Psychiatric diagnosis is fundamental to the understanding of mental illness. Without it, the study, assessment, and treatment of psychopathology would be in disarray. In this chapter, we examine (a) the raison d’être underlying psychiatric diagnosis, (b) widespread misconceptions regarding psychiatric diagnosis, (c) the present system of psychiatric diagnosis and its strengths and weaknesses, and (d) fruitful directions for improving this system.

A myriad of forms of abnormality are housed under the exceedingly broad umbrella of mental disorders. Indeed, the current psychiatric classification system contains well over 300 diagnoses (American Psychiatric Association [APA], 2013). The enormous heterogeneity of psychopathology makes a formal system of organization imperative. Just as in the biological sciences, where Linnaeus’ hierarchical taxonomy categorizes fauna and flora, and in chemistry, where Mendeleev’s periodic table orders the elements, a psychiatric classification system helps to organize the bewildering subforms of abnormality. Such a system, if effective, permits us to parse the variegated universe of psychological disorders into more homogeneous, and ideally more clinically meaningful, categories.

From the practitioner’s initial inchoate impression that a patient’s behavior is aberrant to later and better-elaborated case conceptualization, diagnosis plays an integral role in the clinical process. Indeed, the essential reason for initiating assessment and treatment is often the observer’s sense that “something is just not quite right” about the person. Meehl (1973) commented that the mental health professional’s core task is to answer the question: “What does this person have, or what befell him, that makes him different from those who have not developed clinical psychopathology?” (p. 248). Therein lies the basis for psychiatric diagnosis.

General Terminological Issues

Before proceeding, a bit of terminology is in order. It is crucial at the outset to distinguish two frequently confused terms: classification and diagnosis. A system of classification is an overarching taxonomy of mental illness, whereas diagnosis is the act of placing an individual, based on a constellation of signs (observable indicators, like crying in a
depressed patient), symptoms (subjective indicators, like feelings of guilt in a depressed patient), or both, into a category within that taxonomy. Classification is a prerequisite for diagnosis.

Another key set of terminological issues concerns the distinctions among syndrome, disorder, and disease. As Kazdin (1983) observed, we can differentiate among these three concepts based on our levels of understanding of their pathology—the underlying physiological changes that may accompany the condition—and etiology, that is, causation (Gough, 1971; Lilienfeld, Waldman, & Israel, 1994).

At the lowest rung of the hierarchy of understanding there are syndromes, which are typically constellations of signs and symptoms that co-occur across individuals (syndrome means “running together” in Greek). In syndromes, neither pathology nor etiology is well understood, nor is the syndrome’s causal relation to other conditions established. Antisocial personality disorder is a relatively clear example of a syndrome because its signs (e.g., the use of an alias) and symptoms (e.g., lack of remorse) tend to covary across individuals. Nevertheless, its pathology and etiology are largely unknown, and its causal relation to other conditions is poorly understood (Lykken, 1995). In contrast, some authors (e.g., Lilienfeld, 2013; but see Lynam & Miller, 2012) argue that psychopathic personality (psychopathy) may not be a classical syndrome. These researchers contend that psychopathy is instead a configuration of several largely independent constructs, such as boldness, coldness, and disinhibition, that come together in an interpersonally malignant fashion (Patrick, Fowles, & Krueger, 2009; see also Vitale & Newman, Chapter 16, this book).

In other cases, syndromes may also constitute groupings of signs and symptoms that exhibit minimal covariation across individuals but that point to an underlying etiology (Lilienfeld et al., 1994). For example, Gerstmann’s syndrome in neurology (Benton, 1992) is marked by four major features: agraphia (inability to write), acalculia (inability to perform mental computation), finger agnosia (inability to differentiate among fingers on the hand), and left-right disorientation. Although these indicators are negligibly correlated across individuals in the general population, they co-occur dependably following certain instances of parietal lobe damage.

At the second rung of the hierarchy of understanding there are disorders, which are syndromes that cannot be readily explained by other conditions. For example, in the present diagnostic system, obsessive-compulsive disorder (OCD) can be diagnosed only if its symptoms (e.g., recurrent fears of contamination) and signs (e.g., recurrent hand washing) cannot be accounted for by a specific phobia (e.g., irrational fear of dirt). Once we rule out other potential causes of OCD symptoms, such as specific phobia, anorexia nervosa, and trichotillomania (compulsive hair pulling) we can be reasonably certain that an individual exhibiting marked obsessions, compulsions, or both, suffers from a well-defined disorder (APA, 2000, p. 463).

At the third and highest rung of the hierarchy of understanding there are diseases, which are disorders in which pathology and etiology are reasonably well understood (Kazdin, 1983; McHugh & Slavney, 1998). Sickle-cell anemia is a prototypical disease because its pathology (crescent-shaped erythrocytes containing hemoglobin S) and etiology (two autosomal recessive alleles) have been conclusively identified (Sutton, 1980). For other conditions that approach the status of bona fide diseases, such as Alzheimer’s disease, the primary pathology (senile plaques, neurofibrillary tangles, and granulovacular degeneration) has been identified, while their etiology is evolving but incomplete (Selkoe, 1992).
With the possible exception of Alzheimer’s disease and a handful of other organic conditions, the diagnoses in our present system of psychiatric classifications are almost exclusively syndromes or, in rare cases, disorders (Kendell & Jablensky, 2003). This fact is a sobering reminder that the pathology in most cases of psychopathology is largely unknown, and their etiology is poorly understood. Therefore, although we genuflect to hallowed tradition in this chapter by referring to the major entities within the current psychiatric classification system as mental “disorders,” readers should bear in mind that few are disorders in the strict sense of the term.

Functions of Psychiatric Diagnosis

Diagnosis serves three principal functions for practitioners and researchers alike. We discuss each in turn.

**Diagnosis as Communication**

Diagnosis furnishes a convenient vehicle for communication about an individual’s condition. It allows professionals to be reasonably confident that when they use a diagnosis (such as dysthymic disorder) to describe a patient, other professionals will recognize it as referring to the same condition. Moreover, a diagnosis (such as borderline personality disorder) distills relevant information, such as frantic efforts to avoid abandonment and chronic feelings of emptiness, in a shorthand form that aids in other professionals’ understanding of a case. Blashfield and Burgess (2007) described this role as “information retrieval.” Just as botanists use the name of a species to summarize distinctive features of a specific plant, psychologists and psychiatrists rely on a diagnosis to summarize distinctive features of a specific mental disorder (Blashfield & Burgess, 2007). Diagnoses succinctly convey important information about a patient to clinicians, investigators, family members, managed care organizations, and others.

**Establishing Linkages to Other Diagnoses**

Psychiatric diagnoses are organized within the overarching nosological structure of other diagnoses. Nosology is the branch of science that deals with the systematic classification of diseases. Within this system, most diagnostic categories are arranged in relation to other conditions; the nearer in the network two conditions are, the more closely related they ostensibly are as disorders. For example, social anxiety disorder (social phobia) and specific phobia are both classified as anxiety disorders in the *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*; APA, 2013), and are presumably more closely linked etiologically than are social anxiety disorder and narcissistic personality disorder, the latter of which is classified as a personality disorder in *DSM-5*. Thus, diagnoses help to locate the patient’s presenting problems within the context of both more and less related diagnostic categories.

**Provision of Surplus Information**

Perhaps most important, a diagnosis helps us to learn new things; it affords us surplus information that we did not have previously. Among other things, a diagnosis allows us
to generate predictions regarding case trajectory. As Goodwin and Guze (1996) noted, perhaps hyperbolically, “diagnosis is prognosis” (Kendler, 1980). The diagnostic label of bipolar I disorder describes a distinctive constellation of indicators (e.g., one or more manic or mixed episodes) that discriminates the course, rate of recovery, and treatment response from such related conditions as major depression and bipolar II disorder, the latter of which is marked by one or more episodes of hypomania and disabling depression. But a valid diagnosis does considerably more than predict prognosis. Robins and Guze’s (1970) landmark article delineated formal criteria for ascertaining whether a diagnosis is valid. Validity refers to the extent to which a diagnosis measures what it purports to measure. More colloquially, validity is truth in advertising: A valid diagnosis is true to its name in that it correlates in expected directions with external criteria. Specifically, Robins and Guze outlined four requirements for the validity of psychiatric diagnoses. According to them, a valid diagnosis offers information regarding:

1. Clinical description, including symptomatology, demographics, precipitants, and differences from seemingly related disorders. The last-named task of distinguishing a diagnosis from similar diagnoses is called differential diagnosis.
2. Laboratory research, including data from psychological, biological, and laboratory tests.
3. Natural history, including course and outcome.
4. Family studies, especially studies examining the prevalence of a disorder in the first-degree relatives of probands—that is, individuals identified as having the diagnosis in question.

As a further desideratum, some authors have suggested that a valid diagnosis should ideally be able to predict the individual’s response to treatment (Waldman, Lilienfeld, & Lahey, 1995). Nevertheless, this criterion should probably not be mandatory given that the treatment of a condition bears no necessary implications for its etiology. For example, although both schizophrenia and nausea induced by food poisoning generally respond to psychopharmacological agents that block the action of the neurotransmitter dopamine, these two conditions spring from entirely distinct causal mechanisms. Some authors (e.g., Ross & Pam, 1996) have invoked the felicitous phrase ex juvantibus reasoning (reasoning backward from what works) to describe the error of inferring a disorder’s etiology from its treatment. Headaches, as the hoary example goes, are not caused by a deficiency of aspirin in the bloodstream.

There is reasonably strong evidence that many mental disorders fulfill Robins and Guze’s (1970) criteria for validity. When these criteria are met, the diagnosis offers additional information about the patient, information that was not available before this diagnosis was made. For example, if we correctly diagnose a patient with schizophrenia, we have learned that this patient:

- Is likely to exhibit psychotic symptoms that are not solely a consequence of a severe mood disturbance.
- Has a higher than expected likelihood of exhibiting abnormalities on several laboratory measures, including indices of sustained attention, smooth pursuit eye tracking, and detection of biological motion (Kim, Park, & Blake, 2011).
- Has a higher than average probability of having close biological relatives with schizophrenia and schizophrenia-spectrum disorders, such as schizotypal and paranoid personality disorders.
FUNCTIONS OF PSYCHIATRIC DIAGNOSIS

- Is likely to exhibit a chronic course, with few or no periods of entirely normal functioning, but approximately a 30% chance of overall improvement.
- Is likely to respond positively to medications that block the action of dopamine, although this is most likely to be the case for the positive symptoms (e.g., delusions, hallucinations) of the disorder (Subotnick et al., 2011).

Andreasen (1995) extended the Robins and Guze (1970) framework to incorporate indicators from molecular genetics, neurochemistry, and functional and structural brain imaging as additional validating indicators for psychiatric diagnoses (Kendell & Jablensky, 2003). Her friendly amendment to the Robins and Guze criteria allows us to use endophenotypic indicators to assist in the validation of a diagnosis. Endophenotypes are typically biomarkers or laboratory indicators, that is, “measurable components unseen by the unaided eye along the pathway between disease and distal genotype” (Gottesman & Gould, 2003, p. 636; Waldman, 2005). They are often contrasted with exophenotypes, the traditional signs and symptoms of a disorder.

We can view the process of validating psychiatric diagnoses within the overarching framework of construct validity (Cronbach & Meehl, 1955; Loevinger, 1957; Messick, 1995; but see Borsboom, Cramer, Kievit, Zand Scholten, & Franic, 2009, for a different perspective on construct validity), which refers to the extent to which a measure assesses a hypothesized attribute of individuals. As Morey (1991) noted, psychiatric classification systems are collections of hypothetical constructs; thus, the process of validating psychiatric diagnoses is also a process of construct validation. More broadly, we can conceptualize most or even all psychiatric diagnoses as open concepts (Meehl, 1977, 1990). Open concepts are marked by (a) fuzzy boundaries, (b) a list of indicators (signs and symptoms) that are indefinitely extendable, and (c) an unclear inner nature.

Recalling that psychiatric diagnoses are open concepts helps us to avoid the perils of premature reification of diagnostic entities (Faust & Miner, 1986). For example, the present diagnostic criteria for schizophrenia are not isomorphic with the latent construct of schizophrenia; they are merely fallible, albeit somewhat valid, indicators of this construct. Yet, the past few decades have occasionally witnessed a troubling tendency to reify and deify the categories within the current classification system, with some authors regarding them as fixed Platonic essences rather than as rough approximations to the true state of nature (Ghaemi, 2003; Michels, 1984). This error is manifested, for example, when journal or grant reviewers criticize researchers for examining alternative operationalizations of mental disorders that depart from those in the current diagnostic manual (see section Psychiatric Classification from DSM-I to the Present, later in this chapter). It is also manifested by the common error of referring to measures of certain psychiatric conditions as “gold standards” (see Skeem & Cooke, 2010, for a discussion of this tendency in the field of psychopathy), a phrasing that is erroneous in view of the fact that all indicators of psychopathology are at present fallible and provisional.

In a classic article, Cronbach and Meehl (1955) adopted from neopositivist philosophers of science the term nomological network to designate the system of lawful relationships conjectured to hold between theoretical entities (states, structures, events, dispositions) and observable indicators. They selected the network metaphor to emphasize the structure of such systems in which the nodes of the network, representing the postulated theoretical entities, are connected by the strands of the network, representing the lawful relationships hypothesized to hold among the entities (Garber & Strassberg, 1991).
For Cronbach and Meehl (1955), construct validation is a progressive and never-ending process of testing the links between hypothesized strands of the nomological network, especially those that connect latent constructs, which include psychiatric diagnoses (e.g., schizophrenia and major depression), to manifest indicators, which include the external criteria (e.g., laboratory tests and family history) laid out by Robins and Guze (1970). The more such construct-to-manifest indicator links are corroborated, the more certain we can be that our conception of the diagnosis in question is accurate. From this perspective, the approach to diagnostic validation outlined by Robins and Guze is merely one specific instantiation of construct validation.

One limitation of the Robins and Guze (1970) approach to construct validation is its exclusive emphasis on external validation, that is, the process of ascertaining the construct’s associations with correlates that lie outside of the construct itself. As Skinner (1981, 1986; also Loevinger, 1957) observed, internal validation, ascertaining the construct’s inner structure, is also a key component of construct validation. Internal validation can help investigators to test hypotheses regarding a construct’s homogeneity (versus heterogeneity) and factor structure (Waldman et al., 1995). For example, if analyses suggest that a diagnosis consists of multiple and largely independent subtypes, the validity of the diagnosis would be called into question. Alternatively, factorial validity (i.e., the extent to which the factor structure of a diagnosis comports with theoretical predictions) can inform debates regarding the validity of a diagnosis. For example, factor analyses of attention-deficit/hyperactivity disorder (ADHD) generally support the separation of inattention from impulsivity and hyperactivity, as implied by the DSM criteria for this disorder (Martel, Von Eye, & Nigg, 2010).

In summary, valid psychiatric diagnoses serve three primary functions:

1. They summarize distinctive features of a disorder and thereby allow professionals to communicate clearly with one another.
2. They place each diagnosis under the umbrella structure of other diagnoses. This nosological framework links one diagnosis to both more and less related diagnoses.
3. They provide practitioners and researchers with surplus information regarding a diagnosed patient’s clinical profile, laboratory findings, natural history, family history, and possible response to treatment; they may also offer information regarding endophenotypic indicators.

Misconceptions Regarding Psychiatric Diagnosis

Beginning psychology graduate students and much of the general public hold a plethora of misconceptions regarding psychiatric diagnosis; we examine five such misconceptions here. Doing so will also permit us to introduce a number of key principles of psychiatric diagnosis. As we will discover, refuting each misconception regarding psychiatric diagnosis affirms at least one important principle.

**Misconception #1: Mental Illness Is a Myth**

The person most closely associated with this position is the late psychiatrist Thomas Szasz (1960), who argued famously for over 40 years that the term *mental illness* is a false
and misleading metaphor (Schaler, 2004). For Szasz, individuals whom psychologists and psychiatrists term mentally ill actually suffer from problems in living (that is, difficulties in adjusting their behaviors to the demands of society). Moreover, Szasz contended that mental health professionals often apply the mental illness label to nonconformists who jeopardize the status quo (Sarbin, 1969; Szasz, 1960). This label serves as a convenient justification for forcing maladjusted, malcontented, and maverick members of society to comply with prevailing societal norms.

Specifically, Szasz maintained that medical disorders can be clearly recognized by a lesion to the anatomical structure of the body, but that the disorder concept cannot be imported to the mental realm because there is no such lesion to indicate deviation from the norm. According to him only the body can become diseased, so mentally ill people do not suffer from an illness akin to a medical disorder.

It is undeniable that psychiatric diagnoses are sometimes misapplied. Nevertheless, this legitimate pragmatic concern must be logically separated from the question of whether the mental illness concept itself exists (Wakefield, 1992). We should recall the logical principle of *abusus non tollit usum* (abuse does not take away use): Historical and sociological misuses of a concept do not negate its validity. Wakefield (1992) and others (e.g., Kendell, 1975) have observed that the Szaszian argument is problematic on several fronts. Among others, it assumes that medical disorders are in every case traceable to discernible lesions in an anatomical structure, and that all lesions give rise to medical disorders. Yet identifiable lesions cannot be found in certain clear-cut medical diseases—such as trigeminal neuralgia and senile pruritis—and certain identifiable lesions, such as albinism, are not regarded as medical disorders (Kendell, 1975; Wakefield, 1992). Szasz’s assertion that identifiable lesions are essentially synonymous with medical disorders is false; therefore, his corollary argument that mental disorders cannot exist because they are not invariably associated with identifiable lesions is similarly false.

**MISCONCEPTION #2: PSYCHIATRIC DIAGNOSIS IS MERELY PIGEONHOLING**

According to this criticism, when we diagnose people with a mental disorder, we deprive them of their uniqueness: We imply that all people within the same diagnostic category are alike in all important respects.

To the contrary, a psychiatric diagnosis does nothing of the sort; it implies only that all people with that diagnosis are alike in at least one important way. Psychologists and psychiatrists are well aware that even within a given diagnostic category, such as schizophrenia or bipolar I disorder, people differ dramatically in their race and cultural background, personality traits, interests, and cognitive skills (APA, 2013).

**MISCONCEPTION #3: PSYCHIATRIC DIAGNOSES ARE UNRELIABLE**

Reliability refers to the consistency of a diagnosis. As many textbooks in psychometrics remind us, reliability is a prerequisite for validity but not vice versa. Just as a bathroom scale cannot validly measure weight if it yields dramatically different weight estimates for the same person over brief periods of time, a diagnosis cannot validly measure a mental disorder if it yields dramatically different scores on measures of psychopathology across times, situations, and raters.
Because validity is not a prerequisite for reliability, extremely high reliability can exist without validity. A researcher who based diagnoses of schizophrenia on patients’ heights would end up with extremely reliable but entirely invalid diagnoses of schizophrenia.

There are three major subtypes of reliability. Contrary to popular (mis)conception, these subtypes are frequently discrepant with one another, so high levels of reliability for one metric do not necessarily imply high levels for the others.

Test-retest reliability refers to the stability of a diagnosis following a relatively brief time interval, typically about a month. In other words, after a short time lapse, will patients receive the same diagnoses? Note that we wrote brief and short in the previous sentences; marked changes following lengthy time lapses, such as several years, may reflect genuine changes in patient status rather than the measurement error associated with test-retest unreliability.

In general, we assess test-retest reliability using either a Pearson correlation coefficient or, more rigorously, an intraclass correlation coefficient. Intraclass correlations tend to provide the most stringent estimates of test-retest reliability because, in contrast to Pearson correlations, they are influenced not merely by the rank ordering and differences among people’s scores, but by their absolute magnitude.

Our evaluation of the test-retest reliability of a diagnosis hinges on our conceptualization of the disorder. We should anticipate high test-retest reliability only for diagnoses that are traitlike, such as personality disorders, or that tend to be chronic (long-lasting), such as schizophrenia. In contrast, we should not necessarily anticipate high levels of test-retest reliability for diagnoses that tend to be episodic (intermittent), such as major depression.

Internal consistency refers to the extent to which the signs and symptoms comprising a diagnosis hang together—that is, correlate highly with one another. We generally assess internal consistency using such metrics as coefficient alpha (Cronbach, 1951) or the mean interitem correlation. Cronbach’s alpha can overestimate the homogeneity of a diagnosis, however, if this diagnosis contains numerous signs and symptoms, because this statistic is affected by test length (Schmidt, Le, & Ilies, 2003). We should anticipate high levels of internal consistency for most conditions in the current classification system given that most are syndromes, which are typically constellations of signs and symptoms that covary across people.

Interrater reliability is the degree to which two or more observers, such as different psychologists or psychiatrists, agree on the diagnosis of a set of individuals. High interrater reliability is a prerequisite for all psychiatric diagnoses, because different observers must agree on the presence or absence of a condition before valid research on that condition can proceed.

Many early studies of psychiatric diagnosis operationalized interrater reliability in terms of percentage agreement, that is, the proportion of cases on which two or more raters agree on the presence or absence of a given diagnosis. Nevertheless, measures of percentage agreement tend to overestimate interrater reliability. Here’s why: Imagine two diagnosticians working in a setting (e.g., an outpatient phobia clinic) in which the base rate (prevalence) of the diagnosis of specific phobia is 95%. The finding that they agree with each other on the diagnosis of specific phobia 95% of the time would hardly be impressive and could readily be attributed to chance. As a consequence, most investigators today operationalize interrater reliability in terms of the kappa coefficient, which assesses the degree to which raters agree on a diagnosis after correcting for chance, with chance being the base rate of the disorder in question. Nevertheless, the kappa coefficient often
provides a conservative estimate of interrater reliability, as the correction for chance sometimes penalizes raters for their independent expertise (Meyer, 1997).

Many laypersons and even political pundits believe that psychiatric diagnoses possess low levels of reliability, especially interrater reliability. This perception is probably fueled by high-profile media coverage of dueling expert witnesses in criminal trials in which one expert diagnoses a defendant as schizophrenic, for example, and another diagnoses him as normal. After the widely publicized 1982 trial of John Hinckley, who was acquitted on the basis of insanity for his attempted assassination of then-president Ronald Reagan, political commentator George Will maintained (on national television) that the disagreements among expert witnesses regarding Hinckley’s diagnosis merely bore out what most people already knew: that psychiatric diagnosis is wildly unreliable (Lilienfeld, 1995).

Yet there is a straightforward explanation for such disagreement: Given the adversarial nature of our legal system, the prosecution and defense typically go out of their way to find expert witnesses who will support their point of view. This inherently antagonistic arrangement virtually guarantees that the interrater reliabilities of experts in criminal trials will be modest at best.

Certainly, the interrater reliability of psychiatric diagnoses is far from perfect. Yet for most major mental disorders, such as schizophrenia, mood disorders, anxiety disorders, and alcohol use disorder (alcoholism), interrater reliabilities are typically about as high—intraclass correlations between raters of 0.8 or above, out of a maximum of 1.0—as those for most well-established medical disorders (Lobbestael, Leurgans, & Amtz, 2011; Matarazzo, 1983). Still, the picture is not entirely rosy. For many personality disorders in particular, interrater reliabilities tend to be considerably lower than for other conditions (Maffei et al., 1997; Zimmerman, 1994), probably because most of these disorders comprise highly inferential constructs (e.g., lack of empathy) that raters find difficult to assess during the course of brief interviews.

**Misconception #4: Psychiatric Diagnoses Are Invalid**

From the standpoint of Szasz (1960) and other critics of psychiatric diagnosis (Eysenck, Wakefield, & Friedman, 1983), psychiatric diagnoses are largely useless because they do not provide us with new information. According to them, diagnoses are merely descriptive labels for behaviors we do not like. Millon (1975) proposed a helpful distinction between psychiatric labels and diagnoses; a label simply describes behaviors, whereas a diagnosis helps to explain them.

When it comes to a host of informal pop psychology labels, like sexual addiction, Peter Pan syndrome, codependency, shopping disorder, Internet addiction, and road rage disorder, Szasz and his fellow critics probably have a point. Most of these labels merely describe collections of socially problematic behavior and do not provide us with much, if any, new information (McCann, Shindler, & Hammond, 2003). The same may hold for some personality disorders in the current classification system. For example, the diagnosis of dependent personality disorder, which has been retained in DSM-5 (APA, 2013), arguably appears to do little more than describe ways in which people are pathologically dependent on others, such as relying excessively on others for reassurance and expecting others to make everyday life decisions for them.

Yet, as we have already seen, many psychiatric diagnoses, such as schizophrenia, bipolar I disorder, and panic disorder, do yield surplus information (Robins & Guze, 1970;
Waldman et al., 1995) and, therefore, possess adequate levels of validity. Nevertheless, because construct validation, like all forms of theory testing in science, is a never-ending process, the validity of these diagnoses is likely to improve over time with subsequent revisions to the present classification system.

**MISCONCEPTION #5: PSYCHIATRIC DIAGNOSES STIGMATIZE PEOPLE, AND OFTEN RESULT IN SELF-FULFILLING PROPHECIES**

According to advocates of labeling theory, including Szasz (1960), Sarbin (1969), and Scheff (1975), psychiatric diagnoses produce adverse effects on labeled individuals. They argue that diagnostic labels not only stigmatize patients, but also frequently become self-fulfilling prophecies, leading observers to interpret ambiguous and relatively mild behaviors (e.g., occasional outbursts of anger) as reflecting serious mental illness.

A sensational 1973 and widely cited (over 2,300 citations as of this writing) study by Rosenhan appeared to offer impressive support for labeling theory. Rosenhan, along with seven other normal individuals, posed as pseudopatients (fake patients) in 12 U.S. psychiatric hospitals (some of the pseudopatients presented at more than one hospital). They informed the admitting psychiatrist only that they were hearing a voice saying, “empty,” “hollow,” and “thud.” All were promptly admitted to the hospital and remained there for an average of 3 weeks, despite displaying no further symptoms or signs of psychopathology. In 11 of 12 cases, they were discharged with diagnoses of schizophrenia in remission (the 12th pseudopatient was discharged with a diagnosis of manic depression in remission).

Rosenhan (1973) noted that the hospital staff frequently interpreted pseudopatients’ innocuous behaviors, such as note taking, as indicative of abnormality. In case summaries, these staff also construed entirely run-of-the-mill details of pseudopatients’ life histories, such as emotional conflicts with parents during adolescence, as consistent with their present illness. These striking results led Rosenhan to conclude that psychiatric labels color observers’ perceptions of behavior, often to the point that they can no longer distinguish mental illness from normality.

Even today, some writers interpret Rosenhan’s findings as a resounding affirmation of labeling theory (e.g., Slater, 2004; but see Spitzer, Lilienfeld, & Miller, 2005, for a critique of Slater, 2004). Yet, the evidence for labeling theory is less impressive than it appears. As Spitzer (1975) observed, the fact that all 12 of Rosenhan’s pseudopatients were released with diagnoses in remission (meaning showing no indications of illness) demonstrates that the psychiatrists who treated them were in all cases able to distinguish mental illness from normality. Spitzer went further, demonstrating in a survey of psychiatric hospitals that in-remission diagnoses of previously psychotic patients are exceedingly infrequent, showing that the psychiatrists in Rosenhan’s study successfully made an extremely rare judgment with perfect consensus.

Although incorrect psychiatric diagnoses can engender stigma, at least in the short run (Harris, Milich, Corbitt, Hoover, & Brady, 1992; Milich, McAninich, & Harris, 1992), there is scant evidence to support the popular claim that correctly applied psychiatric diagnoses do so. The lion’s share of the research suggests that stigma is a consequence not of diagnostic labels, but rather of disturbed and sometimes disturbing behavior that precedes labeling (Link & Cullen, 1990; Ruscio, 2004). For example, within 30 minutes or less, children begin to react negatively to children with attention-deficit/hyperactivity
What Is Mental Disorder?

Our discussion up to this point presupposes that the boundaries of the higher-order concept of “disorder,” including mental disorder, are clear-cut or at least reasonably well delineated.¹ To develop a classification system of disorders, one must first be able to ascertain whether a given condition is or is not a disorder. Yet the answer to the question of how best to define disorder, including mental disorder, remains elusive (Gorenstein, 1992). The issues here are of more than academic interest, because each revision of psychiatry’s diagnostic manual has been marked by contentious disputes regarding whether such conditions as ADHD, premenstrual dysphoric disorder, and, more recently, binge eating disorder, attenuated psychosis syndrome, and hypersexual disorder are genuine disorders (Frances & Widiger, 2012; Wakefield, 1992). The fact that homosexuality was removed from the formal psychiatric classification system in 1974 by a majority vote of the membership of the American Psychiatric Association (Bayer & Spitzer, 1982) further demonstrates that these debates are frequently resolved more by group consensus than by scientific research.

Here we evaluate several influential attempts to delineate the boundaries of disorder. As we will discover, each approach has its limitations but each captures something important about the concept of disorder. As we will also discover, these approaches differ in the extent to which they embrace an essentialist as opposed to a nominalist view of disorder (Ghaemi, 2003; Scadding, 1996). Advocates of an essentialist view (Widiger & Trull, 1985) believe that all disorders share some essence or underlying property, whereas advocates of a nominalist view (Lilienfeld & Marino, 1995, 1999; Rosenhan & Seligman, 1995) believe that the higher-order concept of disorder is a social construction that groups together a variety of largely unrelated conditions for the purposes of social or semantic convenience.

Statistical Model

Advocates of a statistical model, such as Cohen (1981), equate disorder with statistical rarity. According to this view, disorders are abnormal because they are infrequent in the general population. This definition accords with findings that many mental disorders are indeed rare; schizophrenia, for example, is found in about 1% of the population across much of the world (APA, 2013).

¹ In our discussion of the definition of disorder, we use the term disorder, including mental disorder, generically to refer to all medical and psychological conditions and do not distinguish disorder from disease (Wakefield, 1992).
Yet, a purely statistical model falls short on at least three grounds. First, it offers no guidance for where to draw cutoffs between normality and abnormality. In many cases, these cutoffs are scientifically arbitrary. Second, it is silent on the crucial question of which dimensions are relevant to abnormality. As a consequence, a statistical model misclassifies high scores on certain adaptive dimensions (like intelligence, creativity, and altruism) as inherently abnormal. Moreover, it does not explain why high scores on certain dimensions (e.g., anxiety) but not on others (e.g., hair length) are pertinent to psychopathology. Third, by definition a statistical model assumes that all common conditions are normal (Wakefield, 1992). Yet the common cold is still an illness despite its essentially 100% lifetime prevalence in the population, and the Black Death (bubonic plague) was still an illness in the mid-1300s despite wiping out approximately one third of the European population.

**SUBJECTIVE DISTRESS MODEL**

Proponents of a subjective distress model maintain that the core feature distinguishing disorder from nondisorder is psychological pain. This model unquestionably contains a large kernel of truth; many serious mental illnesses (such as major depression, obsessive-compulsive disorder, generalized anxiety disorder, and gender identity disorder) are marked by considerable distress, even anguish.

The subjective distress model also falls short of an adequate definition of mental illness, because it fails to distinguish ego-dystonic conditions (those that conflict with one’s self-concept) from ego-syntonic conditions (those that are consistent with one’s self-concept). Although most mental disorders (such as major depression and generalized anxiety disorder) are typically ego-dystonic, some (such as antisocial personality disorder and bipolar I disorder, at least in its manic phase) are largely or entirely ego-syntonic, because individuals with these conditions frequently see little or nothing wrong with their behavior. They experience little or no distress in conjunction with their condition, and frequently seek treatment only when demanded by courts or significant others, or when their condition is complicated by a secondary condition that generates interpersonal difficulties (e.g., alcohol use disorder). Moreover, approximately half of patients with schizophrenia and other severe psychotic conditions are afflicted with anosognosia, meaning that they are unaware of the fact that they are ill (Amador & Paul-Odouard, 2000).

**BIOLOGICAL MODEL**

Proponents of a biological model (Kendell, 1975) contend that disorder can be defined in terms of a biological or evolutionary disadvantage to the organism, such as reduced life span or fitness (i.e., the ability to pass on genes to subsequent generations). Indeed, some mental disorders are associated with biological disadvantages; for example, major depression is associated with a dramatically increased risk for completed suicide (Joiner, 2006), and between 5% and 10% of patients with anorexia nervosa eventually die from complications due to starvation (Goodwin & Guze, 1996).

A biological model, however, also falls prey to numerous counterexamples. For example, being a soldier in front-line combat is not a disorder despite its average adverse effect on longevity and fitness. Conversely, some relatively mild psychological conditions, such as specific phobia, are probably not associated with decreased longevity or fitness, yet are still mental disorders.
WHAT IS MENTAL DISORDER?

NEED FOR TREATMENT

One parsimonious definition is simply that disorders are a heterogeneous class of conditions all characterized by a perceived need for medical intervention on the part of health (including mental health) professionals (Kraupl Taylor, 1971). Like other definitions, this definition captures an important truth: Many or most mental disorders, such as schizophrenia, bipolar I disorder, and obsessive-compulsive disorder, are indeed viewed by society as necessitating treatment. Nevertheless, this definition also falls victim to counterexamples. For example, pregnancy clearly is associated with a perceived need for medical intervention, yet it is not regarded as a disorder.

HARMFUL DYSFUNCTION

In an effort to remedy the shortcomings of extant models of disorder, Wakefield (1992) proposed a hybrid definition that incorporates both essentialist and nominalist features. According to Wakefield, all disorders, including all mental disorders, are harmful dysfunctions: socially devalued (harmful) breakdowns of evolutionarily selected systems (dysfunctions). For example, according to Wakefield, panic disorder is a mental disorder because it (a) is negatively valued by society and often by the individual afflicted with it, and (b) reflects the activation of the fight-flight system in situations for which it was not evolutionarily selected, namely those in which objective danger is absent. In other words, panic attacks are false alarms (Barlow, 2001). Wakefield’s operationalization of disorder has its strengths; for example, it acknowledges (correctly) that most and perhaps all disorders are viewed negatively by others. The concept of disorder, including mental disorder, is clearly associated with social values. As Wakefield (1992) noted, however, social devaluation is not sufficient to demarcate disorder from nondisorder, claims by Szasz (1960) to the contrary. For example, rudeness, laziness, slovenliness, and even racism are viewed negatively by society, but are not disorders (for a dissenting view regarding racism, see Poussaint, 2002). Therefore, Wakefield contends that something else is necessary to distinguish disorder from nondisorder, namely evolutionary dysfunction.

Nevertheless, the dysfunction component of Wakefield’s analysis appears to fall prey to counterexamples. In particular, many medical disorders appear to be adaptive defenses against threat or insult. For instance, the symptoms of influenza (flu), such as vomiting, coughing, sneezing, and fever, are all adaptive efforts to expel an infectious agent rather than failures or breakdowns in an evolutionarily selected system (Lilienfeld & Marino, 1999; Neese & Williams, 1994). Such counterexamples appear to falsify the harmful dysfunction analysis. Similarly, many psychological conditions appear to be adaptive reactions to perceived threat. For example, in contrast to other forms of specific phobia, blood/injection/injury phobia is marked by a coordinated set of dramatic parasympathetic reactions—especially rapid decreases in heart rate and blood pressure—that were almost surely evolutionarily selected to minimize blood loss (Barlow, 2001). Although these responses may not be especially adaptive in the early 21st century, they were adaptive prior to the advent of Band-Aids, tourniquets, and anticoagulants (Lilienfeld & Marino, 1995).

ROSCHIAN ANALYSIS

An alternative approach to defining disorder is radically different. According to a Roschian analysis, the attempt to define disorder explicitly is sure to fail because disorder
is intrinsically undefinable (Gorenstein, 1992). Drawing on the work of cognitive psychologist Eleanor Rosch (Rosch, 1973; Rosch & Mervis, 1975), advocates of a Roschian analysis contend that the concept of mental disorder lacks defining (i.e., singly necessary and jointly sufficient) features and possesses intrinsically fuzzy boundaries. In this respect, mental disorder is similar to many other concepts. For example, the concept of a chair lacks strictly defining features (e.g., a human-made object with four legs that someone can sit on does not succeed as a defining feature, because one can sit on a table and many chairs do not have four legs) and displays unclear boundaries. In addition, the concept of mental disorder, like many other concepts, is organized around a prototype that shares all of the features of the category. Just as certain chairs (e.g., a typical office chair) are more chairlike than others (e.g., a beanbag), certain mental disorders (e.g., schizophrenia, panic disorder) are more disorder-like than others (e.g., hypersexual disorder, premenstrual dysphoric disorder). Not surprisingly, it is at the fuzzy boundaries of disorder where controversies concerning whether a psychological condition is really a disorder most frequently arise. According to the Roschian analysis, these controversies are not only inevitable, but are also not resolvable by scientific data.

Even if the Roschian analysis is correct (for criticisms of this approach, see Cooper, 2011; Wakefield, 1999; and Widiger, 1997), it would not imply that specific mental disorders themselves are not amenable to scientific inquiry. As Gorenstein (1992) noted, the concept of a drug is inherently undefinable; there are no scientific criteria for deciding whether caffeine, nicotine, and many other widely used but addictive substances are drugs. Yet, this problem has not stopped psychopharmacologists from studying specific drugs’ properties, modes of action, or behavioral effects. Nor should the absence of an explicit definition of mental disorder preclude psychopathology researchers from investigating the diagnosis, etiology, treatment, and prevention of schizophrenia, major depression, panic disorder, and other conditions.

Psychiatric Classification From DSM-I to the Present

Prior to the 1950s, the state of psychiatric classification in the United States was largely disorganized, as no standard system was in place for operationalizing specific mental disorders. Indeed, prior to World War I, there was scant interest in developing a systematic classification of mental disorders (Grob, 1991), and even after World War I it took over three decades to put in place a consensual system of classification. As a consequence, what one diagnostician meant by major depression might bear minimal correspondence to what another diagnostician meant by the same term.

DSM-I AND DSM-II

This situation gradually began to change in 1918, when the U.S. Bureau of the Census released the Statistical Manual for the Use of Institutions of the Insane, which divided mental disorders (largely psychoses) into 22 groups; this manual was revised 10 times through 1942 (Grob, 1991). It was not until 1952, however, that the American Psychiatric Association released the first edition of its Diagnostic and Statistical Manual of Mental Disorders, abbreviated as DSM-I (APA, 1952). Although DSM-I was a slim 132 pages in length, it was a landmark. For the first time, it offered reasonably clear, albeit brief, descriptions of major psychiatric diagnoses, thereby facilitating interrater reliability
among clinicians and researchers. Here, for example, was the description for “Manic Depressive Reaction, Depressed type” (later to become major depression) in *DSM-I*:

> Here will be classified those cases with outstanding depression of mood and with mental and motor retardation and inhibition; in some cases there is much uneasiness and apprehension. Perplexity, stupor or agitation may be prominent symptoms, and may be added to the diagnosis as manifestations. (APA, 1952, p. 25)

*DSM-II* appeared 16 years later (APA, 1968) and was similar in approach and scope to *DSM-I*, although it provided somewhat greater detail concerning the signs and symptoms of many diagnoses.

Despite their strengths, *DSM-I* and *DSM-II* suffered from several notable weaknesses, three of which we discuss here:

1. The interrater reliabilities of many of their diagnoses were still problematic, probably because these manuals consisted of global and often vague descriptions of mental illnesses that necessitated considerable subjective judgment on the part of diagnosticians. For example, returning to the description of manic depressive reaction, depressed type, *DSM-I* is silent on what qualifies as “outstanding depression,” and how much motor retardation and inhibition are necessary for the diagnosis.

2. *DSM-I* and *DSM-II* were not theoretically agnostic. In particular, they were influenced by psychoanalytic concepts of mental disorders and often made references to defense mechanisms and other concepts derived from Freudian theory. As a consequence, diagnosticians whose orientation was not psychoanalytic, such as behaviorists, cognitive behaviorists, or humanistic-existential psychologists, found these classification systems difficult to use. Interestingly, a recent effort to revive psychoanalytic conceptions within psychiatric diagnosis emerged in the form of the *Psychodynamic Diagnostic Manual* (Alliance of Psychoanalytic Organizations, 2006), which conceptualizes mental disorders largely in terms of psychodynamic concepts, such as unconscious drives and defense mechanisms. Nevertheless, this manual appears to have exerted little impact outside of psychoanalytic circles.

*DSM-I* and *DSM-II* also conceptualized mental disorders largely from the perspective of psychiatrist Adolph Meyer (1866–1950), who regarded most forms of psychopathology as aberrant reactions to life events (Lief, 1948), hence the use of the term *reaction* in the diagnosis of manic depressive reaction, depressed type and many other *DSM-I* and *DSM-II* diagnoses. Nevertheless, this assumption was based more on plausible theoretical conjecture than on evidence.

3. Despite their Meyerian emphasis, *DSM-I* and *DSM-II* focused almost exclusively on patients’ mental disorders per se, and largely neglected to consider contextual factors, such as co-occurring medical conditions, life stressors, and adaptive functioning, which can play key roles in the etiology and maintenance of psychopathology.

### DSM-III and Beyond

Largely in response to these criticisms, the American Psychiatric Association, with psychiatrist Robert Spitzer at the helm, released *DSM-III* in 1980. As most historians of psychiatric classification and diagnosis now recognize, *DSM-III* was an important revision
of the diagnostic manual; it represented a radical change in thinking and approach from all that came before, and has provided the template for all that has come since (Klerman, 1984; Mayes & Horwitz, 2005), including the newest revision, *DSM-5* (APA, 2013). In this respect, it was every bit as much a landmark, if not more, than was *DSM-I*. Coming in at a hefty 494 pages, a nearly fourfold increase from *DSM-II*, *DSM-III* not only dramatically increased the coverage of mental disorders—from 163 to 224—but also presented far more detailed guidelines than did its predecessors for establishing diagnoses. The operational and philosophical approach of *DSM-III* is often termed neo-Kraepelinian (Compton & Guze, 1995) because it followed in the footsteps of the great German psychiatrist Emil Kraepelin (1856–1926), who grouped and differentiated psychological conditions on the basis on their signs, symptoms, and natural histories. By and large, this neo-Kraepelinian approach has been carried over in the development of *DSM-5*.

**Diagnostic Criteria, Algorithms, Hierarchical Exclusion Rules, and a Multiaxial Approach**

In accord with its neo-Kraepelinian emphasis, *DSM-III* instituted several major changes in psychiatric classification and diagnosis. First and foremost, *DSM-III* introduced: (a) standardized diagnostic criteria and (b) algorithms, or decision rules, for each diagnosis. For a number of major disorders, the interrater reliability and validity of the proposed *DSM-III* criterion sets were pilot-tested in systematic field trials. Rather than merely describing each diagnosis as *DSM-I* and *DSM-II* had done, *DSM-III* explicitly delineated the signs and symptoms comprising each diagnosis and the method by which these signs and symptoms needed to be combined to establish each diagnosis. In these respects, it was influenced heavily by the pioneering efforts of the St. Louis group at Washington University (including Robins, Guze, Winokur, and other giants of descriptive psychopathology), who had introduced preliminary diagnostic criteria and algorithms for 14 major mental disorders in the early 1970s (the “Feighner criteria”; see Feighner et al., 1972). Another major precursor of *DSM-III* was the Research Diagnostic Criteria (RDC), which expanded the Feighner criteria by adding criteria for several other disorders (Spitzer, Endicott, & Robins, 1978).

As an example of *DSM-III*’s highly structured approach to diagnosis, we can examine its criteria for a major depressive episode (MDE). To meet diagnostic criteria for MDE, *DSM-III* required that clients (1) experience “dysphoric mood or loss of interest or pleasure in all or almost all activities” (p. 213; with dysphoric mood described in terms of seven symptoms, including depression, hopelessness, and irritability), and (2) experience at least four of eight signs and symptoms, such as poor appetite, insomnia, loss of energy, difficulty thinking and concentrating, nearly every day for at least a 2-week period. Compare the specificity of these criteria with the skimpy and highly impressionistic description in *DSM-I* presented earlier.

*DSM-III* also outlined hierarchical exclusion rules for many diagnoses; such rules prevent clinicians and researchers from making these diagnoses if other diagnoses can account for their clinical picture. For example, *DSM-III* forbade clinicians and researchers from making a diagnosis of major depressive episode if the episode was superimposed on schizophrenia, a schizophreniform disorder, or a paranoid disorder, or if it appeared to be due to either an organic mental disorder (e.g., hypothyroidism) or uncomplicated bereavement (a prolonged grief reaction). Among other things, hierarchical exclusion rules remind diagnosticians to “think organic”: that is, to rule out potential physical causes of mental disorders before diagnosing them (Morrison, 1997).
**Psychiatric Classification From DSM-I to the Present**

*DSM-III*’s use of diagnostic criteria, algorithms, and hierarchical exclusion rules has been maligned by many commentators as the “Chinese menu” approach to diagnosis (choose three from column A, two from column B, four from column C). Despite these criticisms, there is evidence that this approach markedly decreased the subjectivity of diagnostic decision making and increased the interrater reliabilities of many diagnoses (Spitzer, Forman, & Nee, 1979). However, some authors argue that these increases were exaggerated by *DSM-III*’s proponents (Kirk & Kutchins, 1992).

The interrater reliability of *DSM* diagnoses has also been enhanced by the development of structured and semistructured diagnostic interviews, such as the Structured Clinical Interview for *DSM* (SCID; First, Spitzer, Gibbon, & Williams, 2002), which are coordinated explicitly around *DSM* criteria. These interviews consist of standardized questions—to be read verbatim by interviewers—and required and suggested follow-up probes with which to assess specific diagnostic criteria. For example, the SCID provides the following question to assess the criterion of current, unexpected panic attacks in the *DSM-IV* diagnosis of panic disorder: “Have you ever had a panic attack, when you suddenly felt frightened or anxious or suddenly developed a lot of physical symptoms?” If the respondent replies yes, the SCID instructs the interviewer to ask, “Have these attacks ever come on completely out of the blue—in situations where you didn’t expect to be nervous or uncomfortable?” (First et al., 2002).

Finally, *DSM-III* adopted a multiaxial approach to diagnosis, which was dropped from *DSM-5* (APA, 2013). In a multiaxial approach, clients are described along a series of axes (that is, dimensions). A multiaxial approach forces clinicians to adopt a more holistic approach to diagnosis by considering variables in addition to the individuals’ mental disorders. In *DSM-III* (and its revision), the first two axes are restricted to mental illnesses (with Axis I containing major mental disorders, such as schizophrenia and major depression and Axis II containing personality disorders and mental retardation), and the last three axes assess other dimensions (e.g., medical conditions, life stressors, global adaptive functioning) often relevant to psychological adjustment.

Nevertheless, *DSM-5* jettisoned the multiaxial system of its predecessors, largely in response to the fact that the rationale for the Axis I–Axis II distinction was never grounded in high-quality scientific evidence (Harkness & Lilienfeld, 1997). As discussed later, there is increasing evidence that some Axis I conditions, including mood and anxiety disorders, are underpinned by dimensions (e.g., high levels of negative emotionality) similar or identical to those that underpin many Axis II conditions. Moreover, there is no compelling evidence for a qualitative (black-and-white) difference between Axis I and Axis II conditions.

**Theoretical Agnosticism**

In sharp contrast to its predecessors, *DSM-III* was agnostic with respect to etiology (with the principal exception of one diagnosis, posttraumatic stress disorder, which required the presence of a traumatic event ostensibly tied to the resulting symptoms of the disorder). In particular, *DSM-III* assiduously shunned concepts, such as defense mechanisms, that were tied to psychoanalysis or other specific theoretical orientations. By doing so, it permitted practitioners and researchers of varying persuasions to use the manual with equal ease and comfort. It also facilitated scientific progress by allowing researchers to pit differing theoretical orientations against each other to determine which offers the most scientifically supported etiological explanations for specific disorders (Wakefield, 1998).
**ISSUES IN DIAGNOSIS**

**DSM-III-R AND DSM-IV**

*DSM-III-Revised (DSM-III-R)*, which appeared in 1987, and *DSM-IV*, which appeared in 1994 (and in a more expanded text revision in 2000, which with minor exceptions was identical to *DSM-IV* in its diagnostic criteria), retained all of the major features and innovations of *DSM-III* (APA, 1987, 1994, 2000). Nevertheless, they continued to increase the *DSM*’s coverage of psychopathology; *DSM-IV*, which was 943 pages long, contained 374 diagnoses (APA, 2000).

*DSM-III-R* and *DSM-IV* gradually moved away from a monothetic approach to diagnosis, emphasized in much of *DSM-III*, toward a polythetic approach. In a monothetic approach, the signs and symptoms are singly necessary and jointly sufficient for a diagnosis. In contrast, in a polythetic approach the signs and symptoms are neither necessary nor sufficient for a diagnosis.

The potential disadvantage of a polythetic approach is extensive heterogeneity at the symptom and (perhaps) etiological levels. In *DSM-IV*, for example, 256 different symptom combinations are compatible with a diagnosis of borderline personality disorder. It is implausible that the etiologies of all of these combinations are similar, let alone identical. It is even possible for two people to meet criteria for obsessive-compulsive personality disorder yet share no criteria (Widiger, 2007). Nevertheless, many scholars argue that the potential disadvantage of symptomatic heterogeneity is outweighed by the higher interreliability of the polythetic approach (Widiger, Frances, Spitzer, & Williams, 1991). In a monothetic approach, a disagreement about the presence or absence of only one criterion necessarily leads to a disagreement about the presence or absence of the diagnosis. In contrast, in a polythetic approach, such disagreement often has no impact on levels of agreement about the presence or absence of the diagnosis, because raters can still agree on the presence or absence of the diagnosis even if they disagree on one or more specific criteria.

The shift toward a polythetic approach is also an implicit nod to the fact that few, if any, signs and symptoms of psychopathology are pathognomonic. A pathognomonic indicator can be used by itself to establish the presence or absence of a disorder. For example, Koplik’s spots—tiny spots in the mouth that look much like grains of sand surrounded by red rings—are essentially pathognomonic for measles. A sign or symptom can in principle be one-way pathognomonic, meaning that it is a perfect inclusion test (the sign or symptom’s presence always indicates the presence of the disorder) or two-way pathognomonic, meaning that it is both a perfect inclusion test and a perfect exclusion test (the sign’s or symptom’s presence always indicates the presence of the disorder, and the sign’s or symptom’s absence always indicates the absence of the disorder). With the possible exception of certain organic brain disorders, no DSM diagnoses boast a one-way pathognomonic indicator.

*DSM-III-R* and *DSM-IV* also witnessed a relaxation of many, though not all, of *DSM-III*’s hierarchical exclusion rules (Pincus, Tew, & First, 2004). This change largely reflected the paucity of research evidence concerning the causal primacy of certain disorders above others. In addition, many of these exclusion rules proved difficult to apply in practice, because they required subjective and highly inferential judgments of causal primacy on the part of diagnosticians.

Finally, *DSM-IV* added an appendix for culture-bound syndromes, recognizing the fact that some conditions vary, or at least vary markedly in their expression, across cultures (Draguns & Tanaka-Matsumi, 2003). Most of these culture-bound syndromes...
are widely known in non-Western cultures, although their etiology and relation to conditions diagnosed in Western cultures are poorly understood. For example, koro, an epidemic condition observed in parts of China and Malaysia, is marked by abrupt and intense fears that the penis (in men) or vulva or breasts (in women) are receding into the body. Still other culture-bound syndromes appear to be variants of diagnoses that we readily recognize in Western culture. For example, taijin kyofusho, common in Japan, refers to a fear of offending others by one’s appearance, body odor, nonverbal behavior, and so on. It may be a subspecies of social phobia that is especially prevalent in cultures, especially in Asia, that stress group harmony above individual autonomy (Kleinknecht, Dinnel, Tanouye-Wilson, & Lonner, 1994).

**DSM-5**

Following the publication of *DSM-IV* in 1994 and its text revision in 2000, a great deal of data accumulated regarding the prevalence and correlates of various *DSM* diagnoses. In an effort to accommodate these new data, *DSM-5*, spearheaded by psychiatrists David Kupfer and Darrel Regier, was published in May 2013 (APA, 2013) amid a host of controversies. By and large, *DSM-5* retained most of the major categories of *DSM-IV* but, as noted earlier, dropped the multiaxial system. In addition, one goal of *DSM-5* was to stem the tide of the perceived proliferation of new diagnoses by relying on rigorous validity data for potential new conditions. As will be discussed later, however, *DSM-5* has already been widely criticized on a number of grounds, including its decision to lower the threshold for several diagnostic categories (Batstra & Frances, 2012a). Moreover, others have charged that the field trials for *DSM-5* were inadequate, focusing largely on clinical feasibility, with scant examination of the validity of new diagnostic categories or the potential effects of alterations in extant categories on the prevalence of *DSM* disorders (Frances & Widiger, 2012).

**Criticisms of the Current Classification System**

Recent versions of the diagnostic manual, including *DSM-IV* and *DSM-5*, have helped to place the field of psychopathology on firmer scientific grounds, largely because they have established reasonably reliable operationalizations for most mental disorders and furthered the development of standardized instruments, such as structured psychiatric interviews, to assess these disorders. The theoretical agnosticism of recent *DSM*s has also facilitated research comparing the scientific support for competing theoretical conceptualizations of psychopathology (Wakefield, 1998). Despite the undeniable advances of *DSM-III* and its progeny, many critics have argued that these manuals are scientifically problematic in several respects. Here we examine four key criticisms of the *DSM-5* classification

---

2 One frequent criticism of the *DSM* revision process (for example, Caplan, 1995) that we do not discuss at length here is the reliance on committee consensus in settling on both (a) the inclusion and exclusion of specific disorders from the manual and (b) the diagnostic criteria for specific disorders, largely because we find this criticism to be without substantial merit. Although expert consensus inevitably introduces subjective and political considerations into the diagnostic revision process (Ghaemi, 2003; Kirk & Kutchins, 1992) and has almost certainly resulted in flawed decisions, it is almost surely superior to a system in which one appointed expert adjudicates scientifically complex disputes without the benefit of input from other experts. As Widiger and Clark (2000) observed, “no diagnostic manual can be constructed without a group of fallible persons interpreting the results of existing research” (p. 948). To paraphrase Winston Churchill’s famous wisecrack about democracy, the *DSM* revision process is probably the worst system possible except for every other system.
system, many of which also applied to DSM-III, DSM-III-R, and DSM-IV: comorbidity, medicalization of normality, neglect of the attenuation paradox, and unsupported retention of a categorical model.

**Comorbidity**

DSM-5, like DSM-III and its other progeny, is marked by high levels of co-occurrence and covariation among many of its diagnostic categories, a phenomenon known, perhaps misleadingly, as comorbidity (Caron & Rutter, 1991; Lilienfeld et al., 1994; Pincus et al., 2004). We say misleadingly because it is premature in most cases to assume that comorbidity reflects the overlap among etiologically distinct conditions, as opposed to slightly different variants of the same underlying condition (Drake & Wallach, 2007). Although comorbidity is frequent among all mental disorders, it is especially rampant among personality disorders (Widiger & Rogers, 1989). In one analysis based on multiple sites, patients who met criteria for one personality disorder on average met criteria for approximately two additional personality disorders—with 10% meeting criteria for four or more personality disorders (Stuart et al., 1998). One patient in a research study met criteria for all 10 DSM personality disorders (Widiger et al., 1998).

The extent of comorbidity among mental disorders, including personality disorders, is often underestimated in routine clinical practice (Zimmerman & Mattia, 2000), in part because of a phenomenon known as diagnostic overshadowing. Diagnostic overshadowing refers to the tendency for a more florid disorder to draw attention away from less florid co-occurring disorders, thereby leading diagnosticians to either overlook them or attribute them to the more florid disorder. For example, the dramatic symptoms of borderline personality disorder frequently lead clinicians to underdiagnose commonly co-occurring but less salient conditions, such as narcissistic and dependent personality disorders (Garb, 1998). The full extent of comorbidity among personality disorders typically becomes evident only when structured and semistructured diagnostic interviews, which force assessors to inquire about all diagnostic criteria, are administered (Zimmerman & Mattia, 2000).

There are multiple potential explanations for comorbidity—some primarily substantive, others primarily methodological (for reviews, see Cramer, Waldorp, van der Maas, & Borsboom, 2010; Klein & Riso, 1993; and Lilienfeld, 2003). On the substantive front, one disorder (e.g., generalized anxiety disorder) may predispose to another disorder (e.g., dysthymic disorder), the two disorders may mutually influence each other, or both disorders may be slightly different expressions (formes frustes) of the same latent liability, such as neuroticism or negative emotionality. On the methodological front, comorbidity may result from overlapping diagnostic criteria or from clinical selection bias (du Fort, Newman, & Bland, 1993), that is, the tendency for psychiatric patients with one disorder, such as alcohol use disorder, to seek treatment only when they develop a co-occurring disorder, such as major depression. Comorbidity can also arise from Berksonian bias (Berkson, 1946), a selection bias resulting from the tendency of people with multiple conditions to be selected for research. For example, individuals with both specific phobia and major depression, most of whom are experiencing intense distress, may be more likely to volunteer for research on anxiety disorders than are people with specific phobia alone, who often exhibit relatively mild impairment. In addition, comorbidity can be produced by “logical errors” (Guilford, 1936), that is, mistakes stemming from the tendency of diagnosticians to assume that two largely unrelated conditions are correlated.
CRITICISMS OF THE CURRENT CLASSIFICATION SYSTEM

Whatever its causes, extensive comorbidity is potentially problematic for the DSM, because an ideal classification system yields largely mutually exclusive categories with few overlapping cases (Lilienfeld, Van Valkenberg, Larntz, & Akiskal, 1986; Sullivan & Kendler, 1998). As a consequence, such comorbidity may suggest that the current classification system is attaching multiple labels to differing manifestations of the same underlying condition. Defenders of the current classification system are quick to point out that high levels of comorbidity are also prevalent in organic medicine, and often indicate that certain conditions (e.g., diabetes) increase individuals’ risk for other conditions (e.g., blindness), a phenomenon that Kaplan and Feinstein (1974) termed pathogenetic comorbidity. Nevertheless, in stark contrast to organic medicine, in which the causal pathways contributing to pathogenetic comorbidity are often well understood, the causal pathways contributing to pathogenetic comorbidity in the domain of psychopathology generally remain unknown.

MEDICALIZATION OF NORMALITY

A number of critics have raised concerns that recent DSMs, DSM-5 in particular, have overmedicalized normality (Sommers & Satel, 2005). They have done so, these authors contend, in two ways: (1) increasing the number of diagnoses and (2) lowering the threshold for a number of extant diagnoses. In this way, recent DSMs, including DSM-5, may risk opening the floodgates to a pathologizing of largely normative behaviors, emotions, and thoughts. Probably the most vocal critic in this regard has been psychiatrist Allen Frances, who was the principal architect of DSM-IV. In a number of publications, Frances and others have decried DSM-5’s apparently lowered diagnostic thresholds for a number of conditions, as well as its introduction of new and largely unvalidated disorders (Batstra & Frances, 2012a).

Historically, one dramatic change from DSM-I to DSM-IV was the massive increase in the sheer number of diagnoses, a trend potentially reversed by DSM-5. Some critics have argued that this increase reflects the tendency for successive editions of the DSM to expand their range of coverage into new and largely uncharted waters (Houts, 2001). Many of these novel diagnoses, which describe relatively mild problems, may be of questionable validity. For example, the new DSM-5 diagnosis of disruptive mood dysregulation disorder, which is intended to capture many cases of what some authors believe to be pediatric bipolar disorder, has been harshly criticized by Frances (2012) and others for “turn[ing] temper tantrums into a mental disorder.” Another potential example is the new DSM-5 category of minor neurocognitive disorder, which some authors contend may unduly pathologize mild forgetfulness and other largely normative cognitive problems often associated with aging.

As Wakefield (2001) noted, however, there is little evidence that DSM actually expanded its range of coverage from DSM-III to DSM-IV. Although it is unclear at present, the same conclusion may hold for DSM-5. As Wakefield observed, most increases in the number of diagnoses across previous DSMs, since possibly stabilized by DSM-5, reflect an increased splitting of broader diagnoses into progressively narrower subtypes.

The distinction between splitting and lumping derives from biological taxonomy (Mayr, 1982) and refers to the difference between two classificatory styles: the tendency to subdivide broad and potentially heterogeneous categories into narrower and presumably more homogeneous categories (splitting) or the tendency to combine narrow and presumably more homogeneous categories into broad and potentially heterogeneous
categories (lumping). For example, given evidence that bipolar I disorder and bipolar II disorder are related (although by no means identical) conditions with relatively similar family histories, laboratory correlates, prognoses, and treatment response, should we keep these diagnoses separate or combine them into a more encompassing, albeit more heterogeneous, category? In the case of autism and allied conditions, the developers of DSM-5 elected to embrace a lumping approach, combining several conditions, such as autistic disorder, Asperger’s disorder, and childhood disintegrative disorder, into the broader domain of what is now termed autism spectrum disorder. Nevertheless, some authors have argued that this change will incorrectly exclude children with milder forms of autism spectrum conditions from the DSM (McPartland, Reichow, & Volkmar, 2012).

The splitting preferences of the architects of DSM-III, DSM-III-R, and DSM-IV in particular have been widely maligned (Houts, 2001). Herman van Praag (2000) even humorously “diagnosed” the DSM’s predilection for splitting as the disorder of nosologomania (also see Ghaemi, 2003). Nevertheless, a preference for splitting is defensible from the standpoint of research and nosological revision. A key point is that the relation between splitting and lumping is asymmetrical: If we begin by splitting diagnostic categories, we can always lump them later if research demonstrates that they are essentially identical according to the Robins and Guze (1970) criteria for validity. Yet, if we begin by lumping it would often be difficult or impossible to split later. As a consequence, we may overlook potentially crucial distinctions among etiologically separable subtypes that bear differing implications for treatment and prevention.

At the same time, it is unclear whether DSM-5’s new diagnoses, such as disruptive mood regulation disorder, reflect a splitting of the diagnostic pie into narrower slices or an enlargement of the pie. If the latter, DSM-5 may indeed risk extending the umbrella of pathology to relatively mild and normative problems.

As noted earlier, a second way in which recent DSMs, including DSM-5, may overmedicalize normality is by lowering the threshold for a number of conditions (Batstra & Frances, 2012a; Frances & Widiger, 2012). For example, by increasing the age of onset from 7 to 12 years of age and decreasing the proportion of symptoms for the diagnosis, DSM-5 appears to have made it easier to meet criteria for ADHD (Batstra & Frances, 2012b). Even more controversial was the decision in DSM-5 to remove the bereavement criterion for major depression, allowing individuals to be diagnosed with this condition as soon as 2 weeks following the death of a loved one (APA, 2013).

We believe the concerns of Frances and others regarding the potential overmedicalization of normality are important and worth raising. Nevertheless, the ultimate question is whether the changes in DSM-5 increase or decrease the construct validity of the resultant disorders, not whether they alter the prevalence of individuals diagnosed with these disorders (cf. Batstra & Frances, 2012a). The answer to the latter question will surely vary by disorder, and at present awaits clarification in light of future research.

**Neglect of the Attenuation Paradox**

Much of the impetus behind DSM-III was the laudable attempt to increase the reliability of psychiatric diagnosis and, thereby, place the fields of psychiatry and clinical psychology on firmer scientific footing. The importance of the reliability of diagnostic categories continues to be recognized in DSM-5. Nevertheless, reliability is only a means to an end,
Criticisms of the Current Classification System

namely validity; moreover, as noted earlier, validity is limited not by reliability per se, but by its square root (Meehl, 1986). Therefore, diagnoses of even modest reliability can, in principle, achieve high levels of validity.

Ironically, efforts to achieve higher reliability, especially internal consistency, can sometimes produce decreases in validity, a phenomenon that Loewinger (1957) referred to as the attenuation paradox (also see Clark & Watson, 1995). This paradox can result when an investigator uses a narrowly circumscribed pool of items to capture a broad and multifaceted construct. In such a case, the measure of the construct may exhibit high internal consistency yet low validity, because it does not adequately tap the full breadth and richness of the construct.

Some authors have argued that this state of affairs occurred with several DSM diagnoses. Putting it a bit differently, they have suggested that DSM-III and its descendants sacrificed validity at the altar of reliability (Vaillant, 1984). For example, the current DSM diagnosis of antisocial personality disorder (ASPD) is intended to assess the core interpersonal and affective features of psychopathic personality (psychopathy) delineated by Cleckley (1941), Karpman (1948), and others. Indeed, the accompanying text of DSM-IV even referred misleadingly to ASPD as synonymous with psychopathy (APA, 2000, p. 702). Because the developers of DSM-III (APA, 1980) were concerned that the personality features of psychopathy—such as guiltlessness, callousness, and self-centeredness—were difficult to assess reliably, they opted for a diagnosis emphasizing overt and easily agreed on antisocial behaviors—such as vandalism, stealing, and physical aggression (Hare, 2003; Lilienfeld, 1994). These changes may have resulted in a diagnosis with greater internal consistency and interrater reliability than the more traditional construct of psychopathy (although evidence for this possibility is lacking). Nevertheless, they may have also resulted in a diagnosis with lower validity, because the DSM diagnosis of ASPD largely fails to assess the personality features central to psychopathy (Lykken, 1995; Skeem, Polaschek, Patrick, & Lilienfeld, 2011). Indeed, accumulating evidence suggests that measures of ASPD are less valid for predicting a number of theoretically meaningful variables—including laboratory indicators—than are measures of psychopathy (Hare, 2003; also see Vaillant, 1984, for a discussion of the reliability trade-off in the case of the DSM-III diagnosis of schizophrenia).

Unsupported Retention of a Categorical Model

Technically, the DSM is agnostic on the question of whether psychiatric diagnoses are truly categories in nature, or what Meehl (Meehl & Golden, 1982) termed taxa, as opposed to continua or dimensions. Taxa differ from normality in kind, whereas dimensions differ in degree. Pregnancy is a taxon, as a woman cannot be slightly pregnant; in contrast, height is almost always a dimension (although certain rare taxonic conditions, like hormonal abnormalities, can lead to heights that differ qualitatively from the general population). The opening pages of DSM-IV state: “There is no assumption that each category of mental disorder is a completely discrete entity with absolute boundaries dividing it from other mental disorders or from no mental disorder” (p. xxxi). Yet at the measurement level, the DSM embraces an exclusively categorical model, classifying individuals as either meeting criteria for a disorder or not meeting them. In a highly contentious move, DSM-5 passed on an opportunity to embrace a dimensional model of personality disorders,
leaving the current categorical model in place and relegating a proposed dimensional alternative to Section III of the manual (devoted to provisional criterion sets meriting future consideration).

The *DSM* categorical model is problematic for at least two reasons. First, there is growing evidence from taxometric analyses (Meehl & Golden, 1982)—namely, those that allow researchers to ascertain whether a single observed distribution is decomposable into multiple independent distributions—that many or even most *DSM* diagnoses are underpinned by dimensions rather than taxa (Kendell & Jablensky, 2003), with schizophrenia and schizophrenia-spectrum disorders being notable probable exceptions (Lenzenweger & Korfine, 1992). This is particularly true for most personality disorders (Cloninger, 2009; Trull & Durrett, 2005), including antisocial personality disorder (Marcus, Lilienfeld, Edens, & Poythress, 2006). Even many or most other mental disorders—such as major depression (Slade & Andrews, 2005); social anxiety disorder (Kollman, Brown, Liverant, & Hofmann, 2006); and ADHD (Marcus, Norris, & Coccaro, 2012)—appear to be dimensional as opposed to taxonic in structure.

Second, setting aside the ontological issue of taxonicity versus dimensionality, there is good evidence that measuring most disorders (especially personality disorders) dimensionally by using the full range of scores almost always results in higher correlations with external validating variables than does measuring them categorically in an all-or-none fashion (Craighead, Sheets, Craighead, & Madsen, 2011; Markon, Chmielewski, & Miller, 2011; Ulrich, Borkenau, & Marneros, 2001). Such findings are not surprising given that artificial dichotomization of variables almost always results in a loss of information and, hence, statistical power (Cohen, 1983; MacCallum, Zhang, Preacher, & Rucker, 2002).

### The DSM: Quo Vadis?

In some respects, *DSM-III-R* and *DSM-IV* were disappointments, as they did not resolve many of the serious problems endemic to *DSM-III* (Ghaemi, 2003). If anything, comorbidity in *DSM-III-R* and *DSM-IV* mushroomed due to the dismantling of many hierarchical exclusion rules (Lilienfeld & Waldman, 2004). Moreover, some diagnostic categories (e.g., dependent personality disorder) of questionable validity remained. It is too early to tell whether *DSM-5* will help to resolve these and other problems. *DSM-5* presents both challenges and opportunities: challenges because many conceptual and methodological quandaries regarding psychiatric diagnosis remain unresolved, and opportunities because a new manual opens the door for novel approaches to the classification of psychopathology.

With these considerations in mind, we sketch out two promising future directions for psychiatric diagnosis: adoption of a dimensional approach and the incorporation of endophenotypic markers into psychiatric diagnosis (see Widiger & Clark, 2000, for other proposals for *DSM-5* and future *DSMs*).

### A Dimensional Approach

The accumulating evidence for the dimensionality of many psychiatric conditions, particularly personality disorders, has led many authors to suggest replacing or at least supplementing the *DSM* with a set of dimensions derived from the basic science of personality (Krueger et al., 2011; Widiger & Clark, 2000). One early candidate for a
dimensional model is the five-factor model (FFM; Goldberg, 1993), which consists of five major dimensions that have emerged repeatedly in factor analyses of omnibus (broad) measures of personality: extraversion, neuroticism, agreeableness, conscientiousness, and openness to experience (the FFM, incidentally, can easily be recalled using the waterlogged mnemonics of OCEAN or CANOE). These five dimensions also contain lower-order facets that provide a fine-grained description of personality; for example, the FFM dimension of extraversion contains facets of warmth, gregariousness, assertiveness, excitement seeking, and so on (Costa & McCrae, 1992).

The framers of DSM-5 considered a dimensional model for personality disorders influenced substantially by the work of Harkness (see Harkness & McNulty, 1994) and others. In this model, five broad dimensions of antagonism, detachment, negative affectivity (similar to but broader than neuroticism), disinhibition, and psychoticism would have been used to describe all personality variation in the abnormal range. Nevertheless, this bold proposal was ultimately vetoed by the American Psychiatric Association board of trustees, in part because its clinical feasibility was deemed to be insufficiently demonstrated. As noted earlier, however, these dimensions appear in Section III of the current manual in an effort to encourage further research with an eye toward DSM-6.

In addition to clinical feasibility, there are other potential objections to a dimensional model. For example, there is disagreement regarding both the precise nature and number of the personality dimensions to be used, with some authors advocating for alternative (e.g., three-dimensional) models. Another objection derives from the often neglected distinction between basic tendencies and characteristic adaptations in personality psychology (Harkness & Lilienfeld, 1997; McCrae & Costa, 1995). Basic tendencies are core personality traits, whereas characteristic adaptations are the behavioral manifestations of these traits. A large body of personality research suggests that basic tendencies can often be expressed in a wide variety of different characteristic adaptations depending on the upbringing, interests, cognitive skills, and other personality traits of the individual. For example, the scores of firefighters on a well-validated measure of the personality trait of sensation seeking (a construct closely related to, although broader than, risk taking) are significantly higher than those of college students, but comparable to those of incarcerated prisoners (Zuckerman, 1994). This finding dovetails with the notion that the same basic tendency, in this case sensation seeking, can be expressed in either socially constructive or destructive outlets, depending on yet unidentified moderating influences.

The distinction between basic tendencies and characteristic adaptations implies that personality dimensions may never be sufficient to capture the full variance in personality disorders. This is because these dimensions (basic tendencies) do not adequately assess many key aspects of psychopathological functioning, many of which can be viewed as maladaptive characteristic adaptations (Sheets & Craighead, 2007). This theoretical conjecture is corroborated by findings that the FFM dimensions do not account for a sizable chunk of variance in many DSM personality disorders. For example, in one study the correlations between FFM prototype scores of DSM personality disorders (derived from expert ratings of the FFM facets most closely associated with each disorder) and structured interview-based measures of these disorders were high for some disorders (e.g., avoidant personality disorder, $r = .67$) and modest and even negligible for others (e.g., obsessive-compulsive disorder, $r = .13$; Miller, Reynolds, & Pilkonis, 2004). The latter finding may reflect the fact that some obsessive-compulsive traits—such as perfectionism—may be adaptive in certain settings and, therefore, may not lead inevitably to personality pathology. Moreover, Skodol et al. (2005) reported that the
dimensions of the Schedule for Nonadaptive and Adaptive Personality (SNAP; Clark, 1993)—a measure that assesses many pathological behaviors associated with personality disorders—displayed incremental validity above and beyond the FFM dimensions in distinguishing among DSM personality disorders (also see Reynolds & Clark, 2001). This finding suggests that the FFM overlooks crucial distinctions captured by the SNAP, perhaps in part because the SNAP assesses not only basic tendencies but also the maladaptive characteristic adaptations of many personality disorders (Lilienfeld, 2005).

The findings reviewed here imply that a dimensional model may be useful in capturing core features of many DSM personality disorders. Nevertheless, they raise the possibility that personality dimensions may not be sufficient by themselves to capture personality pathology, because they cannot tell us whether individuals’ behavioral adaptations to these dimensions are adaptive or maladaptive, nor the phenotypic (behavioral) manifestations these adaptations have assumed.

**Endophenotypic Markers**

As noted earlier, considerable recent interest has focused on the use of endophenotypes in the validation of psychiatric diagnoses (Andreasen, 1995; Waldman, 2005). Nevertheless, endophenotypic markers have thus far been excluded from DSM diagnostic criterion sets, which consist almost entirely of the classical signs and symptoms of disorders (exophenotypes). This omission is noteworthy, because endophenotypes may lie closer to the etiology of many disorders than exophenotypes do.

This situation may change in coming years with accumulating evidence from studies of biochemistry, brain imaging, and performance on laboratory tasks; this evidence holds the promise of identifying more valid markers of certain mental disorders (Widiger & Clark, 2000). To take just two examples, many impulse control disorders (e.g., pathological gambling, intermittent explosive disorder) appear to be associated with low levels of serotonin metabolites (Moeller, Barratt, Dougherty, Schmitz, & Swann, 2001) and major depression is frequently associated with left frontal hypoactivation (Henriques & Davidson, 1991).

Nevertheless, at least two potential obstacles confront the use of endophenotypic markers in psychiatric diagnosis, the first conceptual and the second empirical. First, the widespread assumption that endophenotypic markers are more closely linked to underlying etiological processes than exophenotypic markers (Kihlstrom, 2002) is just that: an assumption. For example, the well-replicated finding that diminished amplitude of the P300 (a brain event–related potential appearing approximately 300 milliseconds following stimulus onset) is dependably associated with externalizing disorders—such as conduct disorder and substance dependence (Patrick et al., 2006)—could reflect the fact that P300 is merely a sensitive indicator of attention. As a consequence, diminished P300 amplitude could be a downstream consequence of the inattention and low levels of motivation often associated with externalizing disorders. This possibility would not necessarily negate the incorporation of P300 amplitude into diagnostic criterion sets, although it could raise questions concerning its specificity to externalizing disorders, let alone specific externalizing disorders.

Second, no endophenotypic markers yet identified are close to serving as inclusion tests for their respective disorders. Even smooth pursuit eye movement dysfunction, which is perhaps the most dependable biological marker of schizophrenia, is present only
in anywhere from 40% to 80% of patients with schizophrenia, so it would miss many individuals with the disorder. It may come closer, however, to serving as a good exclusion test, as it is present in only about 10% of normal individuals (Clementz & Sweeney, 1990; Keri & Janka, 2004). Thus, although endophenotypic markers may eventually add to the predictive efficiency of some diagnostic criteria sets, they are likely to be fallible indicators, just like traditional signs and symptoms. These markers also hold out the hope of assisting in the identification of more etiologically pure subtypes of disorders; for example, schizophrenia patients with abnormal smooth pursuit eye movements may prove to be separable in important ways from other patients with this disorder.

Until recently, most of the proposals to implement endophenotypic markers were limited to supplementing the diagnosis of existing psychiatric categories, such as major depression or bipolar I disorder. A more radical proposal emanates from the recent initiative supported by the National Institute of Mental Health (NIMH) to develop Research Domain Criteria (RDoC) as a full-fledged alternative to the DSM and similar diagnostic manuals. At this point in time, RDoC is more of an envisioned research approach than a proposed system. Nevertheless, its goal is to identify well-established psychobiological systems that undergird psychopathology (Morris & Cuthbert, 2012), along with promising markers of these systems. Examples of such systems might include reward systems, fear systems, impulse control systems, and working memory. In turn, each of these systems could be measured using indicators at different levels of analysis, including observable behavior, self-report measures, laboratory measures, and brain imaging findings (Insel et al., 2010; Sanislow et al., 2010). Ultimately, such a system could supplement or even supplant the extant DSM system, but as of this writing progress along these lines remains preliminary.

**Summary and Future Directions**

We conclude the chapter with 10 take-home messages:

1. A systematic system of psychiatric classification is a prerequisite for psychiatric diagnosis.
2. Psychiatric diagnoses serve important, even essential, communicative functions.
3. A valid psychiatric diagnosis gives us new information—for example, it tells about the diagnosed individual’s probable family history, performance on laboratory tests, natural history, and perhaps response to treatment—and it also distinguishes that person’s diagnosis from other, related diagnoses.
4. The claim that mental illness is a myth rests on a misunderstanding of the role of lesions in medical disorders.
5. Prevalent claims to the contrary, psychiatric diagnoses often achieve adequate levels of reliability and validity, and do not typically pigeonhole or stigmatize individuals when correctly applied.
6. There is no clear consensus on the correct definition of mental disorder, and some authors have suggested that the higher-order concept of mental disorder is intrinsically undefinable. Even if true, this should have no effect on the scientific investigation, assessment, or treatment of specific mental disorders (e.g., schizophrenia, panic disorder), which undeniably exist.
I S S U E S  I N  D I A G N O S I S

7. Early versions of the diagnostic manual (*DSM-I* and *DSM-II*) were problematic because they provided clinicians and researchers with minimal guidance for establishing diagnoses and required high levels of subjective judgment and clinical inference.

8. *DSM-III*, which appeared in 1980, helped to alleviate this problem by providing diagnosticians with explicit diagnostic criteria, algorithms (decision rules), and hierarchical exclusion criteria, leading to increases in the reliability of many psychiatric diagnoses.

9. The current classification system, *DSM-5*, is a clear advance over *DSM-I* and *DSM-II*. Nevertheless, from initial reports, *DSM-5* continues to be plagued by a variety of problems, especially extensive comorbidity, reliable diagnoses that are nevertheless of questionable validity, and retention of a categorical model in the absence of compelling scientific evidence.

10. Fruitful potential directions for psychiatric classification include a dimensional model of personality to replace or supplement the existing categorical system of personality disorders and the adoption of endophenotypic markers for diagnostic purposes.
References


Cooper, R. (2011). Mental health and disorder. In H. T. Have, R. Chadwick, & E. M. Meslin (Eds.), *The


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 T. Millon & G. L. Klerman (Eds.), *Contemporary directions in psychopathology:*
REFERENCES


Wakefield, J. C. (1998). The DSM’s theory-neutral nosology is scientifically progressive: Response to


