Defining Mental Disorder

Jerome Wakefield and His Critics

The title-level DOI for this work is:

doi:10.7551/mitpress/9949.001.0001

9 Psychiatric Disorders and the Imperfect Community: A Nominalist HDA

Peter Zachar

I. The Concept of Psychiatric Disorder: Why Wakefield Matters[1]

There has been extensive scholarly discussion about how to define psychiatric disorder (Gert and Culver 2004; Graham 2010; D. Murphy 2006; Wakefield 1992b). Developing an official definition of psychiatric disorder became important, in part, because of the disagreement in American psychiatry during the 1970s about whether homosexuality is a disorder. The *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) definition (developed by Robert Spitzer and first published in the *DSM-III* and revised in later editions) was proposed to support the exclusion of homosexuality from the class of psychiatric disorders (Bayer 1981; Zachar and Kendler 2012). It was not, however, offered to *justify* that exclusion. The primary justification for the exclusion was an empirical one, specifically, the discovery that gay male relationships are not more compulsive and short term in nature than are heterosexual relationships.

Prior to the 1970s, the conventional argument for the pathological nature of homosexuality was that such relationships lacked the depth and commitment of mature sexual relationships. Once this was shown to be false, it became evident that the mental health benefits of sexual relationships in general can accrue to homosexual relationships; that is, sex between two men or two women can have positive effects on their mental health. Being gay was not inherently distressful, nor did it necessitate social or occupational dysfunction. For these reasons, making distress or impairment definitional of a psychiatric disorder supported removing homosexuality from the classification system. In its original formulation, the *DSM* definition began as follows: in *DSM-III*, each of the mental disorders is conceptualized as a clinically significant behavioral or psychological syndrome or pattern that occurs in an individual and that is typically associated with either a painful symptom (distress) or impairment in one or more important areas of functioning (disability) (American Psychiatric Association 1980, 6).

Our working concept of psychiatric disorder, therefore, has significant cultural implications, and defining it is more than an intellectually entertaining puzzle.

Consider the following nominalist definition: *psychiatric disorder is a name for what psychiatrists treat*. A definition of this sort was once suggested by Lilienfeld and Marino (1995). In making this proposal, they were calling attention to the heterogeneity of psychiatric disorders. The definition implies that for the class of psychiatric disorders as a whole, there is no essence or set of necessary and sufficient properties that all of them share and that distinguish them from other medical disorders and from normality.

The problem with such an austere definition is that, as happened in the 1850s, a community of psychiatrists could label recurrent attempts to escape slavery a psychiatric disorder, and no one could reject that label by arguing that runaways do not "really" have a psychiatric disorder. According to the austere definition, if these psychiatrists decide to conceptualize recurrent escape attempts as a "compulsion" and to treat it, then it is a psychiatric disorder. Under the guidance of such a relativistic concept, any kind of political or cultural dissident could be labeled as disordered by a community of psychiatrists.

Szasz's (1961) opposition to psychiatry as a medical specialty is based on his claim that psychiatric disorders are disliked because they represent not legitimate diseases but social norm violations. Consider slavery again. In 1851, the American physician Samuel Cartwright proposed that slaves who evidenced a rebellious desire to run away had a psychiatric disorder that he named *drapetomania*. For those slaves who did succeed in absconding, Cartwright claimed that misery in the form of an even worse disorder called *dysaesthesia aesthiopis* (or rascality) would follow them because they were not constituted to cope with freedom. But all was not lost, he said: "With the advantages of proper medical advice, strictly followed, this troublesome practice that many negroes have of running away, can be almost entirely prevented" (Cartwright 1851/2004, 34).

According to the Szaszian view, the only difference between Cartwright's including drapetomania under the umbrella of psychiatric disorder and the modern psychiatric community's abhorrence of Cartwright's proposal is that the modern psychiatric community holds different values than did Cartwright. The concept of psychiatric disorder, says Szasz (1960/2004), is an abstract name for those problems in living that society considers deviant and deserving of remediation.

1.1 The Harmful Dysfunction Analysis

In light of the Szaszian critique, one of the purposes of a conceptual definition of psychiatric disorder is to help psychiatrists demarcate valid disorders from all other problems in living. With respect to this goal, the most philosophically influential analysis of "psychiatric disorder" is Wakefield's (1992a, 1992b, 2000, 2004) *harmful dysfunction* (HD) model. Wakefield combines the metaphysical essentialism of Kripke (1972) and early Putnam (1975) with the psychological essentialism of Medin and Ortony (1989) under the name *black-box essentialism*. According to this view, the nature of a psychiatric disorder should be subject to scientific authority just as the nature of gold is subject to scientific authority.

Humans noticed and started working with gold at least 4,000 years ago. At various points in history, there was occasional disagreement about the criteria of "real" gold, but people were generally consistent in what they took to be gold. In the twentieth century, scientists discovered that every atom of gold (defined as the metallic element between platinum and mercury on the periodic table) has seventy-nine protons in its nucleus. "The element having seventy-nine protons in the nucleus" was arguably the object of people's talk about gold from the very beginning, even though it was hidden from view—or *in the black box*. The concept of gold *indirectly referred* to the element having seventy-nine protons, but the empirical meaning of the term was not properly specified until scientists discovered atomic structures.

Ancient people also noticed and named behavioral aberrations such as melancholia and mania, although, unlike gold, the concept of "mental/psychiatric disorder" is a term of art that is linked to the medical profession. Within medicine and related professions, there is some intuitive consistency in the concept's use, but according to Wakefield, the meaning of "psychiatric disorder" can and should be clarified just as the meaning of gold was clarified. The harmful dysfunction model is proposed as such a clarification.

Wakefield agrees that psychiatric disorders represent norm violations as the Szaszians claim, but he also argued that "dysfunction" is an objectivist concept—referring to the failure of some biological or psychological mechanism to perform as it was designed to perform during evolution. In Wakefield's synthesis of objectivism and normativism,[2] the attribution of "psychiatric disorder" to a particular condition involves a judgment on the part of mental health professionals that there exists an objective psychological dysfunction that, in addition, is harmful to its bearer and deserving of treatment. Murphy (2006) dubbed this the two-stage picture.

An important aspect of Wakefield's model is the concept of *natural function*. In evolutionary theory, natural functions are adaptive capacities such as vision and temperature regulation whose contribution to a species' survival explains why the mechanism underlying those capacities were selected during evolution (Millikan 1984; Wright 1973). According to this approach, the eyes were designed through natural selection for seeing; if they cannot see (due to something like cataracts), then there is a dysfunction (i.e., a failure of their naturally selected function).

One must also understand that Wakefield's model, like many other evolutionary models in psychology, is an interactive and contextual, not a reductionist, model. For example, many natural psychological functions were selected because they are adaptive responses to social and psychological situations. According to Wakefield, the underlying biological mechanisms for intense sadness may be the same in a grief reaction and a depressive disorder, but a grief reaction is a normal selected response to bereavement, whereas a depressive disorder occurs in response to situations in which intense sadness would not have been selected.

The key problem with the harmful dysfunction model is that it offers limited empirical guidance in distinguishing disorders from nondisorders because identifying objective natural functions depends on conceptual analysis, not factual evidence. Samuel Cartwright's own argument for drapetomania was predicated on the *inability* of some slaves to accept the submissiveness that he speculated represented natural functioning for black Africans enslaved in the United States.

As argued by Richardson (2007), there is not enough information about the selection pressures that were operating during human evolution, particularly on the evolution of the brain, to support empirically based theories of natural function. Wakefield (2001) contends that careful reasoning can reveal what natural psychological functions exist, but one has to worry that reason unconstrained by evidence can be marshalled to defend many different conclusions.

For example, Horwitz and Wakefield (2012) use a conceptual analysis of what we should and should not be expected to do to identify what lies within our biologically designed, naturally selected range of behaviors. According to them, talking to family members without intense anxiety lies in this range, but handling snakes without intense anxiety does not. Only psychiatric symptoms that interfere with what we should naturally be expected to do are considered objective dysfunctions. In this analysis, the distinction between disordered and normal is being made not by discovering an objective dysfunction but by reasoning.

The HD analysis cannot, therefore, empirically do what it was proposed to do, *factually* demarcate valid psychiatric disorders from the larger class of problems-in-living. It is quite likely that no model could do so given all the different considerations that might be deemed relevant in considering something to be dysfunctional and harmful.

1.2 Essentialism versus Empiricism

Spitzer's definition of mental disorder was a listing of features, not an abstract concept such as *harmful dysfunction*. Wakefield's conceptual analysis—that something has "really" gone awry inside the person and that it is harmful to its bearer—is parsimonious and useful. It is also an important advance in our thinking following the challenges posed by the Szaszian critique. For good reasons, Wakefield's analysis has become the de facto definition of mental disorder in psychiatry.

In adopting Wakefield's concept, however, psychiatrists and psychologists have also, maybe unwittingly, adopted a de facto essentialism. This would not displease Wakefield. In a penetrating analysis of the work of the eminent psychologist and philosopher Paul Meehl, Wakefield (2004) argues that when Meehl gave up strict operationalism in favor of scientific realism and construct validity, he made a mistake in not also abandoning empiricism. As an empiricist, Meehl continued to advocate for treating scientific concepts as open. The notion of an open concept was promulgated by the

philosopher Arthur Pap (1958) in his critique of the analytic-synthetic distinction. This neo-empiricist notion, says Wakefield, is a myth.

According to Meehl (1986), *DSM* depression is an open diagnostic concept because the cluster of signs and symptoms in the *DSM* are, at best, indirect measures of an underlying pathology and its associated etiology. Clearly, Meehl construes depression in a medical model framework as a disease entity that results from an underlying pathological process, and Wakefield makes a good point that the medical model's conventional notion of a disease entity coheres with essentialism (perhaps even an essentialism that is more reductionist than Wakefield would prefer).

But Meehl also said that a psychological disorder such as depression is different from an infectious disease. It is also different from gold. For instance, key features of depression such as low positive emotionality and cognitive distortion are *conceptual interpretations* of behavior. The notion of an open concept refers to how the meaning of an abstract, dispositional concept such as "depression" is distributed (Meehl 1978). It is distributed among more observable indicators (e.g., lack of positive emotionality) and other theoretical concepts by which depression is implicitly defined (e.g., psychiatric disorder, cognitive distortion, and object loss). The meaning of an open concept cannot be defined *only* by a set of measurements (or partial definitions). Furthermore, open concepts are potentially extendable so that a new measurement *may* also become part of our definition of the concept. An open concept refers to what it is that the different operationalizations of it have in common, but it is not reducible to any of those operationalizations.[3]

The network of concepts that indirectly define an open concept such as depression is called the concept's surplus meaning. Meehl's notion of construct validation is, in part, about clarifying those surplus meanings that are of interest to us (Cronbach and Meehl 1955). Such meanings and the generalizations they allow can evolve as new facts are discovered and related concepts are modified (including our causal hypotheses). The goal is to calibrate our understanding of the concept so that it is adequate to both facts and well-supported theories (Zachar 2012), but the extendable/open nature of these concepts challenges the essentialist goal of treating them as rigid designators.

For example, the discovery that most cases of depression are precipitated by stress in the previous six months may lead us modify the concept of depression one way, and the discovery that cases of depression that lack precipitants are more treatment resistant may lead us to modify the concept (and related concepts such as psychiatric disorder) in another way. Each of these decisions could lead to elucidating different causal trajectories for "depression." If we narrow the depression construct by eliminating cases with clear precipitants, we can tell ourselves that this (and its causal story) is what we were "really" referring to all along, but that is a post hoc, even Whiggish assertion.

From a neo-empiricist standpoint, essentialism is an excessive metaphysical elaboration that is needlessly grafted onto this complicated network of "observations" and

allied concepts. Some developmental psychologists have argued that essentialism is a cognitive bias that emerges by the time we are five or six years old (Gelman 2003, 2004; Gelman and Wellman 1991). The bias is reinforced by the science curriculum in high school and college because using essentialist frameworks makes scientific concepts easier to understand. The essentialist framework thereafter becomes accepted as a scientific ideal—a model of what a real science is. As scientists gain experience in their own domains of expertise, however, they increasingly adopt nonessentialist thinking. For example, as one learns more about depression, the population of depressive symptom clusters is more likely to be seen in a nonessentialist way as: (a) the result of multiple causal trajectories, (b) with no necessary and sufficient set of causes that are identity determining, and (c) regularly overlapping with normality, anxiety disorders, obsessiveness, and psychosis.

Doubtlessly, the essentialist bias makes Wakefield's concept attractive, whereas the relativism of Lilienfeld and Marino's nominalism leaves psychiatric classification too ungrounded. If philosophical empiricists do not want to cede the ground to the essentialism (or to Meehl's putative crypto-essentialism) but avoid extreme nominalism and relativism, a variation upon Wakefield's, Meehl's, and Lilienfeld and Mariono's analyses is needed. The variation I propose is called the imperfect community model.

II. The Imperfect Community Model

2.1 The Experience of Dysfunction

In early onset Alzheimer's disease, the experience of dysfunction includes getting lost while driving in familiar places or continually forgetting recent events. Such experiences are salient examples of a *decline-in-functioning* that is developmentally unexpected and not a part of the typical course of life (Zachar 2011; Zachar and Kendler 2010). They are intrusive and unwanted failures of capacities that used to be there. Declines-in-functioning should also occur across multiple contexts—they travel with the person.

There are three important differences between this minimalist notion of dysfunction-as-decline and Wakefield's more ontologically elaborate concept of objective natural dysfunction. First, its objectivity does not depend upon speculation about natural functions. Rather, declines-in-functioning are objective in two different senses: (a) they are often intersubjectively confirmable, and (b) denying that they have occurred, although common, is a distortion. Eventually, people who are open to the evidence are compelled to accept that an important change has occurred, no matter what they may prefer to be the case.

Second, this minimalist notion is also normative. Wakefield separates dysfunction from harm, but these concepts are tightly integrated in noncontroversial examples of disorder such as Alzheimer's disease. The affected person experiences declines that *should not* have happened. They are *unwanted* declines. They represent something being

broken. They are capacity *failures*. One can understand the attractiveness of stipulating that dysfunctions are out there and those that are harmful are disorders, but that does not seem consistent with how we come to identify dysfunctions.

What does making dysfunction both objective and normative do to the two-stage picture? The second stage is the attribution of disorder. I join Wakefield in using the term "disorder" as a general concept that encompasses diseases (e.g., tuberculosis), injuries (e.g., broken bones), vulnerability conditions (e.g., hypertension), and numerous painful states such as tension headaches that can be associated with "the sick role." They actively or potentially interfere with functioning and are reasonable targets for treatment. To name something a disorder, practically speaking, is to say that it a potential target for treatment. As we will see shortly, however, what unites psychiatric disorders is not only a belief that they are deserving of treatment but also the kinds of symptoms that characterize them.

Third, if we examine the set of things currently called psychiatric disorders, it is clear that a decline-in-functioning is not an essence. It is neither necessary for the attribution of psychiatric disorder nor sufficient. Many cases of intellectual disability (and other neurodevelopmental disabilities) do not manifest as declines in functioning.

2.2 The Domain of Psychiatry

Berrios (1996) reports that at the beginning of the nineteenth century, the main categories of psychiatric disorder were melancholia, mania, phrenitis, delirium, paranoia, lethargy, carus, and dementia. All these conditions are unambiguous examples of declines-in-functioning, and those conditions causing the greatest degree of impairment would presently be called psychotic conditions. People who become psychotic represent a psychiatrically vulnerable population. If one examines these cases over time, in addition to the florid psychotic symptoms such as hallucinations and delusions, one sees panic, obsessiveness, hypochondriasis, mood instability, impulsivity, and lack of empathy—in fact, much of our extant psychiatric symptom space.

In their studies of folk conceptions of mental disorder, Haslam (2005) and his colleagues propose that behavior that is unexpected, hard to understand or explain, and owned by the person (as opposed to compelled by an outside agent) is seen as pathological in all societies (Giosan et al. 2001; Haslam et al. 2007). "Pathologizing" refers to a sense that something is not right with the person—an inference that is easiest to make if there is a change/decline from a previous level of functioning.

Social constructionists sometimes suggest that in other cultures, people who become psychotic are valued and given meaningful roles like that of the shaman (Silverman 1967). This is better considered a myth (Boyer 2011; Haslam et al. 2007). For example, my colleague Jim Phillips spends part of his year working in Ayacucho, Peru—a rural city in the Andean Mountains. He claims that psychosis in the Ayacucho looks much like psychosis in the United States and that no one is inclined to give it a positive spin.

Similar observations about Eskimos in Alaska and the Yorubas in rural Nigeria were offered independently by the anthropologist Jane Murphy (1976).

In the domain of psychiatric disorders, psychotic states are *exemplars* in Medin's (1989) sense of the term. The larger domain was assembled, initially, in reference to them. Historically, psychiatry as a field developed in the nineteenth century when the exemplary psychotic disorders managed by the doctors who worked in mental asylums (called alienists) were expanded upon by the addition of the functional disorders of the neurologists that occupied some of the same symptom space as psychosis. To the extent that these combined symptom clusters explain why the discipline of psychiatry first appeared, their inclusion in the domain cannot be simply relativized to the idiosyncratic preferences of small communities of psychiatrists.

This development is usually discussed with respect to how the psychological approach associated with Freud came to replace the organic model of the alienists, but for our purposes, the important thing was the expansion of the symptom domain to cover the kinds of problems encountered in both the inpatient settings of the alienists and the outpatient settings of the neurologists. The link between the two settings was the group of premorbid and residual symptoms that resided in the penumbra of the psychoses.

Particularly in the United States, there was a major expansion of psychiatry into the outpatient population after World War II—in the 1950s and thereafter. The establishment of the clinical and counseling psychology specialties in the Veterans Administration hospitals and on college campuses at this time was also important. It is crucial to point out that this expansion cannot be simply attributed to the activity of mental health professionals because people with psychiatric symptoms actively sought out both treatments and diagnoses. In many respects, the expansion in the number of diagnostic constructs in the *DSM-III* was a belated recognition of this new reality.

The result of this mélange of functional disorders is an imperfect community—meaning that there is no set of properties that all psychiatric disorders share and that distinguish them from nondisorders. The "conditions" that were added to the psychiatric domain overlapped with the psychosis cluster in a variety of ways. These include but are not limited to the following:

Decline-in-functioning and other statistically abnormal developmental trajectories

The presence of reality distortion

Suicidal ideation

Confusion and other cognitive difficulties

Intrusive thoughts

Difficult to control impulses and compulsions

Agitation, anger, and excitement

Excessive anxiety and fear

Emptiness and anhedonia

Somatic preoccupations

Seeming more amenable to the skill set of psychiatry than other medical specialties

Interestingly, anthropologist Jane Murphy also indicated that among both the Eskimos and the Yorubas, a greater number of people suffer from the kinds of symptoms that psychiatrists would call depression-anxiety. These symptoms are considered different from being "crazy." Although Eskimos and the Yorubas do not lump this cluster of symptoms under a single name like "depression" or "neurosis," both groups consider them problems that are under the purview of the shaman/healer.

The "imperfect" part of the community of psychiatric disorders has been eloquently described by Allen Frances (2013):

> Some mental disorders describe short-term states, others life-long personality; some reflect inner misery, others bad behavior; some represent problems rarely or never seen in normals, others are just slight accentuations of the everyday; some reflect too little self-control, others too much; some are intrinsic to the person, others are culturally determined; some begin early in infancy, others emerge only late in life; … some are clearly defined, others not; and there are complex permutations of all of these possible differences. (17)

Although imperfect, the notion of *a community* suggests that the collection is not simply random or arbitrary. The various symptoms and symptom clusters are included as members for reasons. The domain of psychiatry should not be limited primarily to psychosis but also include what was added in the merging of the disorders of the alienists and the neurologists, of inpatient and outpatient, and of decline, distress, and disability into the imperfect community of psychiatric disorders.

III. The Causal Network Approach

Rachel Cooper (2005) claims that the concept of psychiatric disorder refers to unwanted psychological-behavioral conditions just as the concept of weed refers to unwanted plants. Cooper also notes that although "weed" is a heterogeneous category, the same cannot be said for particular kinds of weeds. For example, a dandelion is a kind of weed. Dandelions also have shared underlying properties, and generalizations about them can be made. She suggests that the same can be said for psychiatric disorders such as major depressive disorder and schizophrenia.

Wakefield argues that particular disorders such as major depressive disorder and schizophrenia, if valid, are the expressions of underlying psychopathological structures that represent design failures. According to Wakefield (2004), talk about these disorders directly refers to their symptomatic manifestations but indirectly refers to their underlying mechanisms. The mechanisms represent what the disorders really are.

3.1 Latent Variables versus Causal Networks

In psychometrics, the hidden patterns that causally produce observable symptoms are called *latent variables*.[4] When depicted visually, latent variables are represented as circles with causal arrows pointing at squares, which represent observed variables (see figure 9.1).

In clinical psychology, latent variables are considered to represent the psychopathological reality behind the appearances. They are causally important, the same from case to case, and make disorders what they are (identity determining). As a result, they correspond to the philosopher's notion of real essences. Although this essentialist model still remains largely promissory, it continues to hold sway—and understandably so.

Essentialism, however, is not scientifically necessary. A group of psychologists associated with the psychological methods program at the University of Amsterdam, including Han van der Maas, Denny Borsboom, and Angélique Cramer, argues that latent variables do not have to be interpreted as referring to real essences. Consider the latent variable called *psychometric g*. This variable is a mathematical index of the positive correlations that exist between different measures of cognitive ability. It is often conceptualized as a psychological ability called "general intelligence," which refers to what it is that all cognitive abilities share. In the realist interpretation of latent variables, the positive correlations between the abilities exist because they are all the outcomes of a shared causal entity represented by *g*.

According to van der Maas et al. (2006), an alternative to a causally potent latent variable (or common cause) model is a model in which cognitive abilities are in direct causal relationships with each other. For example, being able to process information quickly might have positive effects on working memory. Cognitive abilities can enter into mutual interactions in a variety of ways. Some people may naturally have high abilities across the board, whereas others are gifted in one or two areas—such as

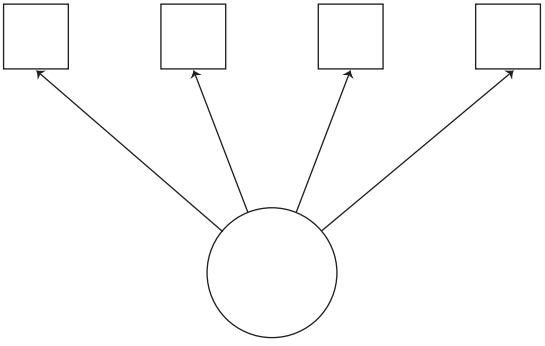

Figure 9.1
Visual representation of a latent variable model.

processing speed and attention capacity—but these skills permeate through the ability network and raise scores on tests of general intelligence. For example, in neuropsychology settings, temporary problems with focused attention just after a brain injury will depress scores on other cognitive abilities. An assessment of lasting deficits cannot occur until attention improves.

Van der Maas and his colleagues simulated data sets that were consistent with both the common cause scenario and the mutual interaction scenario and discovered that the latent variable model "fit" both of them. This means that psychometric *g* will mathematically appear if the positive correlations between the variables are the result of direct causal relationships rather than the result of an underlying common cause. Both scenarios can be analyzed to produce the shared correlations that are lumped together as psychometric *g*.

One implication of this research is that the psychological concept of general intelligence as the ability to perform well across multiple cognitive domains is an empirically supported phenomenon, but it need not be the result of an underlying causal entity called *g*. Another implication is that the relevant causal structure from which a latent variable emerges does not have to be a universal (or the same from case to case).

Likewise, in psychiatry and psychology, latent variables are interpreted realistically—meaning that the cluster of symptoms that constitute depression is considered correlated because they are manifestations of a shared underlying psychopathological process (Borsboom et al. 2003; Kendler et al. 2011). Furthermore, the more reliable the symptomatic criteria, the better they are supposed to be at estimating a person's *true score* on the underlying variable. In contrast, for causal networks, the symptoms hold together because they are in direct, possibly causal, relationships with each other (Borsboom 2008). For example, rather than both sleep problems and fatigue being manifestations of a single underlying cause called "depression," sleep problems (SP) likely directly influence the level of fatigue (F). In addition, such factors as depressed mood (DM) and loss of interest (LI) are central symptoms, meaning they enter into a high number of mutual relationships with other symptoms in the network. As a result of these connections, when central symptoms are activated, it is more likely that other symptoms will follow. A pathological state of depression would represent the emergence of feedback loops between symptoms that become self-sustaining (see figure 9.2).

Abandoning the realist interpretation of latent variables in favor of symptom networks, however, does not make depression a theoretical fiction. Depression is instead understood as the activation of a network within the larger symptom space of psychiatric disorders. According to Borsboom (2008), requiring five out of nine symptoms for a diagnosis does not indicate the presence of an underlying entity called depression. Instead, it indicates the extent to which the symptom network (named depression) has been entered.

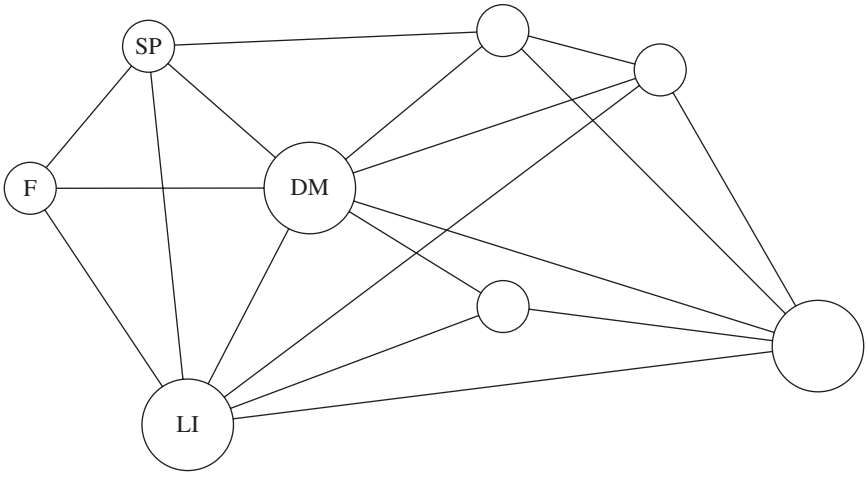

Figure 9.2
A causal network model for major depressive disorder.

3.2 Comorbidity

The network model also offers a new understanding of comorbidity. In traditional medicine, comorbidity is defined as the simultaneous occurrence of two causally independent diseases such as liver cancer and heart disease (Feinstein 1970). Presumably, the presence of one disease has consequences for the development and treatment of the second. The problem in psychiatry is that such co-occurrences tend not to be independent. *Psychiatric comorbidity* refers to complicated, multisymptomatic cases that tend to occur in vulnerable populations (Klein and Riso 1993; Neale and Kendler 1995; Zachar 2009).

According to one very influential latent variable model, the high rate of comorbidity between a depressive episode and generalized anxiety disorder (GAD) is explained with reference to a common vulnerability factor—the personality trait of neuroticism (Clark 2005; Kahn et al. 2005). In contrast, the causal network approach conceptualizes comorbidity in terms of the relationships between symptoms within the larger network of psychiatric symptomatology (Borsboom et al. 2011).

Using data from the National Comorbidity Survey Replication, Cramer et al. (2010) mapped reciprocal relationships between the symptoms in both the depression and GAD clusters. They discovered that some symptoms have connections to symptoms in both networks. They labeled these bridge symptoms.

In depression and GAD, the bridges connecting the two networks include sleep problems (SP), fatigue (F), concentration problems (CP), and irritability (I). For example, the central symptom of depressed mood (DM) has multiple relationships with

other depression symptoms. It is also connected to several bridge symptoms and, through them, to symptoms in the GAD network. DM even does double duty as a bridge symptom by being directly connected to chronic anxiety (CA), which is itself a central symptom in the GAD network. In this model, comorbidity is the result of a spreading activation process. In more vulnerable persons, once activated, a symptom network stays activated via feedback loops (see figure 9.3).

In traditional medical classification, good diagnostic criteria are both sensitive and specific indicators of a disorder. For this reason, a symptom such as irritability is not an ideal criterion for depression because it is sensitive to depression but not specific to depression. Highly anxious people are also irritable. Within the network perspective, however, rather than being ignored because they are not specific to a single disorder, overlapping symptoms contribute to our understanding of how complicated cases might develop. When bridge symptoms are ignored, the gaps between clusters look larger (or more "real") than they are.

An important implication of the symptom network model is that diagnosticians should be attending not only to the diagnostic categories for which a patient meets criteria but also to the number of symptoms activated. If two separate individuals each meet four criteria for a major depressive episode, neither would be diagnosed. But if the first person meets criteria composed of symptoms that are also bridges to another network such as the anxiety disorders network, he might be experiencing considerably more social and occupational dysfunction than the second person who meets criteria for fewer bridge symptoms. Not all subthreshold conditions are the same. Consider

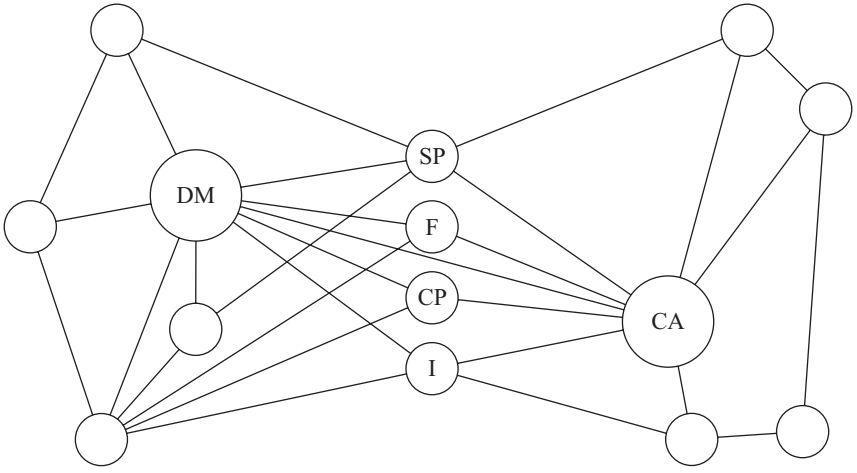

Figure 9.3
A causal network mode for the comorbidity of depression and generalized anxiety disorder.

this statement from the *DSM-5*: "the boundaries between many disorder 'categories' are more fluid over the life course than the *DSM-IV* recognized, and many symptoms assigned to a single disorder may occur, at varying levels of severity, in many other disorders" (American Psychiatric Association 2013, 5).

Given the empirically demonstrated patterns of comorbidity, it would not be unexpected for a person diagnosed with depression to experience anxiety-related symptoms that are not typically listed as falling under the depression concept. From an essentialist standpoint, these extra symptoms are accidental rather than essential properties of a patient's depression. From a symptom network perspective, these symptoms may be an integral part of the symptom cluster for that person.

It should be noted that the network model does not eliminate underlying causal structures. For example, a symptom such as a sleep disturbance can be understood with respect to a multiplicity of underlying mechanisms at many levels of analysis (genetic, physiological, anatomical, etc.). In addition, researchers could also investigate whether the causal relationship between sleep problems and concentration problems involves relations between two sets of underlying mechanisms, that is, the presence of direct causal relations between endophenotypes.

A symptom such as "sleep problem" is also a conceptual abstraction that summarizes a variety of symptoms. Particular kinds of sleep problems (early awakening, difficulty falling asleep, etc.) are themselves the result of underlying causal mechanisms. If the imperfect community is a swarm made up of points that represent a cluster of symptoms, for multifaceted symptoms such as "sleep problems," we can expect that a plurality of underlying nested mechanisms is present. In such a multilevel "bushy" network, some of the basic insights of essentialism such as the importance of underlying causal properties are preserved, but the conventional essentialist framework in which these properties are seen as identity-determining universals is abandoned.

IV. Identifying Disorders in the Imperfect Community

The disorders of psychiatry are the result of a gradual addition of variations on the symptom clusters of the alienists and, after psychodynamic theories made their mark, variations on the neurotic clusters as well. What we are left with is a large symptom space that can be organized in multiple ways. The *DSM* and the *International Classification of Diseases* (*ICD*) are two ways of organizing the symptom space, but because of the way the domain was built (by the addition of variants on variants), no single organization can model all of the overlapping relationships. It is important to establish standards of adequacy and work to improve the classification system, but that goal does not entail the discovery of a classification that is uniquely privileged in nature.

What does such a model do to resolve the problem of defining psychiatric disorders arbitrarily as "what psychiatrists decide to treat?" Let us consider depression. According

to the essentialist model, valid depressive disorders are caused by dysfunctional mood-regulating mechanisms. From an empiricist standpoint, referring to objective dysfunctions hidden in the black box introduces an unnecessary metaphysical elaboration that distorts the actual basis for the distinction.

Jerry Wakefield has taught us that a careful conceptual analysis can help constrain what psychiatrists and other mental health professionals treat. However, rather than making the distinction using an inferred essentialist criterion such as objective dysfunction, it is more commonly made using a polythetic criterion set (i.e., a collection of conceptual elaborations). As more of these conceptual criteria are met, the more it makes sense to start thinking of a symptom cluster as disordered. Rather than being absolutely present or absent, disorders are a matter of degree.

Considerations that are relevant in making the depressive disorder attribution include (a) the extent to which the person has entered a psychiatric symptom network. The most important criterion is the presence of a decline in functioning, although it is not a necessary criterion. Sometimes symptoms are related to impaired functioning only, not to decline. Also, (b) those symptom networks that are locked in rather than transient and flexible are also more disorder-like. Additionally, (c) more severe symptoms and more complex symptom networks support the disorder attribution. For distressing psychological symptoms such as anhedonia, (d) if there are no compensatory factors that allow the person to continue to function (and flourish), then a disorder attribution is more warranted. It is also important not to limit assessment to a single slice of time because (e) a past history of symptoms and a family history of symptoms alter the base rates and make the disorder attribution more plausible. In these cases, the appearance of milder symptoms might signal a risk for more impaired functioning in the future. In addition, as Horwitz and Wakefield (2007) persuasively argue, the attribution of disorder to a cluster of depression symptoms is more warranted when the depressive symptoms appear out of the blue—for no apparent reason—or if there is a precipitant, the response is excessive and not in proportion to the trigger.

Acknowledgments

Andrea Solomon, Denny Borsboom, Bob Krueger, Ken Kendler, Katie Tabb, and Steve Lobello provided helpful commentary on an early version of this chapter.

Notes

1. Part of this chapter draws on material previously published in Zachar, P. (2014). *A Metaphysics of Psychopathology*. MIT Press. Reprinted with permission here.

2. Both Boorse and Fulford use the term "illness" to describe the confluence of fact (an underlying pathology) and value (being *bad* for its bearer). See Boorse, C. (1975). On the distinction

between disease and illness. *Philosophy and Public Affairs* 5, 49–46, and Fulford, K. W. M. (1989). *Moral Theory and Medical Practice*. Cambridge University Press.

3. Since this chapter was written in 2013, I have thought more about open concepts. See Zachar, P., E. T. Turkheimer, and K. S. Schaffner (2020). Defining and redefining phenotypes: operational definitions as open concepts. In *The Cambridge Handbook of Research Methods in Clinical Psychology*, A. G. C. Wright and M. N. Hallquist (eds.), 5–17. Cambridge University Press.

4. For neo-empiricists, signs and symptoms index a latent variable, and the latent variable is an index of a causal trajectory. A latent variable cannot be eliminated in favor of what indexes it or what it indexes.

References

American Psychiatric Association. 1980. *Diagnostic and Statistical Manual of Mental Disorders*. 3rd ed. American Psychiatric Association.

American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. American Psychiatric Publishing.

Bayer, R. 1981. *Homosexuality and American Psychiatry: The Politics of Diagnosis*. Basic Books.

Berrios, G. E. 1996. *The History of Mental Symptoms*. Cambridge University Press.

Borsboom, D. 2008. Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology* 64(9): 1089–1108.

Borsboom, D., A. O. J. Cramer, V. D. Schmittmann, S. Epskamp, and W. Lourens. 2011. The small world of psychopathology. *PLoS ONE* 6(11): e27407.

Borsboom, D. G., G. J. Mellenbergh, and J. van Heerden. 2003. The theoretical status of latent variables. *Psychological Review* 110(2): 203–219.

Boyer, P. 2011. Intuitive expectations and the detection of mental disorder: A cognitive background to folk-psychiatries. *Philosophical Psychology* 24(1): 95–118.

Cartwright, S. A. 1851/2004. Diseases and physical peculiarities of the negro race. In *Health, Disease, and Illness,* A. L. Caplan, J. J. McCartney, and D. A. Sisti (eds.), 28–39. Georgetown University Press.

Clark, L. A. 2005. Temperament as a unifying basis for personality and psychopathology. *Journal of Abnormal Psychology* 114(4): 505–521.

Cooper, R. 2005. *Classifying Madness: A Philosophical Examination of the Diagnostic and Statistical Manual of Mental Disorders*. Springer.

Cramer, A. O. J., L. Waldrop, H. L. J. van der Mass, and D. Borsboom. 2010. Comorbidity: A network perspective. *Behavioral and Brain Sciences* 33(2–3): 137–150.

Cronbach, L. J., and P. E. Meehl. 1955. Construct validity in psychological tests. *Psychological Bulletin* 52(4): 281–302.

Feinstein, A. R. 1970. The pre-therapeutic classification of co-morbidity in chronic disease. *Journal of Chronic Diseases* 23(7): 455–468.

Frances, A. 2013. *Saving Normal*. William Morrow.

Gelman, S. A. 2003. *The Essential Child*. Oxford University Press.

Gelman, S. A. 2004. Psychological essentialism in children. *Trends in Cognitive Science* 8(9): 404–409.

Gelman, S. A., and H. M. Wellman 1991. Insides and essences: Early understandings of the non-obvious. *Cognition* 38: 213–244.

Gert, B., and C. M. Culver. 2004. Defining mental disorder. In *The Philosophy of Psychiatry: A Companion,* J. Radden (ed.), 415–425. Oxford University Press.

Giosan, C., V. Glovsky, and N. Haslam. 2001. The lay concept of 'mental disorder': A cross-cultural study. *Transcultural Psychology* 38(3): 317–322.

Graham, G. 2010. *The Disordered Mind*. Routledge.

Haslam, N. 2005. Dimensions of folk psychiatry. *Review of General Psychology* 9(1): 35–47.

Haslam, N., L. Ban, and L. Kaufmann. 2007. Lay conceptions of mental disorder: The folk psychiatry model. *Australian Psychologist* 42(2): 129–137.

Horwitz, A. V., and J. C. Wakefield. 2007. *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. Oxford University Press.

Horwitz, A. V., and J. C. Wakefield. 2012. *All We Have to Fear: Psychiatry's Transformation of Natural Anxieties into Mental Disorders*. Oxford University Press.

Kahn, A. A., K. C. Jacobson, C. O. Gardner, C. A. Prescott, and K. S. Kendler. 2005. Personality and comorbidity of common psychiatric disorders. *British Journal of Psychiatry* 186: 190–196.

Kendler, K. S., P. Zachar, and C. Craver. 2011. What kinds of things are psychiatric disorders. *Psychological Medicine* 41(6): 1143–1150.

Klein, D. N., and L. P. Riso. 1993. Psychiatric disorders: Problems of comorbidity. In *Basic Issues in Psychopathology,* C. G. Costello (ed.), 19–66. Guilford.

Kripke, S. 1972. *Naming and Necessity*. Reigel.

Lilienfeld, S. O., and L. Marino. 1995. Mental disorder as a Roschian concept: A critique of Wakefield's 'harmful dysfunction' analysis. *Journal of Abnormal Psychology* 104(3): 411–420.

Medin, D. L. 1989. Concepts and conceptual structure. *American Psychologist* 44(12): 1469–1481.

Medin, D. L., and A. Ortony. 1989. Psychological essentialism. In *Similarity and Analogical Reasoning,* S. Vosniadou and A. Ortony (eds.), 179–195. Cambridge University Press.

Meehl, P. E. 1978. Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology* 46(4): 806–834.

Meehl, P. E. 1986. Diagnostic taxa as open concepts: Metatheoretical and statistical questions about reliability and construct validity in the grand strategy of nosological revision. In *Contemporary Directions in Psychopathology: Toward the DSM-IV,* T. Millon and G. L. Klerman (eds.), 215–231. Guilford.

Millikan, R. G. 1984. *Language, Thought, and Other Biological Categories*. MIT Press.

Murphy, D. 2006. *Psychiatry in the Scientific Image*. MIT Press.

Murphy, J. M. 1976. Psychiatric labeling in cross-cultural perspective. *Science* 191(4231): 1019–1028.

Neale, M. C., and K. S. Kendler. 1995. Models of comorbidity for multifactorial disorders. *American Journal of Human Genetics* 57(4): 935–953.

Pap, A. 1958. *Semantics and Necessary Truth*. Yale University Press.

Putnam, H. 1975. *Mind, Language and Reality: Philosophical Papers*. Vol. 2. Cambridge University Press.

Richardson, R. C. 2007. *Evolutionary Psychology as Maladapted Psychology*. MIT Press.

Silverman, J. 1967. Shamans and acute schizophrenia. *American Anthropologist* 69(1): 21–31.

Szasz, T. 1960/2004. The myth of mental illness. In *Health, Disease, and Illness,* C. A. L. Caplan, J. J. McCartney, and D. A. Sisti (eds.), 43–50. Georgetown University Press.

Szasz, T. S. 1961. *The Myth of Mental Illness*. Harper & Row.

van der Maas, H. L. J., C. V. Dolan, R. P. Grasman, J. M. Witchers, H. M. Huizenga, and M. E. Raijmakers. 2006. A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review* 113(4): 842–861.

Wakefield, J. C. 1992a. The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist* 47(3): 373–388.

Wakefield, J. C. 1992b. Disorder as harmful dysfunction: A conceptual critique of *DSM-III-R*'s definition of mental disorder. *Psychological Review* 99(2): 232–247.

Wakefield, J. C. 2000. Aristotle as sociobiologist: The 'function of a human being' argument, black box essentialism, and the concept of mental disorder. *Philosophy, Psychiatry, and Psychology* 7(1): 17–44.

Wakefield, J. C. 2001. Spandrels, vestigial organs, and such. *Philosophy, Psychiatry, and Psychology* 7(4): 253–269.

Wakefield, J. C. 2004. The myth of open concepts: Meehl's analysis of construct meaning versus black box essentialism. *Applied & Preventive Psychology* 11: 77–82.

Wright L. 1973. Function. *Philosophical Review* 82(2): 139–168.

Zachar, P. 2009. Psychiatric comorbidity: More than a Kuhnian anomaly. *Philosophy, Psychiatry, and Psychology* 16(1): 13–22.

Zachar, P. 2011. The clinical nature of personality disorders: Answering the neo-Szazian critique. *Philosophy, Psychiatry, and Psychology* 18(3): 191–202.

Zachar, P. 2012. Progress and the calibration of scientific constructs: The role of comparative validity. In *Philosophical Issues in Psychiatry II: Nosology-Definition of Illness, History, Validity, and Prospects,* K. S. Kendler and J. Parnas (eds.), 21–34. Oxford University Press.

Zachar, P., and K. S. Kendler. 2010. Philosophical issues in the classification of psychopathology. In *Contemporary Directions in Psychopathology,* T. Millon, R. F. Krueger, and E. Simonsen (eds.), 126–148. Guilford.

Zachar, P., and K. S. Kendler. 2012. The removal of Pluto from the class of planets and homosexuality from the class of psychiatric disorders: A comparison. *Philosophy, Ethics, and Humanities in Medicine* 7: 4.