knowledge, but this jump from semantics to metaphysics needs support.

The intelligence argument is supposed to provide this support. Stanley and Williamson's version of the argument consists of the following two claims:

1. To be skilled at ϕ-ing requires that one intelligently ϕ.
2. Intelligently ϕ-ing requires acting on the basis of propositional knowledge of facts.

[56] This is far from a trivial assumption. For criticisms of the linguistic argument, see: Johnson, 2006; Glick, 2011, 2012; Michaelis, 2011; Ginzburg, 2011; Wiggins, 2012; Douskos, 2013; Abbott, 2013.
[57] Noë (2005) makes a similar point.

Both intellectualists and anti-intellectualists accept claim 1; skilled action is necessarily intelligent action. They differ, however, in how to specify the nature of this 'intelligence' (i.e., claim 2). Ryle (1946) provides a good description of the fact-based intellectualist conception of intelligence:

> [Intellectualist conceptions of intelligence] concentrate on the discovery of *truths* or *facts*, and they either ignore the discovery of *ways* and *methods* of doing things or else they try to reduce it to the discovery of *facts*. They assume that intelligence equates with the contemplation of propositions and is exhausted in this contemplation (p. 4, italics added)

Similarly, Stanley (2011) states that:

> it is only when our behavior is guided by intellectual recognition of truths that it deserves to be called "intelligent" (p. 190).[58]

The intelligence argument is the heart of the intellectualist view. It is based on the assumption that if action is not guided by propositional knowledge, then whatever is doing the guiding will be too *weak* to ground skill. For example, consider Ryle's (1949) *anti*-intellectualist conception of intelligence:

[58] Note that intellectualist theories are only committed to the claim that skill is 'guided' by truths. This does not imply that skilled action requires the prior *consideration* of truths.

> When a person is described by one or other of the intelligence-epithets such as 'shrewd' or 'silly', 'prudent' or 'imprudent', the description imputes to him not the knowledge, or ignorance, of this or that truth, but the ability, or inability, to do certain sorts of things. Theorists have been so preoccupied with the task of investigating the nature, the source and the credentials of the theories that we adopt that they have for the most part ignored the question what it is for someone to know how to perform tasks. In ordinary life, on the contrary, as well as in the special business of teaching, we are much more concerned with people's competences than with their cognitive repertoires, with the operations than with the truths that they learn.

Intellectualists argue that the reason we should be preoccupied with identifying the nature, source, and credentials underlying skilled action is that the focus on mere *ability* leaves us with an overly permissive (and overly unintelligent) conception of skill. For example, being able to successfully perform some act, by itself, does not appear to explain the skillful aspect of paradigmatic skilled actions. A person may be a very successful sports player because of her superior size and strength without being skilled, while a diminutive or physically handicapped person may possess a great deal of sporting skill but achieve little success. Similarly, if a chess player is mistaken about the rules of the game, it is hard to consider her a skilled player (even if she is somehow miraculously lucky and successful). It is also reasonable to expect a skilled dentist or surgeon to know facts about human anatomy (rather than just being successful at poking around based on mistaken assumptions). That skilled action must be guided by specific *facts*, and therefore true beliefs, is supposed to explain this perceived difference between success and skill. It also seems to be the case that differences in ability does not necessarily reflect

differences in skill. For example, tennis players of different genders may be equally as skilled while differing in their ability to hit with pace and run with speed. Intellectualist views are meant to account for these instances where skill and ability appear to come apart by identifying skill with the possession of genuine knowledge of truths, not merely ability to act.

5.3.1 Objections to the Intelligence Argument

Objections to the intelligence argument normally fall within two categories: (1) *thought experiments* that attempt to show cases that intuitively appear to be instances of skilled action, but are not guided by knowledge (because of knowledge-defeating factors such as manipulation, luck or some other Gettier-type situation), and (2) *real-world* cases that attempt to show that propositional knowledge is not necessary for skilled action.[59] We will address both categories in turn.

A common method of resisting the intelligence argument is to construct thought experiments that purport to that show skilled action comes apart from knowledge. For example, Stanley (2015) states:

[59] A third category of objections is inspired by Ryle-type regress arguments (Ryle, 1949; Stanley & Williamson, 2001; Fantl, 2011; Stanley, 2011; Weatherson, 2017). Roughly, these regress arguments start from the premises that (1) *if* intelligent action requires the prior consideration of propositionally represented truths, and (2) *if* the consideration of truths is something that can be done more or less intelligently, then (from 1&2) we appear to have a vicious regress (e.g., Stanley, 2011). The strength of the regress argument depends on whether intellectualists are justified in claiming that individuals act *on the basis of* propositional knowledge (without having to consciously entertain it), and thus denying that they are necessarily committed to something like (1). For the sake of argument, this chapter accepts that the regress can be avoided.

> ‘Possession of any skill requires knowledge about the activity, knowledge that is used to intentionally act. What would a counterexample to this thesis look like? It would be something that we would intuitively call a skill, but does not manifest as the result of decisions based on knowledge’ (2015, p. 322).

Intellectualist believe that no such counterexamples are on offer.[60] Stanley and Williamson (2001, p. 435) provide a representative attempt at such an argument:

> Bob wants to learn how to fly in a flight simulator. He is instructed by Henry. Unknown to Bob, Henry is a malicious imposter who has inserted a randomizing device in the simulator’s controls and intends to give all kinds of incorrect advice. Fortunately, by sheer chance the randomizing device causes exactly the same results in the simulator as would have occurred without it, and by incompetence Henry gives exactly the same advice as a proper instructor would have done. Bob passes the course with flying colors. He has still not flown a real plane. Bob has a justified true belief about how to fly. But there is a good sense in which he does not *know* how to fly.

Intuitions vary here. According to Stanley and Williamson, Bob does not know how to fly, and is thus not a skilled pilot, even though he is consistently able to successfully perform the action (in a simulator). On the other hand, it may seem intuitive that Bob *clearly* knows how to fly.[61] If

[60] This hasn’t stopped people from trying. For example, see: Wallis, 2008; Cath 2011, 2015; Brownstein & Michaelson, 2016.
[61] See, for example: Poston, 2009; Cath, 2011.

Bob can learn any number of new flying moves, control his plane, and adapt flexibly to new maneuvers and techniques, then it is not clear why the presence of luck in his training necessarily defeats his claim to skill (while still plausibly defeating his claim to knowledge). Stanley (2011) attempts to deflate this intuition by arguing that while it may seem natural (for some) to attribute know-how in cases like this, this is only because of the *pragmatics* of know-how attributions. In most cases of know-how attributions (such as whether or not someone knows how to find a restaurant or mend a shirt), we normally only care about whether the belief is true, not how individuals came to it. But, the argument goes, if the stakes are high (like when picking a surgeon or pilot), we typically start to care very much. We want our surgeons to not only be successful, but to also *know* what they are doing. And this is because, in cases like Bob's, 'the agent in fact does not know how to do the relevant action, but we easily allow the attributions because of the 'pragmatics of attributions of knowledge how' (2011. p. 189). According to Stanley (2011):

> Knowledge of how to do something is more valuable than true belief...We would not be as happy with [Bob] as our pilot as we would be with someone trained by aa skilled flight instructor even if we were antecedently assured that their beliefs about how to fly the plane are the same. We still find ourselves choosing the surgeon trained at the better institution, even if we were antecedently convinced that the two surgeons had the same beliefs about how to person the surgery, and the mechanisms that govern their execution equally fluid' (p. 181)

It is possible that Stanley is correct about the contextual differences in folk attributions of

knowledge, though on the other hand, it is unclear why the presence or absence of genuine knowledge should make a difference when there is (as hypothesized) *no* difference in ability. We needn't take a position here. The take-away point should be that intuitive evidence from hypothetical cases is far from definitive.

Stronger potential counterexamples to the intelligence argument come from real-world cases where know-how and knowledge appear to come apart without the (possibly intuition confounding) involvement of external manipulation or luck. The objection (dubbed by Glick as the 'arguments from cognitive science' [2011]) is simple: the cognitive science literature provides numerous cases where individuals intuitively know-how to ϕ but hold mistaken beliefs about the way in which they ϕ, and if knowing-how to ϕ requires that a skilled actor knows that some way *w* is a way to ϕ, then intellectualism is false (or at least problematic).[62] For example, Wallis (2008) argues that the literature on clinical diagnostic judgment shows that while clinicians may be skilled at diagnosing patients, they routinely misdescribe (and, thus, do not know) the methods from which they arrive at the diagnosis. Similarly, Brownstein and Michaelson (2016) note that athletes are routinely mistaken about how they perform certain actions. For example, baseball players often falsely believe that fast-pitched balls can rise while approaching the batter or that curved pitches break harder as they approach the batter (e.g., Shapiro et al., 2010). So, in both cases we have instances where folks are presumably skilled but appear to be 'guided by' false beliefs (and thus not knowledge) about the way in which they act. This, minimally, puts pressure on the intellectualist claim that skill at ϕ-ing requires being guided

[62] For similar arguments from cognitive science, see: Adams, 2009; Bzdak, 2008; Devitt, 2011.

by true beliefs about ways to ϕ.

Stanley's (2011) provides up with an intellectualist response to this type of counter-example:

> Wallis's…point is that someone may know how to ϕ, and on that basis ϕ, yet falsely describe the way they employ to ϕ. It is simply not clear why Wallis thinks this shows that the person in question lacks the propositional knowledge that is knowing how to ϕ, according to the account I have defended. Wallis thinks that if someone has false beliefs about the nature of a thing then they cannot have *de re* knowledge about that thing. But this view is absurd on the face of it. Suppose I falsely believe that what is in fact just the sofa in my living room is in fact an alien from Mars. I can still know that the object is grey, despite my false beliefs about its nature. The fact that I would falsely describe the sofa does not undermine the fact that I have knowledge about it. Similarly, I might very well have false descriptive beliefs about a certain way of ϕ-ing, while retaining my knowledge about that way of ϕ-ing, thought of demonstratively or practically, that it is a way to ϕ (pp. 166-167).

Stanley's point seems to be that *of course* we are often mistaken about our ways of ϕ-ing, but that does not mean that our ϕ-ings are not guided by knowledge of facts. These mistakes only show that the propositional knowledge guiding our behavior is often known only under practical modes of presentation. So, to use an example from Stanley (2011), a punch-drunk boxer may know how to fight southpaws (expressed demonstratively- 'like this') while being unable to articulate this knowledge. We can also assume that the boxer acts from this knowledge even

though he may be prone to confabulating reasons while explaining his behavior (e.g., 'you need to watch out because the left arm is always longer than the right' or 'lefties always have a stronger jab than righties'). None of this is supposed to take away from the fact that the boxer knows how to fight southpaws and that his factive knowledge of propositions about fighting guide his skilled boxing activities. And, importantly, in cases like the batting of balls it is this implicit knowledge that is *guiding* our behavior, not the mistaken descriptive beliefs about how we think we hit baseballs.

There is some plausibility to this response. It is not surprising that we are often unaware of the processes guiding our decision-making and fine-motor skills. And it is not surprising that we often confabulate responses. Neither fact obviously vitiates claims to skill. Again, we needn't take a position here. For the purposes of this chapter, we need only note that intellectualism either has the tools to handle real-world cases where explicit and implicit knowledge attributions diverge, or it is wrong (because it is open to counterexamples).

5.3.2 Intellectualism and The Skill of Mental Healing

On the face of it, there is some tension between intellectualist theories of skill and the skill view of mental health. Intellectualism states that skilled action must be guided by knowledge of truths. The skill view of mental health states that individuals can be skilled self-regulators while holding false beliefs about how they achieve this. This section addresses this tension and shows that the skill view is either compatible with intellectualism, or intellectualism about skill is false.

The act of therapist-lead mental healing is a paradigmatic skill. Psychotherapy is an act that individuals can be better or worse at performing, one that can improve with training and practice, and one that requires intelligently navigating interpersonal interactions. Psychotherapists also may be, and often are, skilled healers while also being mistaken about the causes, or underlying mechanisms, responsible for their successful outcomes. And if most forms of genuine psychotherapy are effective, then the success of a particular psychotherapy has no necessary connection to the truth of its supporting theory; at least most (if not all) psychotherapists are guided by false theories. So, the problem is this: if most (or all) psychotherapy is delivered on basis of false beliefs, and most (or all) healing is based on acceptance of falsities, then intellectualist theories appear to have difficulty explaining how both mental healing and mental can be skilled.

Intellectualist theories have three options. The first option is to deny that mental healing and health are skills. But this is absurd. If intellectualism is committed to the claim that it is likely that most (if not all) a mental healers or psychotherapists over the course of human history have been unskilled, then this is a *reductio ad absurdum* against the standard intellectualist account of skill. While there appears to be little to no variation in the effectiveness of different types of psychotherapy, there are significant differences between the effectiveness of individual therapists (e.g., Luborsky et al, 1997; Baldwin et al., 2007; Del Re et al., 2012). And these findings also align with common sense; some therapists are more skilled than others. This paper takes it as a given that mental healing is a paradigmatic skill, and thus any theory of skill that denies this has been guided to this conclusion by error.

The second option is more plausible. Intellectualist theories of skill can argue that, as in the case of the objections from cognitive science, mental healers are skilled insofar as they are being guided by implicit truths. However, it is not clear that this will work. Therapy, unlike the sporting cases, appears to be explicitly *guided* by false beliefs (as opposed to confabulated cased where beliefs are causally inert false descriptions of action). To see this, consider the delivery of psychodynamic therapy for panic disorder (PTDP).

Assume for the sake of argument that Freudian-inspired PTPD, a controversial but effective treatment, is based on false theoretical claims.[63] PTPD posits that panic symptoms are partly the result of 'intense angry feelings of which they are often totally or partly unware' (including 'unacknowledged rage' which is 'found to be an increasing important part of mental life at the time of panic onset') and unconscious fantasy (Busch et al, 1999, p. 235). Panic attacks are claimed to be an instance of what Freud called 'compromise formations', which symbolically represent a conflicting 'compromise between a forbidden wish and the defense against the wish' (Busch et al, 1999, p. 235). The therapy consists, in part, of practitioners guiding patients to uncover, then alter, the unconscious conflicts (hypothesized to usually involve the fear of separation, anger, or sex) causing the panic symptoms. This process of the discovery and modification of unconscious conflicts is normally facilitated by the identification of childhood causes and the focus on transference (the hypothesized tendency for individuals to unconsciously attribute aspects of important, often formative, relationships to unrelated, and in this case

[63] PTPD, while controversial, is listed as one of three empirically-supported treatments for panic disorder by the Division 12 of the APA.

therapeutic, relationships) in order to 'reexperience conflicts directly with their therapists in order for underlying fantasies to be articulated, understood, and rendered less magical and frightening' (Busch et al, 1999, p. 238).

Consider the following clinical example of a PTPD case formulation from Busch, Milrod, and Singer (1999):

> Ms. W., an 18-year-old woman, was driving from one city to another in order to attend her eighteenth birthday party when she experienced her first panic attack. The attack was so severe that she had to drive off the road, call her mother in the city toward which she was driving, and ask her mother to pick her up on the highway. The process of her mother's finding another person to drive with her who could also drive the car back took several hours, and in the meantime, Ms. W.'s party had to be canceled. At the moment that she experienced the attack, Ms. W. had found herself thinking that her eighteenth birthday was very important to her: it symbolized her "total independence" from her family and a new ability "to get rid of them." In the process of unraveling the onset of her illness later in psychotherapy, it became clear that in her fantasy, turning 18 and being "independent" represented the emotional equivalent of killing off her parents and siblings, all of whom enraged her. The fantasy was so full of conflict for her that she had her first panic attack. The panic symptoms represented both the wish to be alone and independent (suddenly she found herself, in fantasy, feeling entirely alone) and the defense against this wish: a sudden-onset, severe illness that made her "independence" from her family (and the very existence of her birthday celebration) impossible and

> effectively immobilized her escape/fantasy murder plan. Rather than a dangerous murderer, she was now helpless and ineffectual. Additionally, the panic represented a real way in which she effectively punished herself for her homicidal (and unacceptable) thoughts: now she could never be free of her family (p. 235).

As this clinical example demonstrates, PTPD therapeutic interventions are (or at least aim to be) explicitly guided by therapeutic commitments (e.g., uncovering subconscious fantasies and compromise formations). And, importantly, the delivery of PTPD is not a mindless, unintelligent, or reflexive action. Psychodynamic therapy requires therapists to respond to new patients with differing and often complex psychological problems, to formulate case-specific hypothesis about the causes of the patient's panic, and to be able to articulate psychodynamic explanations to the patient in a way that the patient finds convincing and motivating.

Nevertheless, intellectualists can argue that while the act of delivering psychotherapy may be guided in large part by false beliefs (about core beliefs, or conflicting unconscious drives, or whatever), the parts of therapy that are actually producing psychological change (and thus the parts of therapy which are 'skilled') are guided by propositional knowledge demonstratively represented under a practical mode of presentation (e.g., '*this* is a way for one to heal a patient'). This response is possible if we accept some kind of common-factors theory of psychological healing. So, in this case, being skilled at healing would reduce to being skilled at delivering the therapeutic common factors (e.g., something like creating a trusting, emotionally-charged bond and being persuasive). These skills can coherently be described as being guided demonstratively represented propositional knowledge. So, according to this view, therapy would be primarily

guided by false beliefs, but the *skilled* parts of therapy would be guided by justified true beliefs (known under a practical mode of presentation).

This response has an interesting consequence. An advocate of Stanley's implicit knowledge argument would have to claim that successful healers really *know* the underlying truth explaining the common-efficacy of psychotherapy, despite being unable to articulate it (or even know that they know). This would be a surprise to many. It is certainly surprising if it were the case that despite adhering to a particular theoretical orientation and performing theory-specific therapeutic techniques, all successful psychotherapists really had knowledge of the true theory of psychological change deep within their unconscious that only revealed itself through action. This is not absurd, but it is also not obviously true.

Advocates of intellectualism have a third option. It is possible to construct a weaker, and more plausible, intellectualist position that is compatible with the skill view of mental health. Intellectualists can adopt a less intellectually demanding version of intellectualism, and argue that skilled action is guided by propositional attitudes (e.g., something like justified or true beliefs), but not necessarily knowledge.[64] Stanley (2011) explicitly reject this weakened form intellectualism. He claims that:

> to argue that a condition weaker than knowledge is sufficient for skilled action one must not only argue that knowledge is more demanding than is required for skill, but one must

[64] Brogaard (2011) adopts this view.

> also make the case that we use the verb "know" to pick out both the more and the less demanding relation (p. 176).

Stanley, and most intellectualists, seem to think that there is intuitive evidence against thinking that either requirement can be met. This section has raised concerns that they are possibly mistaken about the first claim; there is reason to believe that knowledge is more demanding than skill. The second claim, that the verb 'know' is not used in reference to both genuine knowledge and mere justified or true belief, just begs the question against both weakened intellectualism and anti-intellectualism. Anti-intellectualists routinely use the verb 'know' to pick out both the more or less demanding relations! Going back at least to Ryle (1946), skill has been associated with a kind of 'know how', where the verb 'know' does *not* necessarily refer to knowledge of facts or to *any* propositional states.

Any plausible version of intellectualism, then, is compatible with the skill view of mental health. According to Stanley, "that someone skilled at an activity know how to do that activity is as good a candidate as any to be a conceptual truth" (2011, p. 175). This 'conceptual truth' can be made more or less plausible. If the know-how in question involves the knowledge of truths about how to perform some act, then (strong) intellectualism either makes the surprising prediction that all successful therapists actually know the true theory of psychological healing, or it predicts that the paradigmatic skill of mental healing is not in fact a skill (and is thus likely a false theory of skill). If the know-how in question is merely propositional states that are less demanding than knowledge, then (weak) intellectualism is plausible, but skill turns out to be less of an intellectual achievement then intellectualists normally assume. The relative strength of the two

versions of intellectualism needn't be resolved here. We just need note that the skill view of mental health is consistent with both weakened and strong intellectualism. The skill of mental health can be, and often is, guided by false theoretical assumptions, and the presumed falsity of the beliefs that guide the behavior of psychological healers should not necessarily undermine the attribution of skill to their behavior.

5.4 Anti-Intellectualism.

The skill of mental health view fits unproblematically with anti-intellectualism about skill. Both intellectualists and anti-intellectualists agree that what is important in matters of skilled action is that some action is intelligently performed, they disagree over what constitutes intelligent action. Intellectualist claim that intelligent action is guided by propositional knowledge, anti-intellectualists deny that this must be the case. Anti-intellectualists can accept that propositional knowledge may play a role in the *acquisition* of skill, while still denying skilled actions are necessarily *guided* by knowledge of facts (or any propositional states).[65] So, for example, the explicit learning and following of the rules and methods of a game like chess is normally a prerequisite for becoming a skilled player, but the anti-intellectualists deny that these rules continue to necessarily guide the actions of skilled players (e.g., Ryle, 1949; Dreyfus & Dreyfus, 1980).

[65] It is important to note that while Ryle thinks that the explicit learning of rules *may* play a role in skill development, he does not think that it is necessary for the development of skill (1949, p. 30).

What, then, is intelligent action if not action based on the knowledge of truths? According to Ryle (1946, 1949), intelligence is intimately connected with the ability to learn from experience and modify future actions accordingly. However, this flexibility is not supposed to require a cognitive "shadow-performance" of acting on the basis of propositionally represented knowledge-that. Skillful actors can be thought of as having internalized the rules and propositional knowledge that underwrite their acquisition of skill, rather than acting from them. Consider Ryle's (1949) description of an intelligent reasoner:

> The rules that he observes have become his way of thinking, when he is taking care; they are not external rubrics with which he has to square his thoughts. In a word, he conducts his operation efficiently, and to operate efficiently is not to perform two operations [both acting and considering propositions about acting]. It is to perform one operation in a certain manner or with a certain style or procedure…What is true of arguing intelligently is, with appropriate modifications, true of other intelligent operations. The boxer, the surgeon, the poet and the salesman apply their special criteria in the performance of their special tasks, for they are trying to get things right; and they are appraised as clever, skilful, inspired or shrewd not for the ways in which they consider, if they consider at all, prescriptions for conducting their special performances, but for the ways in which they conduct those performances themselves. Whether or not the boxer plans his manoeuvres before executing them, his cleverness at boxing is decided in the light of how he fights. If he is a Hamlet of the ring, he will be condemned as an inferior fighter, though perhaps a brilliant theorist or critic. Cleverness at fighting is exhibited in the giving and parrying of blows, not in the acceptance or rejection of propositions about blows, just as ability at

> reasoning is exhibited in the construction of valid arguments and the detection of fallacies, not in the avowal of logicians' formulae. Nor does the surgeon's skill function in his tongue uttering medical truths but only in his hands making the correct movements. All this is meant not to deny or depreciate the value of intellectual operations, but only to deny that the execution of intelligent performances entails the additional execution of intellectual operations (p. 48).

Similarly, Dreyfus & Dreyfus (1986) state:

> Most of us know how to ride a bicycle…Most of us are able to ride a bicycle because we possess something called "know-how," which we have acquired from practice and sometimes painful experience. That know-how is not accessible to us in the form of facts and rules. If it were, we could say we "know that" certain rules produce proficient bicycle riding. There are innumerable other aspects of daily life that cannot be reduced to "knowing that." Such experiences involve "knowing how" (pp. 86-87).

Intelligence, according to these views, is something that is *displayed* (somehow) in action and ability, rather than something encoded in the mind as factive attitudes. The anti-intellectualist position, thus, is not committed to a necessary relation between skilled action and knowledge of truths. In most activities, acting on the basis of true beliefs will be beneficial. We should hope that our surgeons and pilots know a great deal about the workings of their instruments. But this does not entail that skilled agents *must* act on the basis of true beliefs. As Ryle notes, the skill of the surgeon is a function of 'his hands making the correct movements', not of 'his tongue

uttering medical truths' (p. 49). The knowledge of truths, according to this view, often helps, but they do so only insofar as they aid intelligent, efficient, and flexible action.

5.4.1 Objections to Anti-Intellectualism

Objections to anti-intellectualism normally take something like the following form. If skilled action is intelligent action, and if intelligent action is (somehow) displayed or enacted in through ability, then differences in ability should reflect differences in skill. However, there appears to be situations where we find differences in ability but no differences in skill. Thus, skilled action cannot be mere 'exhibitions' of ability.[66] Stanley and Williamson (2017) offer us a representative articulation of this view when they state:

> Any view… which identifies skills with mere abilities or propensities to act, robs skill of any cognitive component. Consider the difference between someone who can bench-press a maximum of 100 pounds and someone who can bench-press 150 pounds. We may suppose that both employ the same technique; only brute strength makes the difference between them. Both are equally skilled, but clearly have different abilities. Similarly, there may be no distinction in skill between someone who runs a five minute mile and someone only capable of running a six minute mile. But there is clearly a difference in ability. Any view of skill must account for such cases. In particular, it must explain why strength, speed, and stamina are not themselves skills. Another way of putting the general

[66] For example, see: Stanley, 2011; Stanley & Williamson, 2017; Brogaard, 2011; Bengson & Moffett, 2011.

> demand on an account of skill is that it must explain what is distinctively *mental* about skill. It is presumably this feature that distinguishes skills from abilities merely due to brute strength, speed, and stamina (p. 721).

This is a genuine concern. The most natural anti-intellectualist response to this type of objection is to claim that mere strengths, speed, and stamina are only skilled if they are *intelligently* performed. But this type of response leaves it ambiguous, at best, how we are to differentiate intelligent action from merely successful action (given that intelligence is supposed to be manifested or enacted through successful action). Intellectualists, on the other hand, have a ready explanation: we can differentiate intelligence from mere success by appealing to the difference in knowledge guiding the two types of action.

Ryle (1949) offers the prototypical anti-intellectualist explanation of this difference:

> We observe, for example, a soldier scoring a bull's eye. Was it luck or was it skill?....To decide whether his bull's eye was a fluke or a good shot, we need and he himself might need to take into account more than this one success. Namely, we should take into account his subsequent shots, his past record, his explanations or excuses, the advice he gave to his neighbour and a host of other clues of various sorts. There is no one signal of a man's knowing how to shoot, but a modest assemblage of heterogeneous performances generally suffices to establish beyond reasonable doubt whether he knows how to shoot or not. Only then, if at all, can it be decided whether he hit the bull's eye because he was lucky, or whether he hit it because he was marksman enough to succeed when he tried….

> [The issue] is not one of the occurrence or non-occurrence of ghostly processes [i.e., acting on the basis of knowledge-that], but one of the truth or falsehood of certain 'could' and 'would' propositions and certain other particular applications of them (pp. 45-46).

The idea here is that we can distinguish skilled action from merely successful action by appealing to modal truths. There is some plausibility to this response. So, presumably, the difference in the weight bench-pressed by two people displaying the same technique will not reflect a difference in skill if there was no significant difference in their 'coulds' and 'woulds'. The challenge for anti-intellectualists is to specify which modal truths are relevant for skill. However, this paper need not weigh in on anti-intellectualism's potential for successes at meeting this challenge. The purpose of this chapter is not to argue that anti-intellectualism (or intellectualism) is likely true. Rather, the point here is to demonstrate that the skill view of mental health is compatible with any plausible theory of skilled action (be it intellectualist or anti-intellectualist). It clearly is.

Chapter 6: Conclusion

6.1 Conclusions

This dissertation introduces a novel skill-based theory of mental health and disorder. Chapters 2-3 motivate the skill view of mental *health* by first examining the process of mental *healing.* Chapters 4 and 5 outline and defend the skill view.

Chapter 2 was a case study examining the theoretical foundations of CBT (the dominant explanation for mental healing in North America). Chapter 2 argued that CBT is likely based on mistaken theoretical assumptions. The CBT model is based on the idea that consciously accessible core beliefs filter and assign meaning to all stimuli that impinges on an organism. The goal of CBT is to identify maladaptive core beliefs, and replace them with more adaptive cognitive interpretations. While CBT is clearly effective as a treatment, this chapter argued that there is good reason to be skeptical that its efficacy is due to the causal mechanisms posited by the CBT model. This chapter argued that the specific cognitive schemas posited by the CBT model likely do not play a direct role in the development or treatment of psychological illness. Cognitive schemas, as identified in CBT interventions, are likely to be the result of patient confabulation and epistemically under-supported practitioner-based identification. CBT interventions appear to impose coherence on patients' psychological states, rather than identifying and modifying preexistent causally efficacious core beliefs.

Chapter 3 provided an explanation for the findings of common efficacy in psychotherapy. CBT, and other forms of genuine psychotherapy, appear to be effective despite being based on false theories of psychological change. A prominent explanation for the findings of common efficacy in psychotherapy is to postulate that all successful therapies work by altering maladaptive meanings and providing patients with new, more adaptive meanings. This chapter argued that the 'meaning view' of psychological change is likely mistaken; psychological problems are not normally problems of meaning nor are they directly ameliorated by changes in meaning. This chapter then outlined a skill-based explanation for the findings of the common efficacy of psychotherapy.

Chapter 4 built upon Chapter 3's skill-based explanation of psychological change to develop a novel skill-based naturalist theory of mental health and disorder. Chapter 5 provided metaphysical support for this theory. The novel skill view of mental health provides a naturalist framework for the scientific study and treatment of mental disorders that avoids the explanatory and conceptual failings of traditional naturalist and normative views. The focus on skilled action, rather than the adherence to either biological norms or value norms, allows for a theory of health that avoids pathologizing normal human difference while also providing a non-revisionary accounting of paradigmatic mental disorders.

References

Abbott, B. (2013). Linguistic solutions to philosophical problems: The case of knowing how. *Philosophical Perspectives 27*(1), 1-21.

Adams, M. P. (20009). Empirical evidence and the knowledge-that/knowledge-how distinction. *Synthese, 170*(1), 97-114.

Alford, B. A., & Beck, A. T. (1998). *The Integrative Power of Cognitive Therapy*. New York, NY: Guilford Press.

Allen, N. B., & Badcock, P. B. T. (2003) The social risk hypothesis of depressed mood: evolutionary, psychosocial, and neurobiological perspectives. *Psychological Bulletin, 129*(6), 887-913.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (Rev 3rd ed.). Washington, DC: American Psychiatric Association

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Anderson, T., Lunnen, K.M., & Ogles, M. (2010). Putting models and techniques in context. In B. L. Duncan, S. D. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart and should of change: Delivering what worlds in therapy (2nd ed.)* (pp. 143-166). Washington, DC: American Psychological Association.

Andrews, P. W., & Thomson, J. A., (2009). The bright side of feeling blue: Depression as an adaptation for analyzing complex problems. *Psychological Review, 116*(3), 620-554.

Asay, T. P., & Lambert, M. J. (1999). The empirical case for the common factors in therapy: Quantitative findings. In M. A. Hubble, B. L. Duncan, & S. D. Miller (Eds.), *The heart*

and soul of change: What works in therapy (pp. 23–55). Washington, DC: American Psychological Association.

Baldwin, S. A., Wampold, B. E., & Imel, Z. E. (2007). Untangling the alliance-outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology, 75*, 842-852.

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

Baumeister, R. F., & Heatherton, T. (1996). Self-regulation failure: An overview. *Psychological Inquiry 7*(1), 1-15.

Baumeister, R. F., Heatherton, T., & Tice, M. (1993). When ego threats lead to self-regulation failure: Negative consequences of high self-esteem. *Journal of Personality and Social Psychology, 64*(1), 141–156.

Baumeister, R. F., Heatherton, T., & Tice, M. (1994). *Losing Control.* San Diego, CA: Academic Press.

Baumeister, R. F., & Tice, M. (1990). Point-counterpoints: Anxiety and social exclusion. *Journal of Social and Clinical Psychology 9*(2), 165-195.

Beck, A. T. (1964). Thinking and depression: II. Theory and therapy. *Archives of General Psychiatry, 10*(6), 561-571.

Beck, A. T. (1967). *Depression: Clinical, Experimental, and Theoretical Aspects*. Philadelphia, PA: University of Pennsylvania Press.

Beck, A. T. (1976). *Cognitive Therapy of the Emotional Disorders.* New York, NY: New American Library.

Beck, A. T. (1979). *Cognitive Therapy and the Emotional Disorders.* New York, NY: Penguin.

Beck, A. T. (1987). Cognitive models of depression. *Journal of Cognitive Psychotherapy, 1*(1), 5–37.

Beck, A. T. (1991). Cognitive therapy as *the* integrative therapy. *Journal of Psychotherapy Integration 1*(3), 191-198.

Beck, A. T. (1997). Cognitive therapy: reflections. In J. K. Zeig (Ed.), *The Evolution of psychotherapy: the third conference* (pp. 55-64). New York, NY: Brunner/Mazel.

Beck, A. T. (2004). A Cognitive Continuum. In A. Freeman, M. Mahoney, P. Devito, & D. Martin (Eds.), *Cognition and Psychotherapy* (pp. 107–220). New York, NY: Springer.

Beck, A. T. (2005). The current state of cognitive therapy: a 40-year retrospective. *Archives of General Psychiatry, 62*(9), 953–959.

Beck, A. T., & Alford, B. A. (2009). *Depression: Causes and Treatment.* Philadelphia, PA: University of

Beck, A. T., Davis, D. D., & Freeman, A. (2015). *Cognitive Therapy of Personality Disorders* (3rd ed.). New York, NY: Guilford.

Beck, A. T., & Dozois, D. J. (2011). Cognitive therapy: current status and future directions.

Beck, A. T., & Steer, R. A. (1987). *BDI, Beck depression inventory: manual*. Psychological Corporation San Antonio, TX.

Beck, A. T., Rush, J. A., Shaw, B. F., & Emery, G. (1979). *Cognitive Therapy of Depression*. New York, NY: Guilford Press.

Beck, J. S. (2011). *Cognitive Behavior Therapy,* Second Edition: Basics and Beyond. New York, NY: Guilford Press.

Beevers, C. G., Strong, D. R., Meyer, B., Pilkonis, P. A., & Miller, I. W. (2007). Efficiently assessing negative cognition in depression: An item response theory analysis of the Dysfunctional Attitude Scale. *Psychological Assessment, 19*(2), 199–209.

Bengson, J., & Moffett, M. (2011). Non-propositional intellectualism. In M. Bengson (Ed.), *Knowing how: Essays on knowledge, mind, and action* (pp. 136-160). Oxford: Oxford University Press.

Bergin, A. E., & Garfield, S. L. (1994). Overview, trends, and future issues. In, A. E. Bergin & S. L. Garfield (Eds.), *Handbook of Psychotherapy and Behavior Change* (4th ed., pp. 821-830). New York, NY: Wiley.

Berofsky, B. (2001). If, cans, and free will: The issues. In R. Kane (Ed.) *The Oxford Handbook of Free Will*. Oxford: Oxford University Press.

Bertrand, M., & Mullainathan, S. (2003). *Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination.* National Bureau of Economic Research (Working Paper 9873).

Bloom, P. (2007). Religion is natural. *Developmental Science 10*(1), 147-151.

Bolton, D. (2001). Problems in the definition of 'mental disorder'. *The Philosophical Quarterly, 51*(203), 182-199.

Boorse, C. (1976). What a theory of mental health should be. *Journal for the Theory of Social Behaviour, 6,* 61-84.

Boorse, C. (1977). Health as a theoretical concept. *Philosophy of Science, 44*(4), 542-573.

Boorse, C. (1997). A rebuttal on health. In J. M. Humber and R. F. Almeder (Eds.), *What is Disease?* (pp 1-134). Totowa, NJ: Humana Press.

Boorse, C. (2014). A second rebuttal on health. *Journal of Medicine and Philosophy, 39*(6), 683-724.

Boyer, P. (2003). Religious thought and behavior as by-products of brain function. *Trends in Cognitive Sciences 7*(3), 119-124.

Brach, T. (2003). *Radical Acceptance: Embracing your life with the heart of a Buddha.* New York, NY: Bantam.

Brewin, C. R. (1996). Theoretical foundations of cognitive-behavior therapy for anxiety and depression. *Annual Review of Psychology, 47*(1), 33–57.

Brogaard, B. (2011). Non-propositional intellectualism. In M. Bengson (Ed.), *Knowing how: Essays on knowledge, mind, and action* (pp. 136-160). Oxford: Oxford University Press.

Brownstein, M., & Michaelson, E. (2015). Doing without believing: Intellectualism, knowledge-how and belief-attribution. *Synthese 193*(9), 2815-2836.

Busch, F. N., Milrod, B. L., & Singer, M. B. (1999). Theory and technique in psychodynamic treatment of panic disorder. *The Journal of Psychotherapy Practice and Research, 8*(3), 234.

Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review, 26*(1), 17–31.

Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review, 26*(1), 17–31.

Bydak, D. (2008). On amnesia and knowing-how. *Techné: Research in Philosophy and Technology, 12*(1), 36-47.

Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality – social, clinical, and health psychology. *Psychological Bulletin 92*, 111-135.

Cath, Y. (2011). Knowing how without knowing that. In M. Bengson (Ed.), *Knowing how: Essays on knowledge, mind, and action* (pp. 113-135). Oxford: Oxford University Press.

Cath, Y. (2011). Revisionary intellectualism and Gettier. *Philosophical Studies, 172*(1), 7-27.

Chang, E. C. (1996). Cultural differences in optimism, pessimism, and coping: Predictors of subsequent adjustment in Asian American and Caucasian American college students. *Journal of Counseling Psychology (43)*, 113-123.

Chang, E. C. (2002). Cultural differences in psychological distress in Asian and Caucasian American college students: Examining the role of cognitive and affective concomitants. *Journal of Counseling Psychology*, *49*(1), 47.

Charland, L. C., (2004). Moral treatment and the personality disorders. In J. Radden (Ed.), *The Philosophy of Psychiatry: A Companion* (pp. 64-77). Oxford: Oxford University Press.

Clark, D. A. (2004). *Cognitive-behavioral therapy for OCD*. New York, NY: Guilford Press.

Clark, D. A. & Beck, A. T. (1999). *Scientific Foundations of Cognitive Theory and Therapy*

Clark, D. A. & Beck, A. T. (2010). *Cognitive therapy of anxiety disorders: science and practice*. New York, NY: Guilford Press.

Clouser, K. D., Culver, C. M., & Gert, B. (1981). Malady: a new treatment of disease. *The Hastings Center Report 11*(3), 29-37.

Clouser, K. D., Culver, C. M., & Gert, B. (1997). Malady. In J. M. Humber & R. F. Almeder (Eds.) *What is a disease?* Totowa, NJ: Humana Press

Cooper, R. (2002) Disease. *Studies in the History and Philosophy of Biological and Biomedical Sciences, 33*(2), 263-282.

Cooper, R. (2005). *Classifying Madness.* Dordrecht, The Netherlands: Springer.

Coyne, J. C., & Gotlib, I. H. (1983). The role of cognition in depression: A critical appraisal. *Psychological Bulletin, 94*(3), 472.

Crits-Christoph, P. (1997). Limitations of the dodo bird verdict and the role of clinical trials in psychotherapy research: Comments on Wampold et al. (1997). *Psychological Bulletin, Vol 122(3),* 216-220.

Cuijpers, P., van Straten, A., Anderson, G., & van Oppen, P. (2008). Psychotherapy for depression in adults: A meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology, 76*(6), 909-922.

Cuijpers, P., Smit, F., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2010). Efficacy of cognitive-behavioural therapy and other psychological treatments for adult depression: meta-analytic study of publication bias. *The British Journal of Psychiatry: The Journal of Mental Science, 196*(3), 173–178.

Culver, C. M., & Gert, B. (1982). *Philosophy in Medicine.* New York, NY: Oxford University Press.

Davis, T. (2017). The Goldberg exaptation model: Integrating adaptation and by-product theories of religion. *Review of Philosophy and Psychology 8*(3), 687-708.

Del Re, A. C., Flückiger, C., Horvath, A. O., Symonds, D., & Wampold, B. E. (2012). Therapist effects in the therapeutic alliance-outcome relationship: A restricted-maximum likelihood meta-analysis. *Clinical Psychology Review, 32,* 642-649.

DeRubeis, R. J., Webb, C. A., Tang, T. Z., & Beck, A. T. (2010). Cognitive therapy. In K. S. Dobson (Ed.), *Handbook of Cognitive Behavioral Therapies* (pp. 277-316). New York, NY: Guiford.

Devitt, M. (2011). Methodology and the nature of knowing how. *The Journal of Philosophy, 108*(4), 205-218.

Dobson, K. S. (Ed.). (2010). *Handbook of Cognitive-Behavioral Therapies, Third Edition*. New York, NY: Guilford Press.

Douskos, C. (2013). The linguistic argument for intellectualism. *Synthese 190*(12), 2325-2340.

Dreyfus, S. E., & Dreyfus, H. L. (1980). *A five-stage model of the mental activities involved in directed skill acquisition.* Berkeley, CA: University of California Berkeley Operations Research Center.

Dreyfus, S. E., & Dreyfus, H. L. (1986). Why expert systems do not exhibit expertise. *IEEE expert, 1*(2), 86-90.

Eagle, M. N., & Wolitzky, D. L. (1992). Psychoanalytic theories of psychotherapy. In D. K. Freedheim, H. J. Freudenberger, J. W. Kessler, S. B. Messer, D. R. Peterson, H. H. Strupp, & P. L. Wachtel (Eds.), *History of psychotherapy: A century of change* (pp. 109–158). Washington, DC: American Psychological Association.

Ellis, A. (1962). *Reason and emotion in psychotherapy.* New York, NY: Lyle Stuart.

Ellis, A. (1994). *Reason and emotion in psychotherapy* (Revised and Updated). New York, NY: Carol Publishing.

Fantl, J. ((2011). Ryle's regress defended. *Philosophical Studies, 156*(1), 121-130.

Frank, J. D. (1961). *Persuasion and healing.* Baltimore, MD: Johns Hopkins Press.

Frank, J. D. (1995). Psychotherapy as rhetoric: Some implications. *Clinical Psychology: Science and Practice 2*(1), 90-93.

Frank, J. D., & Frank, J. B. (1991). *Persuasion and healing: A comparative study of psychotherapy (3rd ed.).* Baltimore, MD: Johns Hopkins University Press.

Frankfurt, H. (1982). The importance of what we care about. *Synthese 53*(2), 257-272.

Fredrickson, L. B., Cohn, M. A., Coffey, K. A., Pek, J., & Finkel, S. M. (2008). Open Hearts build lives: Positive emotions, induced through loving-kindness meditation, build consequential personal resources. *Journalof Personality and Social Psychology, 95*(5), 1045-1062.

Freud, S. (2000). *Three Essays on the Theory of Sexuality* (J. Strachey, Trans). New York, NY: Basic Books.

Fujita, K. (2011). On conceptualizing self-control as more than the effortful inhibition of impulses. *Personality and Social Psychology Review 15*(4), 352-366.

Fulford, K. W. M. (1989). *Moral Theory and Medical Practice*. Cambridge: Cambridge University Press.

Gabbard, G. O. (2000). *Psychodynamic psychotherapy in clinical practice,* 3rd. ed. Washington, DC: American Psychological Association.

Garson, J. (2008). Function and Teleology. In A. Plutynski & S. Sarkar (Eds.) *Companion to Philosophy of Biology* (pp. 525-549). Malden, MA: Blackwell.

Garson, J, & Piccinini, G. (2013). Functions must be performed at appropriate rates in appropriate situations. *The British Journal for the Philosophy of Science, 65*(1), 1-20.

Gipps, R. G. T. (2013). Cognitive behavior therapy: a philosophical appraisal. In K. Fulford, M. Davies, R. Gipps, G. Graham, J. Sadler, G. Stanghellini, & T. Thorton (Eds.), *The Oxford Handbook of Philosophy and Psychiatry*. Oxford: Oxford University Press.

Glick, E. (2011). Two methodologies for evaluating intellectualism. *Philosophy and Phenomenological Research, 83*(2), 398-434.

Glick, E. (2012). Abilities and know-how attributions. In J. Brown & M. Gerken (Eds.), *Knowledge Ascriptions* (pp. 120-139). Oxford: Oxford University Press.

Ginzburg, J. (2011). How to resolve how to. In M. Bengson (Ed.), *Knowing how: Essays on knowledge, mind, and action* (pp. 136-160). Oxford: Oxford University Press.

Goldfried, S. L. (1980). Toward the delineation of therapeutic change principles. *American Psychologist 35*, 991-999.

Graham, G. (2013). *The Disordered Mind: An Introduction to Philosophy of Mind and Mental Illness* (2nd Revised Edition). New York, NY: Routledge.

Greenberg, L. (2010). *Emotion focused therapy. Theory and Practice.* Washington, DC: American Psychological Association.

Greenhouse, W. J., Meyer, B. & Johnson, S. L. (2000) Coping and medication adherence in bipolar disorder. *Journal of Affective Disorders*, *59*(3), 237-241.

Grissom, R. J. (1996). The magic number .7+-.2: Meta-meta-analysis of the probability of superior outcome in comparisons involving therapy, placebo, and control. *Journal of Consulting and Clinical Psychology, 64*, 973-982.

Gruber, J. (2011). Can feeling too good be bad? Positive emotion persistence (PEP) in bipolar disorder. *Current Directions in Psychological Science 20*(4), 217-221.

Gruber, J. Mauss, I. B., & Tamir, M. (2011). A dark side of happiness? How, when, and why happiness is not always good. *Perspectives on Psychological Science 6*(3), 223-233.

Haaga, D. A., Dyck, M. J., & Ernst, D. (1991). Empirical status of cognitive theory of depression. *Psychological Bulletin*, *110*(2), 215–236.

Haidt, J. (2006). *The Happiness Hypothesis*. Cambridge, MA: Basic Books.

Haidt, J. (2012). *The righteous mind*. New York, NY: Pantheon.

Harden, C. D., & Banaji, M. R. (2013). The nature of implicit prejudice: Implications for personal and public policy. In E. Shafir (Ed.), *The behavioral foundations of public policy* (pp. 13-31). New Jersey: Princeton University Press.

Hardin, E. E., & Leong, F. T. L. (2005). Optimism and pessimism as mediators of the relations between self-discrepancies among Asian and European Americans. *Journal of Counseling Psycholgoy 52*(1), 25-35.

Hausman, D. M. (2012). Health, naturalism, and functional efficiency. *Philosophy of Science, 79*(4), 519-541.

Hausman, D. M. (2014). Health and Functional Efficiency. *Journal of Medicine and Philosophy, 39*(6), 634-647.

Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and Commitment Therapy: Model, processes and outcomes. *Behaviour Research and Therapy, 44*(1), 1–25.

Hubble, M. A, Duncan, B. L., & Miller, S. D. (1999). Directing attention to what works. In, M. A. Hubble, B. L. Duncan, & S. D. Miller (Eds.), *The heart and soul of change: What works in therapy* (pp. 407–447). Washington, DC: American Psychological Association.

Hurlburt, R. T. (2011). *Investigating pristine inner experience: Moments of truth.* Cambridge University Press.

Isen, A. M., & Levin, P. F. (1972). Effect of feeling good on helping: Cookies and kindness. *Journal of Personality and Social Psychology, 21*(3), 384–388.

Johansson, P., Hall, L., Sikström, S., & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science, 310*(5745), 116–119.

Johnson, K. (2006). Externalist thoughts and the scope of linguistics. *Protosociology 22,* 19-35.

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Kendell, R. E. (1975). The concept of disease and its implications for psychiatry. *British Journal of Psychiatry, 127*(4), 305-315.

Kendell, R. E., (2002). The distinction between personality disorder and mental illness. *British Journal of Psychiatry, 180*(2), 110-115.

Kingma, E. (2014). Naturalism about health and disease: Adding nuance for progress. *Journal of Medicine and Philosophy 39*(6), 590-608.

Knowles, E. D., Lowery, B. S., & Schaumberg, R. L. (2010). Racial prejudice predicts opposition to Obama and his health care reform plan. *Journal of Experimental Social Psychology, 46*(2), 420-423.

Kraemer, D. M. (2013). Statistical theories of functions and the problem of epidemic disease. *Biology & Philosophy*, *28*(3), 423-438.

Krantz, S., & Hammen, C. L. (1979). Assessment of cognitive bias in depression. *Journal of Abnormal Psychology*, *88*(6), 611.

Lambert, M. J. (1992). Psychotherapy outcome research: Implications for integrative and eclectic therapists. In J. C. Norcross & M. R. Goldried (Eds.), *Handbook of psychotherapy integration* (pp. 94-129). New York, NY: Basic Books.

Lambert, M. J. (2013). The efficicacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change, 6th ed.* (pp. 169-218). New York, NY: Wiley.

Leahy, R. L. (1996). *Cognitive Therapy: Basic Principles and Applications*. New Jersey: Jason Aronson, Incorporated.

Lebowitz, M. S. (2014). Biological conceptualizations of mental disorders among affected individuals: A review of correlates and consequences. *Clinical Psychology: Science and Practice, 21*(1), 67-83.

Leder, G. (2017). Know Thyself? Questioning the theoretical foundations of cognitive behavioral therapy. *Review of Philosophy and Psychology 8*(2), 391-410.

Leichsenring, F., Hiller, W., Weissberg, M., & Leibing, E. (2006). Cognitive-behavioral therapy and psychodynamic psychotherapy: techniques, efficacy, and indications. *American Journal of Psychotherapy, 60(3),* 233-259.

Lilienfeld, S. O., & Marino, L. (1995). Mental disorder as a Roschian concept: a critique of Wakefield's "harmful dysfunction" analysis. *Journl of Abnormal Psychology, 104*, 411-420.

Liu, J., Smeesters, D., & Vohs, K. (2012). Reminders of money elicit feelings of threat and reactance in response to social influence. *Journal of Consumer Research, 38(6),* 1030-1046.

Luborsky, L., McLellan, A. T., Diguer, L., Woody, G., & Seligman, D. A. (1997). The psychotherapist matters: comparison of outcomes across twenty-two therapists and seven patient samples. *Clinical Psychology: Science and Practice, 4*, 53-65.

Luborsky, L., Rosenthal, R., Diguer, L., Andrusyna, T. P., Berman, J. S., Levitt, J. T., et al. (2002). The Dodo Bird Verdict Is Alive and Well—Mostly. *Clinical Psychology: Science and Practice*, *9*(1), 2–12.

Marcus, D. K., O'Connell, D. O., Norris, A. L., Sawaqdeh, A. (2014). Is the Dodo bird endangered in the 21st century? A meta-analysis of treatment comparison studies. *Clinical Psychology Review, 34*(7), 519-530.

Margolis, J. (1976). The concept of disease. *The Journal of Medicine and Philosophy 1*(3), 238-255.

Margolis, J. (1976). The concept of disease. *The Journal of Medicine and Philosophy 1*(3), 238-255.

Margolis, J. (1980). The concept of mental illness: a philosophical examination. In B. A. Brody & H. T. Engelhardt (Eds.), *Mental Illness: Law and Public Policy* (pp. 2-23). Dordrecht, Netherlands: Reidel Publishing Company.

Martin, M. (1985). Malady and menopause. *The Journal of Medicine and Philosophy 10*(4), 329-338.

McEachrane, M. (2009). Capturing emotional thoughts: the philosophy of cognitive-behavioral therapy. In Y. Gustafsson, C. Kronqvist, & M. McEachrane (Eds.), *Emotions and Understanding: Wittgensteinian perspectives.* New York: Palgrave Macmillan.

Messer, S. B., & Wampold, B. E. (2002). Let's Face Facts: Common Factors Are More Potent Than Specific Therapy Ingredients. *Clinical Psychology: Science and Practice, 9*(1), 21–25.

Michaelis, L. A. (2011). Knowledge ascription by grammatical construction. In M. Bengson (Ed.), *Knowing how: Essays on knowledge, mind, and action* (pp. 136-160). Oxford: Oxford University Press.

Miller, S., Duncan, B. L., & Hubble, M. A. (2005). Outcome-informed clinical work. In J. C. Norcross & M. R. Goldried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 84-102). New York, NY: Oxford University Press.

Murphy, D. (2006). *Psychiatry in the Scientific Image.* Cambridge, MA: MIT Press.

Murphy, D., & Woolfolk, R. L. (2000). The harmful dysfunction analysis of mental disorder.

Nasar, J. L., & Devlin, A. S. (2011). Impressions of psychotherapists' offices. *Journal of Counseling Psychology, 58*(3), 310–320.

Neenan, M. & Dryden, W. (2005). *Cognitive Therapy.* SAGE Publications.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231–259.

Noë, A. (2005). Anti-intellectualism. *Analysis 65*(288), 278-290.

Norcross, J. C., & Karpiak, C. P. (2012). Clinical psychologists in the 2010s: 50 years of the APA division of clinical psychology. *Clinical Psychology: Science and Practice 19*(1), 1-12.

Nordenfelt, L. Y. (1995). *On the nature of health: an action-theoretic approach*. Dordrecht, Netherlands: Kluwer.

Nosek, B. A, Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., Smith, C. T., Olson, K. R., Chugh, D, Greenwald, A. G., & Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology 18*(1), 36-88.

Nosek, B. A., Banaji, M. & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beleifs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice, 6*(1), 101.

Orlinsky, D. E., & Howard, K. I. (1986). Process and outcome in psychotherapy. In S. L. Garfield, & A. E. Bergin (Eds.) *Handbook of psychotherapy and behavior change*. New York, NY: Wiley.

Pescosolido, B. A., Martin, J. K., Long, J. S., Medina, T. R., Phelan, J. C., & Link, B. G. (2010). "A disease like any other"? A decade of change in public reactions to schizophrenia, depression, and alcohol dependence. *American journal of Psychiatry, 167*, 1321-1330. *Philosophy, Psychiatry, & Psychology 7*(4), 241-252.

Pickard, H. (2009). Mental illness is indeed a myth. In M. Broome, & L. Bortolotti (Eds.), *Psychiatry as Cognitive Neuroscience: Philosophical Perspectives.* Oxford: Oxford University Press.

Prinz, J. (2012). *Beyond human nature: How culture and experience shape our lives.* London, UK: Allen Lane.

Rachman, S., & de Silva, P. (1978). Abnormal and normal obsessions. (*Behaviour Research and Therapy 16*, 411-423.

Radomsky, A. S., Alcolado, G. M., Abramowitz, J. S., Alonso, P., Belloch, A., Bouvard, M., … & Garcia-Soriano, G. (2014). Part 1- You can run but you can't hide: Intrusive thoughts on six continents. *Journal of Obsessive-Compulsive and Related Disorders 3*(3), 269-279.

Poston, T. (2009). Knowing how to be Gettiered. *Philosophy and Phenomenological Research, 79*(3), 743-747.

Rawn, C., & Vohs, K. (2011). People use self-control to risk personal harm: An intra-interpersonal dilemma. *Personality and Social Psychology Review 15*(3), 267-289.

Redlich, F. C., & Freedman, D. X. (1966). *The Theory and Practice of Psychiatry.* New York, NY: Basic Books.

Robinson, M. D., & Clore, G. L. (2002). Beliefs, situations, and their interactions: Towards a model of emotion reporting. *Psychological Bulletin, 128*, 934-960.

Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy: "At last the Dodo said, 'Everybody has won and all must have prizes'". *American Journal of Orthopsychiatry 6*, 412-415.

Ryle, G. (1946). Knowing how and knowing that. *Proceedings of the Aristotelian Society 46*, 1-16.

Ryle, G. (1949). *The Concept of Mind.* Chicago, IL: Chicago University Press.

Scadding, L. G. (1967). Diagnosis: The clinician and the computer. *Lancet, 2,* 877-822.

Scheff, T. J. (1970). Schizophrenia as ideology. *Schizophrenia Bulletin 1*(2), 15-19.

Schnall, S., Benton, J., & Harvey, S. (2008). With a clean conscience: cleanliness reduces the severity of moral judgments. *Psychological Science*, *19*(12), 1219–1222.

Schroeder, S. A. (2013). Healthier or healthier than? *British Journal for the Philosophy of Science 64*, 131-159.

Schwartz, P. H. (2014). Reframing the disease debate and defending the biostatistical theory. *Journal of Medicine and Philosophy*, *39*(6), 572-589.

Schwitzgebel, E. (2008). The unreliability of naive introspection. *Philosophical Review 117*(2), 245-273.

Sedgwick, P. (1973). Illness: Mental and otherwise. *Hastings Center Studies 1*(3), 19-40.

Segal, Z. V., Teasdale, J. D., & Mark, J. (2004). Mindfulness-Based Cognitive Therapy: Theoretical Rationale and Empirical Status. In S. C. Hayes, V. M. Follette, & M. M. Linehan (Eds.), *Mindfulness and acceptance: Expanding the cognitive-behavioral tradition* (pp. 45–65). New York: Guilford Press.

Segal, Z. V., Teasdale, J. D., & Mark, J. (2004). Mindfulness-Based Cognitive Therapy: Theoretical Rationale and Empirical Status. In S. C. Hayes, V. M. Follette, & M. M.

Linehan (Eds.), *Mindfulness and acceptance: Expanding the cognitive-behavioral tradition* (pp. 45–65). New York, NY: Guilford Press.

Segal, Z. V., Williams, J. M. G., & Teasdale, J. D. (2012). *Mindfulness-based cognitive therapy for depression*. New York, NY: Guilford Press.

Shapiro, A., Lu, Z. L., Huang, C. B., Knight, E., & Ennis, R. (2010). Transitions between central and peripheral vision create spatial/temporal distortions: A hypothesis concerning the perceived break of the curveball. *PLoS One, 5*(10), e13296.

Skinner, B. F. (1977) Why I am not a cognitive psychologist. *Behaviorism 5*(2): 1-10.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American psychologist, 32*(9), 752.

Smith, M. L., Glass, G. V., & Miller, T. (1980). *The Benefits of Psychotherapy*. Baltimore, MD: Johns Hopkins.

Solomon, A. & Haaga, D. A. (2004). Cognitive theory and therapy in depression. In M. A. Reinecke & D. A. Clark (Eds.), *Cognitive therapy across the lifespan* (pp.12-40). Cambridge, UK: Cambridge University Press.

Sosis, R. (2009). The adaptationist-byproduct debate on the evolution of religion: Five misunderstandings of the adaptationist program. *Journal of Cognition and Culture 9*(3), 315-332.

Spitzer, R. L., & Endicott, J. (1978). Medical and mental disorder: proposed definition and criteria. In R. L. Spitzer & D. F. Klein (Eds.) *Critical Issues in Psychiatric Diagnosis.* New York, NY: Raven.

Steinpreis, R. E., Anders, K. A., & Ritzke, D. (1999). The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles, 41*, 509-528.

Stanley, J. (2011). *Know How.* Oxford: Oxford University Press.

Stanley, J. (2015). Knowledge, habit, practice, skill. *Journal of Philosophical Research, 40*(Supplement), 315-323.

Stanley, J., & Krakauer, J. W. (2013). Motor skill depends on knowledge of facts. *Frontiers in Human Neuroscience 7*, 1-11.

Stanley, J., & Williamson, T. (2001). Knowing how. *Journal of Philosophy 98*, 411-444.

Stanley, J., & Williamson, T. (2017). Skill. *Nous 51*(4), 713-726.

Teasdale, J. D. (1997). The relationship between cognition and emotion: The mind-in-place in mood disorders. In D. M. Clark & C. G. Fairburn (Eds.), *Science and practice of cognitive behaviour therapy* (pp. 67–93). New York, US: Oxford University Press.

Teasdale, J. D., & Barnard, P. J. (1993). *Affect, cognition, and change: Re-modelling depressive thought* (Vol. xv). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Teasdale, J. D., & Barnard, P. J. (1993). *Affect, cognition, and change: remodelling depressive thought.* New York, NY: Psychology Press.

Thomas, D. L., & Diener, E. (1990). Memory accuracy in the recall of emotions. *Journal of Personality and Social Psychology, 59*(2) 291-297.

Tolin, D. F. (2010). Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. *Clinical Psychology Review,30*(6), 710-720.

Van Inwagen, P. (1990). *Material Beings.* Ithica, NY: Cornell University Press.

Wakefield, J. C. (1992). The concept of mental disorder: on the boundary between biological facts and social values. *American Psychologist, 47*, 373-388.

Wakefield, J. C. (1999). Evolutionary versus prototype analysis of the concept of disorder. *Journal of Abnormal Psychology 108*(3), 374-399.

Wallis, C. (2008). Consciousness, context, and know-how. *Synthese, 160*(1), 123-153.

Wampold, B. E. (2001). Contextualizing psychotherapy as a healing practice: Culture, history, and methods. *Applied and Preventive Psychology*, *10*(2), 69–86.

Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. New Jersey: Lawrence Erlbaum Associates.

Wampold, B. E. (2007). Psychotherapy: the humanistic (and effective) treatment. *The American Psychologist*, *62*(8), 855–873.

Wampold, B. E., & Imel Z.E. (2015). *The great psychotherapy debate: the evidence for what makes psychotherapy work*. New York, NY: Routledge.

Wampold, B. E., Flückiger, C., Del Re, A. C., Yulish, N. E., Frost, N. D., Pace, T. B., Goldberg, S. B., Miller, S. D., Baardseth, T. P., Laska, K. M., & Hilsenroth, M. J. (2017). In pursuit of truth: a critical examination of meta-analyses of cognitive-behavioral therapy. *Psychotherapy Research, 27*(1), 14-32.

Wampold, B. E., Imel, Z. E., Bhati, K. S., & Johnson-Jennings, M. D. (2007). Insight as a Common Factor. In L. G. Castonguay & C. Hill (Eds.), *Insight in psychotherapy* (pp. 119–139). Washington, DC, US: American Psychological Association.

Wampold, B. E., Minami, T., Baskin, T. W., & Callen Tierney, S. (2002). A meta-(re) analysis of the effects of cognitive therapy versus "other therapies" for depression. *Journal of Affective Disorders*, *68*(2), 159–165.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies. *Psychological Bulletin, 122*(3), 203.

Wells, A. (2009). *Metacognitive therapy for anxiety and depression*. New York, NY: Guilford Press.

Weatherson, B. (2017). Intellectual skill and the Rylean regress. *The Philosophical Quarterly, 67*(267), 370-386.

Wenzel, A., Chapman, J. E., Newman, C. F., Beck, A. T., & Brown, G. K. (2006). Hypothesized mechanisms of change in cognitive therapy for borderline personality disorder. *Journal of Clinical Psychology 6*(24), 503-516.

Westbrook, D., Kennerley, H., & Kirk, J. (2011). *An introduction to Cognitive Behaviour Therapy: Skills and Applications.* SAGE Publications.

Wiggins, D. (2012). Practical knowledge: Knowing how to and knowing that. *Mind, 121*(481), 97-130.

Wilson, D. S. (2003). *Darwin's cathedral: Evolution, religion, and the nature of society.* Chicago, IL: University of Chicago Press.

Wilson, T. D. (2002). *Strangers to Ourselves: Discovering the Adaptive Unconscious* (New Ed edition.). Cambridge, Mass.: Belknap Press.

Wolitzky, D. L. & Eagle, M. N. (1992). Psychoanalytic theories of psychotherapy. In Freedheim, D., Donald R., Strupp H., Wachtel P. (Eds.), *History of psychotherapy: A century of change* (pp. 109-158). Washington, DC: American Psychological Association.

Woolfolk, R. L. (1999). Malfunction and mental illness. *The Monist, 82*(4). 658-670.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: clinicl descriptions and diagnostic guidelines.* Geneval: World Health Organization.

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2003). *Schema Therapy: A Practitioner's Guide*. New York: Guilford Press.