The Cycle of Classification: DSM-I Through DSM-5

Roger K. Blashfield,[1] Jared W. Keeley,[2]
Elizabeth H. Flanagan,[3] and Shannon R. Miles[4]

[1]995 Eby Road, Hood River, Oregon 97031; email: blashrk@auburn.edu

[2]Department of Psychology, Mississippi State University, Mississippi State, Mississippi 39762;
email: jkeeley@psychology.msstate.edu

[3]Department of Psychiatry, Yale University School of Medicine, New Haven,
Connecticut 06513; email: elizabeth.flanagan@yale.edu

[4]Department of Psychology, University of Tulsa, Tulsa, Oklahoma 74104;
email: shannon-reynolds@utulsa.edu

Keywords

Abstract

The *Diagnostic and Statistical Manual of Mental Disorders* (DSM) was created
in 1952 by the American Psychiatric Association so that mental health pro-
fessionals in the United States would have a common language to use when
diagnosing individuals with mental disorders. Since the initial publication of
the DSM, there have been five subsequent editions of this manual published
(including the DSM-III-R). This review discusses the structural changes in
the six editions and the research that influenced those changes. Research
is classified into three domains: (*a*) issues related to the DSMs as measure-
ment systems, (*b*) studies of clinicians and how clinicians form diagnoses,
and (*c*) taxonomic issues involving the philosophy of science and metatheo-
retical ideas about how classification systems function. The review ends with
recommendations about future efforts to revise the DSMs.

Contents

THE CYCLE OF CLASSIFICATION: DSM-I THROUGH DSM-5

The classification of psychopathology is integral to the science and practice of clinical psychology as well as all behavioral health disciplines. In the United States, the *Diagnostic and Statistical Manual of Mental Disorders* (DSM; Am. Psychiatr. Assoc. 1952) has been the official American classification scheme since its inception in 1952. The fifth edition of this system was recently

released (DSM-5; Am. Psychiatr. Assoc. 2013a). Understanding the history of the DSM editions is important owing to their influence over diagnostic practice and research. The conceptual and methodological struggles of earlier editions of the DSM still apply, and thus the oft quoted piece of wisdom attributed to Edmund Burke pertains: "Those who do not know history are destined to repeat it." However, the entirety of those influences is too vast to cover in a single article. Thus, the current review focuses on the major environmental pressures that led to the creation of each edition of the DSM as well as the research between editions centering on the themes of measurement, clinicians' diagnostic practices, and the taxonomic underpinnings of the manual.

Prior to 1900, psychiatrists were few and far between and usually relegated to large state hospitals and asylums for the severely mentally ill. Psychoanalysis had not yet been created, and hardly any psychiatrists were engaged in outpatient psychotherapy (Grob 1991). Naturally, these psychiatrists were more interested in the pragmatic aspects of managing an asylum, and were less interested in academic pursuits. Thus, there was little interest in nosology (the branch of science dealing with the classification of disease) beyond how it would be practically useful in managing patients and performing administrative duties. In this context, as Grob (1991) notes, diagnosis was a primary concern for psychiatrists, but only insofar as it served a practical purpose. Psychiatrists were well aware of the problems in defining mental disorder categories, so the classification of mental disorders tended to be general and fluid. Classifications made on the basis of symptom descriptions led to much overlap of diagnostic categories, which often caused the diagnosis of one psychiatrist to be radically different from that of another. The problem of diagnostic agreement among clinicians would continue to plague psychiatric classification for years to come. Further, classifications on the basis of etiology of psychopathology were not possible, for theories of cause were speculative at best. However, psychiatrists of the time did value the role statistics could play in advancing the field. Statistics could shed light on prevalence rates, demographic patterns in mental illness, and disease course, and thus create a case for public policy and increased funding. But the collection of such statistical data requires categories. So, mental disorder classifications were considered a necessary evil and kept as simple as possible to limit any negative effects.

RESEARCH PRIOR TO DSM-I

A study from the 1930s that still has relevance to modern research was performed by a Catholic priest, Thomas V. Moore (1930, 1933). Moore gathered data on 367 psychotic patients from two mental institutions in the Washington, DC area. His descriptive data included 40 symptoms for which he provided prose definitions of what the symptoms meant, scores on cognitive ability tests, and behavior rating scales. By hand computation, Moore performed a factor analysis on the correlation matrix of these variables. Moore (1930) interpreted his results as yielding eight factors that he named: (*a*) cognitive defect, (*b*) catatonic syndrome, (*c*) uninhibited or kinetic syndrome, (*d*) non-euphoric manic syndrome, (*e*) euphoric manic syndrome, (*f*) delusional hallucinatory syndrome, (*g*) syndrome of constitutional hereditary depression, and (*h*) syndrome of retarded depression. These symptom groupings corresponded to similar diagnostic constructs common in inpatient psychiatry at the time and provided evidence in support of grouping symptoms into identifiable syndromes. Moore was ahead of his time when it came to thinking about psychopathology in terms of dimensions, and we begin with him as an example of psychiatry's struggle with how best to measure psychopathology.

However, the factors identified in his research were not particularly stable. Moore (1933) added some additional data, used a different method to calculate item intercorrelations, and generated a new factor analysis solution. The famous factor analytic researcher Thurstone (1934) published his own analysis of the Moore data, which yielded five factors. Two additional reanalyses of the Moore

data were published by Degan (1952) and by Blashfield (1984). Degan argued for nine factors, whereas Blashfield favored five factors similar to those reported by Thurstone. The specific factors are not as important as how these reanalyses highlight the fact that methodological choices and researcher preconceptions likely had a strong influence upon the resulting solution.

DSM-I

After World War II, American psychiatry was embarrassed by the chaotic state of classification in the United States. Four systems were in use across different sectors of the mental health field (Houts 2000). The American Psychiatric Association (APA) decided to overcome this "Tower of Babel" situation by creating a classification that would be acceptable to all members of its organization and that could unify the diagnostic terms of its psychiatrists. The result was the DSM (later renamed the DSM-I because it was the first edition in a series of substantive revisions to this manual).

The DSM-I contained 128 categories and was published as a smallish (132 pages) paperback book that cost $3. Organizationally, the DSM-I had a hierarchical system in which the initial node in the hierarchy was differentiating organic brain syndromes from "functional" disorders. The functional disorders were further subdivided into psychotic versus neurotic versus character disorders. This organization roughly followed the decision-making process of clinicians.

The DSM-I descriptions of disorders were prose paragraphs that incorporated behavioral and trait-like criteria; 93 of the 128 categories in this system had prose descriptions. These descriptions were very short, rarely over 200 words, and added little to what meaning could be derived from the name of the disorder. The terms in the description were relative and left to the interpretation of the clinician, leading to problems with reliability across professionals. An example of a DSM-I description for the diagnosis of psychophysiologic cardiovascular reaction follows. "This category includes such types of cardiovascular disorders as paroxysmal tachycardia, hypertension, vascular spasms, migraine, and so forth, in which emotional factors play a causative role" (Am. Psychiatr. Assoc. 1952, p. 30).

In retrospect, the DSM-I had an inpatient psychiatry focus. This edition focused mainly on the organic and psychotic disorders. The inpatient focus can also be seen by examining the miscellaneous categories in this system such as "Transient, hospitalized only for psychological testing" and "Deceased at the time of examination."

RESEARCH BETWEEN DSM-I AND DSM-II

Measurement

Factor analytic, descriptive research of psychopathology continued after the DSM-I was published. Like Moore's original research, these later investigators performed descriptive studies of inpatients in either state or V.A. hospitals. Wittenborn, Lorr, and Overall were important researchers in this tradition, each of whom conducted a number of studies attempting to empirically create the optimal descriptive representations of severe forms of psychopathology.

Another area of research that began during this time period focused on diagnostic reliability. The prototypical study in this tradition was performed by Phillip Ash (1949) as his Master's thesis in industrial psychology at Pennsylvania State University. He gathered data on three psychiatrists who independently interviewed and diagnosed 52 applicants to the Central Intelligence Agency (CIA). Ash was surprised by the large amount of variance among the psychiatrists in their diagnostic impressions of these different individuals. Ash concluded that psychiatric classification lacked adequate reliability when being used clinically to assign diagnoses.

Ash's study, particularly his provocative finding of inadequate reliability (interclinician agreement), stimulated a number of research studies that were performed in the 1950s through the 1970s. For detailed reviews of these studies, especially their methodological strengths and weaknesses, the interested reader should examine Matarazzo (1978) and Zubin (1967).

Research on Clinicians

Overall (1963) conducted a study on how clinicians viewed psychopathology. At the time, Overall was a well-known researcher using factor-analytic techniques to perform descriptive studies of the psychopathology of long-term psychiatric patients. He was also heavily involved in early studies on the effectiveness of the phenothiazines to treat schizophrenia. In his 1963 study, Overall used a scale he developed titled the Brief Psychiatric Rating Scale (BPRS), which rated the symptoms of these long-term patients on 16 dimensions. Overall asked 28 psychiatrists and 10 clinical psychologists to conceptualize "typical" patients for the 13 functional psychoses recognized in the DSM-I. They then rated these typical patients on the dimensions of the BPRS. A discriminant analysis formed dimensions that optimally separated the 13 categories. In the resulting three-dimensional space, which accounted for 83% of the variance, the diagnoses, as viewed by the clinicians, coalesced into visual clusters. Thus, despite using a dimensional methodology, Overall demonstrated the power of a categorical model for describing the way clinicians think about diagnoses.

Taxonomy

The World Health Organization (WHO) added a psychiatric section to its classification of medical disorders with the sixth version of the *International Statistical Classification of Diseases and Related Health Problems* (ICD-6). However, this psychiatric classification proved to be a political failure because it was ignored by almost every country in the United Nations at that time. A British psychiatrist named Stengel (1959) was asked to perform a thorough analysis of the psychiatric classifications that were used around the world. Stengel found that almost every country in the world had its own classification system, and some European countries had more than one. He was appalled at this multiplicity in diagnostic language. The DSM, from Stengel's perspective, provided a model of how the international community should proceed in trying to create a consensual system that would be adopted by every country in the world.

DSM-II

Stengel's review became a call for action. The WHO funded a series of international committee meetings in which countries around the world worked to create a consensual system. The result was the ICD-8. The American version of the ICD-8 was the DSM-II. Although the DSM-II and the ICD-8 were almost identical, a few differences did exist. The ICD-8 had a category for "hysterical psychosis," which Americans thought was an oxymoron because hysteria was clearly a neurotic disorder. Also Americans held onto a category that originated in military psychiatry called "passive-aggressive personality disorder." Europeans thought that this diagnosis was pejorative and disingenuous. Finally, the DSM-II, like the DSM-I, did have short prose definitions of its categories. The ICD-8 was only a list of approved diagnostic terms. No attempt was made to define the terms in the ICD-8.

The DSM-II had 193 diagnostic categories, of which 120 were defined using short prose presentations. Like the DSM-I, the DSM-II was a paperback manual consisting of 119 pages (costing $3.50) and had a hierarchical organization. Unlike the DSM-I, many of the new categories added in the DSM-II were categories of relevance to outpatient mental health efforts. Anxiety disorders,

depressive disorders, personality disorders (PDs), and disorders of childhood/adolescence were larger subsets than they had been in the DSM-I.

The other interesting change associated with the DSM-II concerned the miscellaneous categories. Because the ICD-8 was trying to create an international system that could be used around the world, the ICD-8 had diagnostic code numbers associated with each category. These code numbers, written in a decimal format, were constant, regardless of how a category was named within a particular language. There were two major types of miscellaneous categories in the ICD-8. The first were category codenames that ended in ".8." The "8" after the decimal point meant that the category was a unique diagnosis for that country. This coding solution allowed different countries to keep diagnostic categories that were commonly used in those countries even though these categories were not part of the consensual international system. The second type of miscellaneous category were those that ended in ".9." These were "wastebasket categories" that included all patients who fit within a particular family of mental disorders but who did not match the definition of any of the specific categories within that family (e.g., the patient had a PD but did not have any of the seven specific PDs listed in the ICD-8). Today, this miscellaneous category is called "not otherwise specified" (NOS). Of note, both kinds of categories are still used in the ICD.

RESEARCH BETWEEN THE DSM-II AND DSM-III

Taxonomy

In 1973, Rosenhan published a provocative paper in *Science* about how a group of colleagues went to different inpatient facilities in the United States requesting admission. They were truthful about themselves during the intake interview except for two things: (*a*) they gave fictitious names so that their admissions would not appear on their future medical records, and (*b*) they reported hearing a voice saying "Empty" or "Thud." All were admitted with a diagnosis of schizophrenia. Their average length of stay in the inpatient facility was nineteen days (the total range was 7 to 52 days). When discharged, most of them were given a diagnosis of "schizophrenia, in remission." Rosenhan and his colleagues noted that most of the patients in the facilities spotted that they were fakes, but none of the pseudopatients were detected by the hospital staff. Rosenhan concluded that inpatient facilities of the time could not differentiate the sane from the insane.

Rosenhan's paper stirred up a firestorm of protest. Robert Spitzer, who became the head of the DSM-III (Am. Psychiatr. Assoc. 1968), wrote a detailed and scathing methodological critique of Rosenhan's study. An entire issue of the *Journal of Abnormal Psychology* was devoted to analyses of the Rosenhan study along with a response by Rosenhan. The debates stimulated by the Rosenhan paper touched on central issues to taxonomy such as how a mental disorder is or is not defined, what methodologies in studying mental disorders are considered scientific versus pseudoscience, and how the validity of particular diagnoses can be ascertained. The Rosenhan paper stimulated strong emotional responses within the field that have persisted even into the contemporary literature (Slater 2004).

Research on Clinicians

Overall & Woodward (1975) performed a follow-up study to Overall (1963). They had psychiatrists assign ratings to clusters of psychopathology and compared these ratings to data from 2,000 actual American and French patients. The descriptions the clinicians gave to psychopathology differed substantially from the patient ratings. These findings suggested that the conceptualizations held in psychiatrists' minds about different diagnostic categories were quite different from the way in which psychopathology actually appeared during patient assessments.

Another important paper that proved to be a stimulus for the changes in the DSM-III was the US/UK international diagnostic project (Kendell et al. 1971). A set of eight videotapes of patients from the United States and Great Britain were shown to groups of American and British psychiatrists. For all eight videotapes, the modal diagnosis by the American clinicians was schizophrenia. In contrast, some of the videotapes, in the opinion of the British psychiatrists, represented patients with manic-depressive disorders, schizophrenia, and personality disorders. In conjunction with Rosenhan (1973), the US/UK project results were considered as further evidence that Americans were sloppy diagnosticians who tended to be over inclusive in their use of schizophrenia as a diagnosis.

Measurement

Aaron Beck was a psychiatrist who studied depressive disorders and who became famous for his advocacy of cognitive-behavioral approaches to the treatment of depression. Relatively early in his research career, Beck and colleagues (1962) performed a study on the reliability of psychiatric diagnosis using outpatients who were being considered for a research trial with his cognitive-behavioral therapy. Beck and his colleagues had the clinicians independently diagnose these outpatients and give reasons why they disagreed. Interestingly, clinicians most frequently felt that the fault for the diagnostic disagreements were the overly broad and nonspecific diagnostic definitions that existed in the DSM-II. A number of subsequent research papers cited Beck et al. (1962) as evidence for a need to change the definitions in the DSM-II.

In 1972, a group of psychiatrists at Washington University in St. Louis proposed a change that they hoped would lead to improvements in diagnostic reliability and also increase the specificity in meaning of contemporary diagnostic concepts (Feighner et al. 1972). In this paper, the Washington University group argued that there existed 15 mental disorder categories having sufficient evidence to assert that these categories were valid. Included in the fifteen were schizophrenia, manic-depressive disorder, homosexuality, and hysteria. These psychiatrists then proposed relatively specific, operational definitions for these categories in the form of diagnostic criteria. They argued that any future research on these fifteen mental disorders should use these diagnostic criteria to identify patient samples. The Feighner et al. (1972) study became the most highly cited paper in the psychiatric research literature of the time.

DSM-III

The classification that resulted from the research described above proved to be a truly revolutionary system. The earlier DSMs used short, broadly worded prose definitions to describe categories. The DSM-III, built on the innovation of the Feighner et al. paper, contained diagnostic criteria to specify the meaning of the categories. In addition, for each category, there was a description of the typical demographic profile of patients experiencing this disorder, a lengthy prose explanation of what the category meant, a description of how to differentiate the target category from any other category with which it might be confused, and a brief discussion of what was known, if anything, about the course and onset of the disorder.

Another innovation to the DSM-III was that the system was multiaxial. Each patient was expected to be diagnosed along five separate axes: (*a*) the descriptive presentation of the patient (i.e., the mental disorder categories), (*b*) the underlying personality and/or intellectual disorder, (*c*) any associated medical disorder that was relevant to the patient's psychiatric presentation, (*d*) the psychosocial stressors in the patient's environment, and (*e*) the patient's highest level of adaptive functioning in the past year. Finally, the DSM-III contained an extensive set of supplementary

Table 1 Description of the editions of the *Diagnostic and Statistical Manual of Mental Disorders*

Edition	Publication date	Number of pages	Number of diagnoses	Revenue for the American Psychiatric Association
DSM-I	1952	132	128	Unknown
DSM-II	1968	119	193	$1.27 million
DSM-III	1980	494	228	$9.33 million
DSM-III-R	1987	567	253	$16.65 million
DSM-IV	1994	886	383	$120 million
DSM-IV-TR	2000	943	383	Unknown
DSM-5	2013	947	541	Unknown

materials (e.g., a diagnostic flowchart) that could be useful to clinicians and to public health officials (e.g., tables showing how the DSM-III categories matched with ICD-8 categories).

There were 228 categories of mental disorders in the DSM-III (163 categories defined using diagnostic criteria) discussed in 494 pages, making the size of the DSM-III much larger than either the DSM-I or DSM-II. The price of the DSM-III increased ninefold ($31.75). As shown in **Table 1**, the revenue generated by the DSM-III also demonstrated a sizeable increase. The higher-order hierarchical system of the DSM-I and DSM-II was dropped. Instead, the categories, some of which were not recognized by the international community, were organized into 19 families of disorders. Examples of exclusively American categories were borderline PD, brief reactive psychosis, and psychogenic pain disorder. Consistent with the ICD, at the end of each chapter on a family of disorders, there were miscellaneous categories for patients whose symptoms met some of the diagnostic criteria in this family but not in a sufficient number to obtain a specific diagnosis. These individuals were given the family disorder diagnosis with an additional specifier of "not otherwise specified."

The explanation for the revolutionary nature of the DSM-III extends far beyond the confines of what a classification does and is. Publishing the DSM-III was part of a paradigm shift in psychiatry (and the mental health field in general). Prior to the DSM-III, psychiatry was dominated by psychoanalytically trained psychiatrists who eschewed the ideas of Kraepelin. These psychoanalysts saw little value to clinical diagnosis for working with psychotherapy patients. In contrast, the main authors of the DSM-III were the leaders of a group that have become known as the neo-Kraepelinians (Compton & Guze 1995). Outcasts within American psychiatry during the 1950s and 1960s, these individuals took over the DSM-III. In doing so, the neo-Kraepelinians attempted to bring psychiatry back to its medical roots. The ideas of the neo-Kraepelinians also fit well with the transition in treatment focus from psychotherapy to the use of medications. Decker (2013) described in detail the internal struggles within American psychiatry that were associated with the birth of the DSM-III. In effect, the neo-Kraepelinians, by creating the DSM-III, changed the entire focus of the mental health field.

RESEARCH BETWEEN THE DSM-III AND THE DSM-III-R

Measurement

Prior to creating the DSM-III, the leader of that innovative classification system, Robert Spitzer, had formed a partnership with the Washington University psychiatrists to create a broader classification system that would serve as the precursor to the DSM-III. This pre-DSM-III was called the Research Diagnostic Criteria (RDC) and focused on categories of psychotic and depressive

disorders (Spitzer et al. 1975). As the DSM-III was being created, early drafts of the RDC classification were made available. The result was a large number of research studies that appeared almost immediately after the DSM-III was published in which there were empirical comparisons of diagnostic criteria from Feighner et al., the RDC, and the DSM-III. These empirical studies often pointedly demonstrated significant flaws in the diagnostic criteria. This body of research stimulated Spitzer and others from the DSM-III to begin working on a revised version of the DSM-III called the DSM-III-R, which was intended to correct the diagnostic criteria.

In addition to the studies analyzing diagnostic criteria, another innovation followed the publication of the DSM-III: the creation of structured interviews. Spitzer, prior to becoming the leader of the DSM-III, had been performing a series of research studies with a psychologist named Jean Endicott to examine interview-based measures of psychopathology. When the RDC were published, Spitzer, Endicott, and Robins created a structured interview to assess the diagnostic criteria in the RDC. After the publication of the DSM-III, Spitzer and his colleagues created the SCID (Structured Clinical Interview for DSM-III-R; Spitzer et al. 1990). Other structured interviews that focused on more specific disorders were quickly developed, and by the year 2000 there were over 240 instruments to measure various aspects of psychopathology and mental disorders (Rush et al. 2000).

The reliability of diagnostic assessment using these new instruments generally was a distinct improvement over what had been found in the pre-DSM-III research. Spitzer & Fleiss's (1974) review of pre-DSM-III reliability research showed estimates of interclinician agreement typically ranging from 0.4 to 0.6. Using structured interviews like the SCID, reliability estimates were distinctly higher, typically in the range of 0.75 to 0.90. Because of the clearly defined method for assigning psychopathology, along with improved reliability, structured interviews would soon dominate the research world although even today they are rarely used in clinical practice.

Research on Clinicians

In the same year that the DSM-III was published, an innovative research study was published by Cantor et al. (1980). This research was proposed on the basis of a prototype model in cognitive psychology to explain how children learn categorical concepts. Cantor et al. (1980) had 13 mental health clinicians list the features that they associated with nine DSM-II diagnostic categories of psychosis. Any feature that was chosen by at least 3 of the 13 clinicians was kept for the final feature list. Then they took twelve case histories of patients that had been given one of four psychotic diagnoses (manic, depressed, paranoid schizophrenia, and undifferentiated schizophrenia). Four cases were considered to be quite prototypical of the four diagnoses (i.e., these four cases contain almost all of the features generated by the 13 clinicians), four were moderately prototypical, and four were not typical (i.e., these four cases had four or less of the defining features generated by the 13 clinicians). These case histories were given to the clinicians to diagnose. The reliability of the diagnoses varied as a function of the prototypicality of the case histories, with the least prototypical cases having the lowest reliability.

Research following the Cantor et al. (1980) study suggested that clinicians did not use diagnostic criteria to make diagnoses (Blashfield & Breen 1989, Morey & Ochoa 1989). Clinicians' diagnoses tended to follow a prototype-matching model rather than a criteria-based model (Clarkin et al. 1983, Horowitz et al. 1981, Livesley 1986).

Taxonomy

Alternative classificatory structures only began to appear around the time of the DSM-III. One stimulus for thinking about the organization of psychiatric classification was the development of

alternative theoretical approaches to biological classification. In particular, there was a movement in the 1960s and 1970s in which computerized systems were developed to create classificatory systems of living organisms. The school of thought associated with this approach was called numerical taxonomy, and the computerized methods that created these classifications were termed cluster analysis. Lorr and colleagues (1963; Lorr 1966) expanded these ideas in an attempt to empirically create classifications of psychotic patients. Other applications of cluster analysis attracted attention in the psychopathology literature. These efforts culminated the later work of Meehl to create "taxometric methods" to investigate the latent structure of psychopathology (Meehl 1995).

Another stimulus for taxonomic thought was that Spitzer and the organizing committee for the DSM-III took the bold step of proposing a tentative definition of the concept of mental disorder. They needed this definition because an explicit goal of the creators of the DSM-III was to avoid speculations about the causal mechanisms (especially theoretical concepts couched in psychoanalytic terms) that explained psychopathology. This definition was also in direct contrast to the antipsychiatry movement that attempted to define a mental disorder as society's way of dealing with undesirable people—by labeling them with a mental disorder to keep them quiet and segregated. The definition in the DSM-III was:

> Each of the mental disorders is conceptualized as a clinically significant behavioral or psychological syndrome or pattern that occurs in an individual and that is typically associated with either a painful symptom (distress) or impairment in one or more important areas of functioning (disability). In addition, there is an inference that there is a behavioral, psychological, or biological dysfunction, and that the disturbance is not only in the relationship between the individual and the society. (Am. Psychiatr. Assoc. 1980, p. 6)

As will become clear below, the DSM-III definition of mental disorder led to an interesting and growing discussion of psychiatric classification by philosophers, cognitive psychologists, social anthropologists, and historians.

DSM-III-R

Because of the research on diagnostic criteria that began appearing almost immediately after the DSM-III was published, the authors of the DSM-III decided to update those criteria. However, as often occurs when a committee is formed, the actions of the committee do not always match exactly the goals that were originally intended. The DSM-III-R (Am. Psychiatr. Assoc. 1987) did contain a number of changes to the diagnostic criteria. The DSM-III-R also contained new diagnostic categories that had not appeared in the DSM-III. Structurally, in terms of its multiaxial system, its use of diagnostic criteria, and its organization of the major families of mental disorders, the DSM-III-R was the same as the DSM-III. But, in terms of its specifics, the classification system changed substantially. The DSM-III-R was not just a revision; it was a new classification system.

The DSM-III-R contained a total of 253 categories (there had been 228 in the DSM-III). Of these, 174 were defined using diagnostic criteria (163 categories in the DSM-III had criteria). The biggest change was in the sleep disorders. The workgroup for this family of disorders had not been able to finish in time for the publication of the DSM-III. Thus, this section of the classification was new in the DSM-III-R. However, other changes did occur. For instance, "inhalant abuse," "inhalant dependence," and "inhalant addiction" were added to the DSM-III-R. "Schizoid disorder of childhood or adolescence" was dropped.

The political struggles concerning the DSM-III centered around a battle between a psycho-analytic faction of the APA and a biologically-oriented faction (the neo-Kraepelinians). When the DSM-III-R was being created, the focus of controversy shifted. Feminists were concerned with proposals by the DSM-III-R committees for new categories such as premenstrual syndrome and masochistic PD. As a result of the controversy, the DSM-III-R added a new appendix to its classifications called "Proposed diagnostic categories needing further study." Contained in this appendix were three categories: late luteal phase dysphoric disorder (the new name for premenstrual syndrome), sadistic PD (to balance masochistic PD), and self-defeating PD (the new name for masochistic PD). Politics had shifted the focus of the controversy from the psychoanalytic tradition to feminism.

RESEARCH BETWEEN THE DSM-III-R AND DSM-IV

Taxonomy

In an important series of publications, Wakefield (1992a,b, 1999) criticized the DSM-III approach to defining the concept of mental disorder. Wakefield argued that an important problem in the DSM-III definition was that it relied on the concept of "dysfunction" but that concept itself was undefined. The danger in defining dysfunction is in avoiding a tautology or circular definition. "A dysfunction occurs when something is not functioning." The trick, then, is to determine some independent means of identifying a function or dysfunction.

Wakefield decided, given his concerns about the vagueness of the DSM-III definition of mental disorder, to offer a definition of his own. Wakefield (1992a) offered two criteria, harm and dysfunction, which he considered necessary and jointly sufficient for the presence of a mental disorder. Wakefield's harmful-dysfunction definition, while not the only attempt to define mental disorder (see Lilienfeld & Marino 1995, 1999), is the most debated and examined in the literature.

The first part of his definition requires that the condition result in harm for the patient, which must be based on a value judgment. There is no such thing as a harmful condition in nature. Nature is neutral to the concerns of any individual, and so there are no justifiable bases to consider a condition harmful or beneficial without first explicitly taking a certain point of view. For the case of mental disorders, harm can come about in several ways: subjective distress, disability, or discomfort; a decrement in functioning in one or more socially defined roles; impeding upon the rights of others; etc. An alternative way of conceptualizing the harm criterion is that it represents a condition of therapeutic concern (Kirmayer & Young 1999, Lilienfeld & Marino 1995, Sadler 1999). In other words, the person, society, or treating clinician values the condition negatively.

The fact that a value judgment is necessary to instantiate any definition of mental disorder has a profound (yet often unsettling to many) implication. Because the definition of mental disorder—and by extension individual mental disorders—is fundamentally based upon a value judgment, it will always be a moving target. The same symptoms might be judged as disordered in one context but not in another. Further, as societal and individual values change over time, some conditions that used to be disordered will no longer be considered abnormal (e.g., homosexuality), and others that were not disordered might become problematic (e.g., Internet use). Thus, there can never be a "final" version of the DSM. Individual disorder definitions might call upon different values in order to be considered disordered. A single definition of mental disorder (as attempted by the DSM-III or Wakefield) contradicts the plurality of the myriad ways a condition might be defined as a disorder. Instead, mental disorder must always be a pluralistic concept, meaning that it has no single definition, but instead a variety of definitions depending upon the context (Ghaemi 2003).

The second component of Wakefield's definition of disorder is the idea of dysfunction. In contrast to the value-laden concept of harm, Wakefield proposed that a dysfunction could be scientifically and objectively determined through an understanding of natural functions via the effects they are "designed" to create, on the basis of evolutionary theory.

However, there are several problems with this definition of dysfunction. First, evolution is not a directed process, and so there is no design for a particular characteristic. Rather, the adaptive value of a characteristic is always a function of the environment in which the organism exists. As environments change, so do the selective pressures, and thus the adaptive value of the characteristic in question.

Further, the argument that evolution can be used to define mental functions may not hold, because mental functions may not be direct adaptations to the environment. Instead, they may be exaptations, or characteristics that evolved for some other purpose but currently serve a particular function (Gould 1991; Lilienfeld & Marino 1995, 1999). In short, Wakefield's essentialist approach to the interpretation of dysfunction has proven untenable when applied to a definition of mental disorder. Despite these problems, the definition of mental disorder has remained relatively unchanged in the DSM.

Research on Clinicians

In the mid-1980s, another group of research studies using mental health clinicians as participants were conducted with relative frequency—studies of gender bias in diagnosis. These studies were stimulated by two earlier papers. The findings of Broverman and colleagues (1970) were interpreted to mean that clinicians tend to pathologize women more relative to men. Warner (1978) used a standard case history that was presented with masculine pronouns to one group of clinicians and with feminine pronouns to another group. Warner (1978) interpreted his results as demonstrating bias by clinicians to use the diagnosis of histrionic personality with women. Both studies had serious methodological flaws. Nonetheless, the conclusions reached by these authors were readily accepted by a number of writers at the time.

A reasonably large literature on gender bias followed. This literature often focused on the PDs (Adler et al. 1990, Becker & Lamb 1994, Fernbach et al. 1989, Ford & Widiger 1989, Hamilton et al. 1986, Henry & Cohen 1983). Many of the findings were inconsistent across researchers, cases, and changes in methodology. The most consistent finding of these studies was that female versions of cases were rated as more histrionic. But even that finding had mixed results.

Reviews of the gender bias literature can be found in Garb (1998), and Widiger & Spitzer (1991). After these initial studies on gender bias in the diagnosis of PDs, the interest in this question decreased and the central issue moved from documenting gender bias to determining the source of differences in diagnosis based on gender: true prevalence differences in the population, clinician knowledge of population base rates, clinician application of societal stereotypes, etc. The various possible reasons for gender differences in diagnosis are very nicely outlined in Widiger (1998). Other careful analyses argue that research findings are not due to bias but instead reflect gender differences in personality traits underlying Axis II diagnoses (Paris 2004). Nonetheless, the mixed results from these various gender bias studies do not show conclusive evidence that clinicians are biased in their diagnosis of PDs.

DSM-IV

The DSM-II had been published in 1968 to coincide roughly with the publication of the ICD-8. The DSM-III was published in 1980; the DSM-III-R came out in 1987; the DSM-IV appeared in

1994 (Am. Psychiatr. Assoc. 1994). Zimmerman (1988), among others, was quite critical of this rate of change. Researchers needed stability in the definition of categories in order to perform useful studies of psychopathology. Clinicians, likewise, were confused by and had difficulty adjusting to changes in the fundamental terminology that organized the diagnostic process. Additionally, the rate of scientific discoveries did not support the rapid changes.

The APA chose a new leader for the DSM-IV, Allen Frances. He first created the 13 workgroups responsible for the various subsections of the DSM-IV. The composition of the workgroups was designed to contain some holdovers from the DSM-III/DSM-III-R, but generally the majority of each workgroup represented a new collection of professionals. The initial task of the workgroups was to perform careful literature reviews associated with the various diagnostic categories to which each workgroup was assigned. These literature reviews were intended to guide the decisions of the workgroups. In addition, the literature reviews served to identify databases that existed in the United States about different mental disorders. The researchers who had created these databases were asked if they would be willing to contribute their data to the DSM-IV workgroups who then used those data to make decisions about which diagnostic criteria needed alteration. The literature reviews were collected in a three-volume series of source books.

The DSM-IV grew to 383 categories. Of these, 201 diagnostic categories were defined using diagnostic criteria. The size of the DSM-IV also grew to 886 pages. The DSM-IV Source Books (three volumes) added 3,010 pages. Another area of considerable growth in the DSM-IV was the appendix for categories needing further study. The DSM-III-R had three. The DSM-IV had seventeen categories that needed further study of which one ("Medication-induced movement disorders") was further subdivided into seven subcategories. There also were three scales (a Defensive Functioning Scale, a Global Assessment of Relational Functioning Scale, and a Social and Occupational Functioning Assessment Scale) that were included in the to-be-studied appendix. A text revision of DSM-IV (DSM-IV-TR) was published in 2000. The intention was not to change any criteria, but a few changes did occur, such as dropping the clinical significance criterion for tic disorders and adjusting it to account for nonconsenting victims in the paraphilias (First & Pincus 2002). Otherwise, only the supporting narrative text was updated (Am. Psychiatr. Assoc. 2000). Interestingly, however, the DSM-IV-TR increased in length by 57 pages and the cost of purchase increased from $48.95 for the DSM-IV to $74.95 for the DSM-IV-TR.

RESEARCH BETWEEN THE DSM-IV AND DSM-5

Taxonomy

As noted above, the value-laden (i.e., harm) aspect of Wakefield's definition of mental disorder attracted relatively little criticism. Sadler (2005) expanded on the idea that values were relevant to psychiatric classification in a book titled *Values and Psychiatric Diagnosis*, which is one of the most important books published on the topic of psychiatric classification since the DSM-III. Sadler highlighted five values and the roles they play in psychiatric nosology: (*a*) aesthetics—how people prefer things to be, in the sense that they "like" or "appreciate" them; (*b*) epistemology—choices about how we know what we know about classification (i.e., what research methods we prefer); (*c*) ethics—what morals the classification upholds; (*d*) ontology—what is the fundamental nature of "things," or in the case of psychiatry, what mental disorders are in a (meta)physical sense; and (*e*) pragmatics—how useful or user-friendly the classification might be. Until the publication of this book, most (although not all) of those assumptions were ignored or taken for granted. This book legitimized the imperative role philosophical discourse plays in the development of a classification of mental disorders.

One example of this sort of work is by Haslam (2002), who discussed a classification of types relevant to psychiatric classification. The inherent idea is that individual mental disorders can be defined in a variety of ways; a taxonomy of "kinds of kinds." These kinds vary between true dimensions on one side and true categories on the other, with a variety of intermediary steps. True dimensions represent a continuous range of values, much like the personality trait "neuroticism." One can create a cut-off point on the dimension to call one tail neurotic and the other normal, but to do so is to create groups that do not reflect the properties of the underlying measured characteristic. Practical kinds are also continua, but there is a pragmatic basis for defining a cut-off point (Zachar 2000), like an IQ score of 70 for defining intellectual disability. A practical kind is defined via some value judgment. Fuzzy kinds, however, do present evidence of a natural demarcation between groups, like an overlapping, bimodal distribution. Although there is evidence that there are separate groups, these groups may overlap to some degree. Many DSM prototype-style categories are examples of fuzzy kinds owing to their features overlapping with neighboring diagnoses. The final two kinds in Haslam's hierarchy represent true categories, in which any overlap between groups is due to nothing more than chance (i.e., it is not a meaningful reflection of measurement characteristics). The difference between a discrete kind and a natural kind is that natural kinds have a definable essence that causes the separation into groups. Discrete kinds happen to be separate without a clearly defined etiology. Although there are strong arguments against ever finding a mental disorder that would qualify as a natural kind (e.g., Zachar 2000), Haslam wants to reserve the possibility that such a condition could be discovered.

A variety of statistical techniques including taxometrics (Meehl 1995, Ruscio et al. 2006) and factor mixture analysis (Muthen & Muthen 2010) have been developed to examine the dimensionality versus taxonicity of diagnostic constructs. Many taxometric studies find that dimensions tend to be favored for a variety of psychiatric disorders (Haslam et al. 2012), especially the PDs (Eaton et al. 2011, Wright et al. 2013), but other studies find justification in the measurement characteristics for retaining groups (Bernstein et al. 2010, Lenzenweger et al. 2008, Picardi et al. 2012).

The point of Haslam's taxonomy is that mental disorder categories might be a variety of kinds. There need not be a unified definition of mental disorder. Other areas of medicine have embraced this plurality in their diagnostic concepts, happily including practical kinds, like hypertension, next to natural kinds, like tuberculosis. The DSM-5 (Am. Psychiatr. Assoc. 2013a) authors incorporated an implicit flexible definition of mental disorder, including dimensional models for some disorders and retaining fuzzy prototypes for others (although their explicit definition of mental disorder remains remarkably similar to that in the DSM-III).

A major debate regarding the DSM-5 was how to incorporate dimensional models of personality into the manual. Research supports that normal and abnormal personality are not distinct from one another, and that personality is best represented on a continuum.

Markon et al. (2005) published an influential paper on the structure of personality which demonstrated that both abnormal and normal personalities (across multiple measures and samples) adhere to a hierarchy that begins with two factors and ends with five. The favored personality model depends upon which level of abstraction one examines. The two highest-order factors are often referred to as externalizing and internalizing, which are similar to Watson and colleagues' (1988) negative affectivity and positive affectivity. The three-factor solution is made of negative emotionality, disinhibition (similar to externalizing), and positive emotionality. The four-factor solution contains negative emotionality and positive emotionality and breaks disinhibition into disagreeable disinhibition and unconscientious disinhibition. Finally, the five-factor solution has negative emotionality, agreeableness (disagreeable disinhibition reversed), conscientiousness (unconscientious disinhibition reversed), extraversion, and openness (these last two factors break from

positive emotionality; Markon et al. 2005). The four- and five-factor solutions have generated the most personality models and appear to be the preferred levels of abstraction.

With extensive research supporting dimensional models of personality, the DSM-5 PD workgroup began the process of selecting a model to use in the new manual. Widiger & Simonsen (2005) compared and contrasted 18 candidate dimensional systems for the DSM-5, all of which were considered potential solutions. However, the DSM-5 PD workgroup chose to develop their own model, which was largely influenced by the Collaborative Longitudinal Personality Disorders Study (CLPS; Gunderson et al. 2000). The National Institute of Mental Health (NIMH) funded this study, and its goals were to examine the stability, course, and outcome of select PDs using a control group of clients with major depressive disorder and no PDs. The study only focused on four of the ten DSM-IV-TR PDs—avoidant, borderline, obsessive-compulsive, and schizotypal—chosen for "conceptual and logistical issues" (Gunderson et al. 2000, p. 303). Obsessive-compulsive personality disorder is a separate, discrete disorder from the other PDs; schizotypal personality disorder is a "prototype of a 'spectrum' disorder" (Gunderson et al. 2000, p. 304); borderline personality disorder is the most distressing and disabling PD; and avoidant personality disorder is the most prevalent of the PDs.

The expected changes for the DSM-5 personality chapter were released on the APA's website in February 2010 (Am. Psychiatr. Assoc. 2010). The original proposal was to replace all the PD categories with a dimensional model of personality. However, the political pressures against such a significant change were powerful, and the next version of the proposal (Am. Psychiatr. Assoc. 2011) involved a hybrid of categories and dimensions. Based on the 2011 proposal, six of the ten categorical diagnoses from the DSM-IV were kept: antisocial, avoidant, borderline, narcissistic, obsessive compulsive, and schizotypal. Each patient would have been rated on a dimensional scale of how much s/he resembled the specific PD. Additionally, five proposed domains, each with trait facets, would be rated on a 0–4 dimensional scale. The proposed domains were negative affectivity, detachment, antagonism, disinhibition, and psychoticism, roughly corresponding to the five-factor solution described above with the exception of psychoticism being conceptualized as qualitatively different than openness. For clients who did not meet criteria for the six categorical diagnoses, a diagnosis of "personality disorder–trait specified" would be given with the traits that describe the client listed.

The proposed changes strove to alleviate concerns about DSM-IV PDs, such as high rates of comorbidity between PDs, limited validity of some PDs, arbitrary diagnostic thresholds, and instability of PD diagnoses (Trull & Durrett 2005). The committee cited six CLPS articles (Bender et al. 2001; Grilo et al. 2004, 2005; Morey et al. 2012; Skodol et al. 2002, 2005) as empirical support for retaining the six categorical diagnoses and creating the new dimensional model (note that the website documenting their rationale has since been deactivated). As stated above, the CLPS only studied four DSM-IV PDs because of methodological reasons, not for empirical ones. No references were given for the exclusion of the remaining PDs, but the committee stated the remaining PDs were dimensional and related to one another.

The proposed changes to the DSM-5 quickly generated opposition. Widiger (2011) described four main concerns about the proposed changes. First, the DSM-5 disregards current literature by ignoring the bipolarity of personality because the proposed dimensions were unipolar; second, it does not include normal ranges of personality, which are not distinct from PDs; and third, the proposal neglects to cover many important aspects of personality even with 37 traits. Fourth, Widiger also took issue with the 2010 proposed changes of turning some of the current PDs, like histrionic and narcissistic PDs, into single traits. Although the DSM-5 PD committee struggled to come to a consensus about which dimensions and traits should be included in the new model, the NIMH continued to push a research agenda that examined the underlying systems that

contribute to psychopathology (e.g., genetics, epidemiology, cognitive psychology, sociology, etc.). The NIMH developed a plan for a new way of classifying psychopathology that could help identify the causal mechanisms underlying or cross-cutting current DSM syndromes. They termed their project the Research Domain Criteria (RDoC; Insel et al. 2010), which encompass multiple levels of analysis, ranging from the molecular to the social, with a particular emphasis on the neural circuitry of psychopathology primarily using genetics and neuroscience methodologies.

Research on Clinicians

After the publication of the DSM-IV, research on clinicians went beyond merely studying what kind of diagnostic decisions clinicians make to studying the mechanisms behind the diagnostic patterns. For instance, Flanagan and colleagues conducted a series of studies using folk taxonomic theory and methods of cognitive psychologists to study clinicians' views of the relationships between diagnostic categories (Flanagan & Blashfield 2006, 2007; Flanagan et al. 2008) and the hierarchical structure of clinicians' taxonomies (Flanagan et al. 2012). These studies found that clinicians' conceptualizations of the relationship among diagnostic categories is more similar to the DSM-I and DSM-II structure in which clinicians first differentiate between organic and nonorganic disorders, and then psychotic and neurotic disorders. In particular, clinicians did not differentiate Axis I and Axis II disorders into separate groups. These studies suggested a startling disconnect between how clinicians think of mental disorders and how they are organized in the current DSM, which could make the current DSM more difficult for clinicians to use than necessary.

The WHO has made a point to study how clinicians understand the relationships among mental disorder categories and to align the ICD-11 with clinicians' conceptualizations (Reed 2010, Reed et al. 2013, Roberts et al. 2012). The samples collected by these studies are broad, being conducted in as many as 64 countries with over 1,800 participants. The findings indicated that clinicians' conceptualizations of mental disorders are strikingly consistent, although not identical to either the DSM or ICD systems (Reed et al. 2013, Roberts et al. 2012).

Furthermore, clinicians' conceptualizations of comorbidity do not match the unstated structure of diagnostic manuals either. Specifically, clinicians follow a multiplicative model of comorbidity, as predicted by cognitive psychological theories, whereby the clinical picture of multiple disorders goes beyond the symptoms present in any one of the disorders (Keeley & Blashfield 2010, Keeley et al. 2013). The DSM and ICD implicitly follow an additive model in which a comorbid clinical picture is the simple addition of the features of the disorders involved.

Another interesting line of research on clinicians' conceptualizations after the DSM-IV was conducted by Woo-kyoung Ahn and colleagues. This research investigated clinicians' views of the underlying bases of mental disorders and how those affect their diagnostic decisions (Ahn et al. 2006, Kim & Ahn 2002). Kim & Ahn (2002) argued that clinicians use theories to represent mental disorders on the basis of how the symptoms are causally related rather than understanding mental disorders as lists of criteria. Clinicians were more likely to assign a disorder to a case vignette if the hypothetical patient displayed symptoms that were more central to the clinician's theory of the disorder. Clinicians were also more likely to remember symptoms from the case vignettes if they were more central to their theories. Later research found that clinicians' views of the biological and psychological bases of disorders are negatively related, such that clinicians believe medication to be more effective for biologically based mental disorders than psychotherapy (Ahn et al. 2009). This research group also investigated clinicians' beliefs about the essences of mental disorders (Ahn et al. 2006). Clinicians thought mental disorders had weaker essences than medical disorders and were hesitant to endorse that mental disorders were "real" or "natural." Expert clinicians were less likely to endorse that a mental disorder had a single, underlying cause.

This research was important as the DSM does not conceptualize disorders as being represented by causal theories, having a biological or psychological basis, or having an underlying essence, but Ahn and colleagues showed that clinicians think that way. Again, the disconnect between the DSM and clinicians' views could make the DSM more difficult for clinicians to use.

DSM-5

The DSM-5 development process began in 1999, and the Internet allowed the world to watch how the manual was built. Drafts of the DSM-5 were posted on the APA's website for people to make comments and suggestions (called the Prelude Project; Am. Psychiatr. Assoc. 2013b). The first draft was posted in February 2010, and in return, 8,000 comments were submitted. The second draft was released in 2011, and 2,000 additional comments were submitted. Although the Prelude Project allowed the DSM-5 to communicate with mental health professionals from around the world, it also opened the door for staunch opposition from anyone who took issue with the manual. In addition, Internet communication in the form of blogs and emails allowed criticism to coalesce in ways that the leaders of the DSM-5 had not anticipated.

David Kupfer was appointed chair of the DSM-5 and began the process with goals of creating a revolutionary manual in terms of matching the classification system to modern molecular biology, cognitive and affective neuroscience, and psychometrics. He assembled 13 workgroups that contained over 500 mental health professionals. Eleven academic medical centers participated in the field trials to test the reliability of the proposed categories.

In physical size, the DSM-5 grew to 947 pages. There were a total of 541 diagnostic categories, an increase of nearly 160 categories compared with the DSM-IV. However, the number of categories defined using diagnostic criteria dropped to 151 compared with the 201 categories in the DSM-IV with diagnostic criteria. Unlike the DSM-IV, there were no "Source Books" published to document the processes used by the DSM-5 workgroups when creating this classification. The rationales and reviews posted on the DSM-5 development website were taken down and cannot be retrieved. The cost of the DSM-5 more than doubled to $199 per hardback copy.

RESEARCH AFTER THE DSM-5

The DSM-5 was published in 2013, so the reactions of the field have yet to be seen. Because it is too soon to tell whether research will support its structure, validity, and reliability, we comment on the goals of the DSM-5 as specified by the APA prior to its publication (Am. Psychiatr. Assoc. 2013a, p. 2). Additionally, we provide a set of recommendations on how to proceed with classification research after the DSM-5.

Goal 1

The first goal of the DSM-5 was to decrease reliance on not otherwise specified (NOS) diagnoses by creating criteria with greater specificity. In most areas of psychopathology, NOS diagnoses are the most common ones given (Kupfer & Regier 2011). Ironically, increasing the specificity of criteria would have the opposite of the desired effect by increasing the number of cases that do not meet the more narrowly defined concept, thereby necessitating a NOS diagnosis. One of the interesting changes in the DSM-5 was that this edition actually had a distinct decrease in the number of diagnostic categories that were defined. This decrease is particularly striking because the number of categories with definitions had consistently increased across the editions. This

decrease reflects the conservative goals of the DSM-5 as this system tried to focus its efforts on disorders that either had the most valid evidence or that had substantial use by clinicians.

Conversely, the total number of diagnostic categories in this classification system increased markedly. Most of this increase was in the form of miscellaneous (administrative) categories, which covered a vast range of reasons why someone might be seen by a mental health professional (e.g., "spouse or partner abuse, psychological, confirmed, initial encounter," "overweight or obesity," "problems related to unwanted pregnancy"). Many of these added categories lacked any definition (other than what is implied by the name of the categories). If a definition was provided, the level of specificity was about the same as the category descriptions in the DSM-I (see above).

Goal 2

The DSM-5 intended to add dimensional measures of symptoms and severity (First 2010, Lopez et al. 2007, Narrow & Kuhl 2011, Regier 2007). Problems that are frequently encountered with categorical diagnoses include high rates of comorbidity, frequent use of NOS diagnoses, and boundary lines between disorders drawn on the basis of tradition rather than empirical data (Jones 2012, Widiger & Samuel 2005).

The debate to use dimensions or categories in classifying mental illness has been an issue since before the DSM-I (Moore 1930, as discussed above). In the clinical domain, dimensional assessments allow one to measure the severity of current symptoms, track the cost of treatment, and make probabilistic estimates of diagnoses on the basis of dimensional screeners (Kessler 2002, Watson 2005). Additionally, dimensional data may offer a more clinically useful picture by providing descriptive information that can be lost in a category that contains heterogeneous patients (Jones 2012). Researchers who test hypotheses have long preferred dimensional ratings because of the greater stability of dimensional scores (Watson 2005), increased statistical power, and decreased need for different cut-off points in diverse samples (Kessler 2002, Kraemer et al. 2004).

There are also scientific advances that support the use of dimensions. Few psychiatric diagnoses have been found to have "zones of rarity" (i.e., distinct boundaries between conditions; Kendell & Jablensky 2003, Kessler 2002). With the progression of large-scale family/adoption and genomewide association studies, disorders once thought to be separate have been shown to actually share underlying genetic similarities (Cross-Disord. Group Psychiatr. Genomics Consort. 2013). Additionally, entire categories of disorders (i.e., the anxiety disorders) likely stem from the same genetic variations but are cut apart into different diagnoses based on phenomenology and tradition (e.g., social phobia, panic disorder, and generalized anxiety disorder; Smoller et al. 2008). What is inherited is likely underlying brain circuits that influence the individual to respond to the environment in an anxious way, rather than a specific disorder. Although the leaders of the DSM-5 supported the move toward a more dimensional system, the internal controversies associated with the DSM-5 were intense around this dimensional versus categorical split. The DSM-5 proposal to create a hybrid categorical-dimensional system for the PDs was rejected by the Board of Trustees of the APA (Am. Psychiatr. Assoc. 2013b, Kupfer et al. 2002). The dimensional components of workgroup proposals (aside from the creation of diagnostic spectra) were included in Section III of the manual, indicating the controversial and unsettled views within the DSM-5.

Goal 3

The third goal was to better align the DSM with the WHO's ICD-11, which has an estimated release date of 2015. Although the ICD-11 has not yet been published, its proposed structure and disorders appear to be very similar to the DSM-5 (with a few exceptions). The development of the

ICD-11 and the DSM-5 progressed very much in parallel, with many individuals participating on corresponding workgroups for both manuals. It is important to note that the DSM is generally used for making diagnostic decisions in the United States; however, providers must still use an ICD diagnostic code number for insurance reimbursements. The United States has been using ICD-9 codes as its official reporting system since 1979 and will officially adopt ICD-10 on October 1, 2014. Sadly, ICD-10 codes correspond to DSM-IV diagnostic concepts. Correspondence between the DSM-5 and the ICD-10 is poor. When clinicians complain about a lack of compatibility between the DSM and the ICD (Remington 2012), it is because the United States is not up to date in the implementation of the ICD. Efforts to make the DSM-5 compatible with ICD-11 are moot if political issues and economic costs keep the United States from quickly updating to the ICD-11.

Goal 4

Researchers have argued that the DSMs have reified diagnoses and stunted research that examines the underlying etiology of mental disorders (Kupfer & Regier 2011). In order for the DSM-5 to reflect the most current scientific evidence available (Kupfer & Regier 2011), select criteria were changed. For example, DSM-IV posttraumatic stress disorder (PTSD) criterion A2 required an individual to feel fear, horror, or hopelessness when experiencing the trauma. This criterion was deleted in the DSM-5 because it showed poor agreement with those actually diagnosed with PTSD after a traumatic event (Pereda & Forero 2012).

Another goal related to updating the manual to reflect current scientific evidence was to address overarching issues of organization throughout the manual. The manual is separated into 22 categories/chapters of diagnoses. The APA has stated that the disorders are reorganized to reflect disorders with similar etiologies (Am. Psychiatr. Assoc. 2013a). For example, obsessive-compulsive disorder (OCD) was moved out of the anxiety disorders and into its own category of obsessive-compulsive and related disorders to reflect the underlying scientific evidence that it belongs on its own spectrum rather than with anxiety disorders (Stein et al. 2011).

Axis I, Axis II, and Axis III disorders are once again intertwined. Much of the work that supports this reorganization comes from personality research, which failed to find biological, psychometric, or psychological evidence for the separation of Axes I and II (Clark 2005, Krueger et al. 2011). In addition, research suggests that clinicians do not separate disorders according to Axis I and II when describing their conceptualizations of how mental disorders fit together (Flanagan & Blashfield 2006).

RECOMMENDATIONS FOR FUTURE DSMS

The DSM editions, in addition to growing in size, have grown in the extent of the controversies surrounding them. The issues with the DSM-III and DSM-III-R were the schism between the biological, proscience wing of the APA with the psychoanalytic, proclinical wing. For the DSM-IV, the strongest critics of that system were the feminists, who viewed many of the DSM changes as potentially destructive to women. With the DSM-5, controversy erupted over the potentially secretive process that was being used to make decisions as well as the corrupting influence of income and its potential revenues. More than once, important decisions that affected the final outcome of a DSM categorization system were made at the level of the Board of Trustees of the APA. The claim that the DSMs represent the best science possible is not credible when the final arbitrator of decision making is not a scientific body but the leaders of an organization with profit motives. For instance, one speculation about why the DSM-5 PD section ended up looking like it did was that adopting already existing dimensional models for this area of psychopathology

was not in the financial interest of the APA because copyrights for those existing measurement instruments were held by others. We offer the following recommendations to the APA for the publication of the next DSM.

Increase Transparency About Finances

We recommend that the APA make all financial records about the DSM, both past and present, publicly available. This would include information about royalties (if any) and honoraria paid to individuals involved with these editions and funding from pharmaceutical or other companies that could experience a financial impact because of DSM-influenced decisions, etc. Simply making this information available would go a long way towards alleviating concerns that financial decisions have driven the DSM process. The APA instituted a mandatory financial disclosure policy for the DSM-5 panel and workgroup members. However, some argue that openness about financial ties to industry does not limit bias, and many DSM-5 task force members reported financial ties to pharmaceutical companies (Cosgrove & Krimsky 2012).

Clarify the Goals of the DSM

The ICD comes in several versions. One is intended for clinicians and uses prototypes in its description of disorders [ICD-10-CM (clinical manual)], whereas the other is intended for researchers and uses operationalized diagnostic criteria [ICD-10-RDC (research diagnostic criteria)]. Thus, the ICD developers recognized that the optimal structure and nature of a classification system might be different for different purposes. The DSM, however, attempts to serve all masters simultaneously and may thereby diminish its own ability to do justice to any. The establishment of diagnostic criteria in the DSM-III made distinctions between individual categories clearer and increased diagnostic reliability. However, those criteria are too specific for clinicians and not specific enough for researchers. Experienced clinicians do not use criteria lists when making diagnostic decisions. Instead, they match new clients to ones they have seen previously using prototype matching (Blashfield & Breen 1989, Kim & Ahn 2002, Luhrmann 2000, Morey & Ochoa 1989). Only novice clinicians refer to criteria lists and present them as justification for a particular diagnosis. If clinicians in practice do not use diagnostic criteria, then their purpose seems limited to research. Perhaps prototypes would be a better way for clinicians to learn and communicate diagnostic categories once the categories are well defined with feature lists (i.e., criteria). We recommend that the DSM clarify its mission—is it designed to aid researchers in communally defining targets for study or to aid mental health practitioners in their work? These different values (Sadler 2005) appear to be in conflict.

Create a Research Agenda

Consider the NIMH's Research Domain Criteria program. Striking criticisms of the conceptualizations and measurements of psychopathology have been published in response to the DSM-5. Most shocking was the rejection of the DSM-5 by the NIMH. The director of the NIMH stated that the DSM-5 was a clinical tool rather than one that would assist researchers (Insel 2013), which is a further indictment that the DSM is not meeting its research goal. As described above, the RDoC are proposed to highlight underlying and cross-cutting conceptualizations of psychiatric symptoms, ideally to move the field toward a better understanding of etiology. The push is to move away from syndrome definitions of conditions, which currently are the best definitions available. Kupfer & Regier (2011) wrote a commentary about the compatibility of the RDoC and the new

DSM-5 and stated that science is not yet at the place that clinicians can order a blood test and diagnose a mental disorder. Although the authors of this paper would argue that the RDoC relies too heavily on neuroscience and needs to include more sophisticated input from other aspects of psychopathology (e.g., emotions, cognitions, personal experience of psychopathology), we also argue that the RDoC has a laudable goal of redefining how psychopathology is conceptualized, which may lead to better understanding of these conditions. However, this discussion again highlights the fact that serving a clinical goal of aiding practitioners in identifying, conceptualizing, and treating conditions may be (increasingly) separate from a productive research classification.

Use the revenue from the DSM-5 to fund a research agenda. Considerable criticism is levied for the amount of revenue the DSMs generate for the APA. These criticisms might be allayed somewhat if the revenue were used to improve the DSM. So far, most of the APA-sponsored research for the DSM focused on reliability, most prominently the field trials (Clarke et al. 2013). What the DSM lacks is improvements in the validity and clinical utility of the diagnoses. The DSMs' revenues could complement the money spent by the NIMH on RDoC by addressing questions not under the purview of the strictly etiological project. We also argue that more emphasis should be given to nonconventional research by investigators such as Peter Zachar and John Sadler to address the broadly impactful but largely ignored assumptions underlying the classifications.

Develop and fund a research agenda about how clinicians understand and use mental disorders. From at least as far back as Overall's study of clinicians, individual researchers have conducted studies of how clinicians make diagnoses. But rarely have grants been funded to extensively study clinicians' conceptualizations of psychopathology and diagnostic decisions. Nor has the APA made it a priority to study clinicians' use of the DSM. The WHO, in contrast, has supported a focused series of research to study clinicians' use of diagnostic categories. These studies have used sophisticated methodologies developed by cognitive psychologists to study how clinicians view the relationships among mental disorder categories. The rationale is that—barring empirical evidence on etiology or phenomenology about the relationships among disorder categories—classificatory decisions about the structure of the manual could be informed by how clinicians organize that information cognitively (Reed 2010). The APA would advance the field's understanding of psychopathology if it developed, funded, and supported a research agenda to study the clinical utility of the DSM categories and to make clinical utility an important consideration when the workgroups design categories. This development would require a paradigm shift for contemporary psychiatry, from largely focusing on neuroscience and the biological validity of diagnostic concepts to recognizing clinicians as worthwhile and informative targets of study.

Reduce Political Bias in the Development of the DSM

Given all of the claims of potential political bias in the DSM development process, it stands to reason that the authors of the next manual should continue to think of approaches to reduce actual or perceived bias in the DSM. Although the suggestions below are tentative, and nothing is likely to remove the problem entirely, we believe they are steps in the right direction.

Determine workgroups on the basis of a voting scheme using Internet-based nominations confirmed by the members of the APA. How the chair of the DSM revision is chosen has never been a transparent process. As described above, the APA has desired "fresh blood" with each version of the DSM and so different "great men" have been chosen to head the committees: the DSM-III/DSM-III-R by Spitzer, the DSM-IV by Frances, and the DSM-5 by Kupfer. Each

of these great men has chosen great (primarily) men to head the various DSM workgroups. They, in turn, chose other great (primarily) men to populate the workgroups. Blashfield & Reynolds (2012) documented how the workgroup members of the personality disorders committee were primarily members of the multimillion dollar CLPS grant funded by the NIMH or people who published papers together. Similar "invisible colleges" were evident among the neo-Kraepelinians who created the DSM-III (Blashfield 1984). A more transparent, inclusive, and democratic method for choosing workgroup members would weaken claims of bias in the DSM. In addition, social psychology has shown how political processes work better with an optimal amount of divergent opinions. Groups that are too homogenous are susceptible to groupthink.

Include consumers, advocates, family members, and policy makers as active, voting members on the committees and workgroups in large enough numbers (e.g., 50%) so they move beyond tokenism and have a significant influence. The Affordable Care Act is currently being implemented in the United States. Central to this act is that "stakeholders" (i.e., patients, family members, care givers) should be focal in the delivery and evaluation of medical services. A medical procedure's helpfulness will be decided on the basis of health indicators (i.e., improved health conditions as well as symptoms) as well as improvement in areas that are important to stakeholders such as employment, housing, functioning, quality of life, social support, religiosity, etc. The APA is antiquated in not having meaningful stakeholder involvement in the development and evaluation of the manual. As indicated above, the APA is quite proud of the stakeholder feedback culled for the DSM-5 through Internet comments. However, soliciting anonymous, out of context, comments from whoever chose to answer the surveys is woefully inadequate. It would be as if to say that all a hospital would have to do to ensure quality control is to solicit patient satisfaction questionnaires—they would not have to document how they use them to control quality and they would not have to include stakeholders on their boards and executive committees. It is the ultimate stigma that mental health stakeholders have been kept out of the development of the DSM. When people have disorders of the mind, they are considered unable to contribute meaningfully to quality control, but that kind of discrimination would not be tolerated in other areas of medicine.

Develop an independent oversight committee. Finally, we recommend that the Board of Trustees delegate all final decision making about future DSM editions to an independent committee. Some (a minority) of these committee members would be appointed by the APA, others would be appointed by other mental health associations, and still others would be appointed by stakeholder groups (e.g., the pharmaceutical industry, organizations of family members of mental health patients, the WHO). Once appointed, this committee would be the final decision-making body regarding any controversies and/or process issues (e.g., for resolving the debate between two alternative scientific views).

SUMMARY

Humans naturally attempt to sort and make sense of their environments, including how to classify psychopathology. Different methods are developed (taxonomies) in an attempt to most accurately represent reality. There are scientific pros and cons to each method, and politics play a role in deciding which methods are used. Throughout the histories of the DSMs, the researchers and clinicians have been struggling with similar issues, and these issues have not been resolved with the newly released DSM-5. We still do not know the etiology of mental disorders or when dimensions are better to use in classifying them as opposed to discrete categories. Theoretical positions oppose

one another, which is good for science in that it allows theories to be falsified (Popper 1985) but bad if money and self-serving biases influence the system more than the data.

The DSM-5 did not meet its revolutionary goals. Just like eager but well-intentioned politicians who take office and then realize the dynamics of Washington, the DSM-5 did not make the changes it said it would. It is working not only with science but also with people and politics. The themes will likely continue as psychiatry inches forward in attempting to understand psychopathology.

DISCLOSURE STATEMENT

LITERATURE CITED

Adler DA, Drake R, Teague G. 1990. Clinician's practices in personality assessment: Does gender influence the use of DSM-III Axis II? *Comp. Psychiatry* 33:125–33

Ahn W, Flanagan EH, Marsh JK, Sanislow CA. 2006. Beliefs about essences and the reality of mental disorders. *Psychol. Sci.* 17:759–66

Ahn W, Proctor C, Flanagan EH. 2009. Mental health clinicians' beliefs about the biological, psychological, and environmental bases of mental disorders. *Cogn. Sci.* 33:147–82

Am. Psychiatr. Assoc. 1952. *Diagnostic and Statistical Manual of Mental Disorders.* Washington, DC: Am. Psychiatr. Publ. 1st ed.

Am. Psychiatr. Assoc. 1968. *Diagnostic and Statistical Manual of Mental Disorders.* Washington, DC: Am. Psychiatr. Publ. 2nd ed.

Am. Psychiatr. Assoc. 1980. *Diagnostic and Statistical Manual of Mental Disorders.* Washington, DC: Am. Psychiatr. Publ. 3rd ed.

Am. Psychiatr. Assoc. 1987. *Diagnostic and Statistical Manual of Mental Disorders.* Washington, DC: Am. Psychiatr. Publ. 3rd ed., rev.

Am. Psychiatr. Assoc. 1994. *Diagnostic and Statistical Manual of Mental Disorders.* Washington, DC: Am. Psychiatr. Publ. 4th ed.

Am. Psychiatr. Assoc. 2000. *Diagnostic and Statistical Manual of Mental Disorders.* Washington, DC: Am. Psychiatr. Publ. 4th ed., text rev.

Am. Psychiatr. Assoc. 2010. *Brief Rationale and Status of the Development of a Trait Dimensional Diagnostic System for Personality Disorder in DSM-5.* Arlington, VA: Am. Psychiatr. Assoc.

Am. Psychiatr. Assoc. 2011. *Rationale for the Proposed Changes to the Personality Disorders Classification in DSM-5.* Arlington, VA: Am. Psychiatr. Assoc.

Am. Psychiatr. Assoc. 2013a. *Diagnostic and Statistical Manual of Mental Disorders.* Washington, DC: Am. Psychiatr. Publ. 5th ed.

Am. Psychiatr. Assoc. 2013b. *From Planning to Publication: Developing DSM-5.* **http://www.psych.org/ File%20Library/Practice/DSM/DSM-5/DSM-5-Development.pdf**

Ash P. 1949. The reliability of psychiatric diagnosis. *J. Abnorm. Soc. Psychol.* 44:272–76

Beck AT, Ward CH, Mendelson M, Mock JE, Erbaugh JK. 1962. Reliability of psychiatric diagnosis: II. A study of consistency of clinical judgments and ratings. *Am. J. Psychiatry* 119:351–57

Becker D, Lamb S. 1994. Sex bias in the diagnosis of borderline personality disorder and posttraumatic stress disorder. *Prof. Psychol. Res. Pract.* 25:55–61

Bender D, Dolan R, Skodol A, Sanislow C, Dyck I, et al. 2001. Treatment utilization by patients with personality disorders. *Am. J. Psychiatry* 158:295–302

Bernstein A, Stickle TR, Zvolensky MJ, Taylor S, Abramowitz J, Stewart S. 2010. Dimensional, categorical or dimensional-categories: testing the latent structure of anxiety sensitivity among adults using factor-mixture modeling. *Behav. Ther.* 41:515–29

Blashfield RK. 1984. *The Classification of Psychopathology.* New York: Plenum

Blashfield RK, Breen MJ. 1989. Face validity of the DSM-III-R personality disorders. *Am. J. Psychiatry* 146:1575–79

Blashfield RK, Reynolds SM. 2012. An invisible college view of the DSM-5 personality disorder classification. *J. Pers. Dis.* 26:821–29

Broverman IK, Broverman DM, Clarkson FE, Rosenkrantz PS, Vogel SR. 1970. Sex-role stereotypes and clinical judgments of mental health. *J. Consult. Clin. Psychol.* 34:1–7

Cantor N, Smith EE, French RD, Mezzich J. 1980. Psychiatric diagnosis as a prototype categorization. *J. Abnorm. Psychol.* 89:181–93

Clark LA. 2005. Temperament as a unifying basis for personality and psychopathology. *J. Abnorm. Psychol.* 114:505–21

Clarke DE, Narrow WE, Regier DA, Kuramoto SJ, Kupfer DJ, et al. 2013. DSM-5 field trials in the United States and Canada, part I: study design, sampling strategy, implementation, and analytic approaches. *Am. J. Psychiatry* 170:43–58

Clarkin JF, Widiger TA, Frances A, Hurt SW, Gilmore M. 1983. Prototypic typology and the borderline personality disorder. *J. Abnorm. Psychol.* 92:263–75

Compton W, Guze S. 1995. The neo-Kraepelinian revolution in psychiatric diagnosis. *Eur. Arch. Psychiatry Clin. Neurosci.* 245:196–201

Cosgrove L, Krimsky S. 2012. A comparison of DSM-IV and DSM-5 panel members' financial associations with industry: a pernicious problem persists. *PLoS Med.* 9(3):e1001190

Cross-Disord. Group Psychiatr. Genomics Consort. 2013. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 381:1371–79

Decker H. 2013. *The Making of the DSM-III: A Diagnostic Manual's Conquest of American Psychiatry*. New York: Oxford Univ. Press

Degan JW. 1952. Dimensions of functional psychosis. *Psychom. Mono.* #6. Richmond, VA: William Byrd

Eaton NR, Krueger RF, South SC, Simms LJ, Clark LA. 2011. Contrasting prototypes and dimensions in the classification of personality pathology: evidence that dimensions, but not prototypes, are robust. *Psychol. Med.* 41:1151–63

Feighner JP, Robins E, Guze SB, Woodruff TA, Winokur G, Munoz R. 1972. Diagnostic criteria for use in psychiatric research. *Arch. Gen. Psychiatry* 26:57–63

Fernbach BE, Winstead BA, Derlega VJ. 1989. Sex differences in diagnosis and treatment recommendations for antisocial personality and somatization disorders. *J. Soc. Clin. Psychol.* 8:238–55

First MB. 2010. Clinical utility in the revision of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM). *Prof. Psychol. Res. Pract.* 41:465–73

First MB, Pincus HA. 2002. The DSM-IV text revision: rationale and potential impact on clinical practice. *Psychiatr. Serv.* 53:288–92

Flanagan EH, Blashfield RK. 2006. Do clinicians see Axis I and Axis II as different kinds of disorders? *Comp. Psychiatry* 47:496–502

Flanagan EH, Blashfield RK. 2007. Clinicians' folk taxonomies of mental disorders. *Philos. Psychiatry Psychol.* 14:249–69

Flanagan EH, Keeley JW, Blashfield RK. 2008. An alternative hierarchical organization of the mental disorders of DSM-IV. *J. Abnorm. Psychol.* 117:693–98

Flanagan EH, Keeley JW, Blashfield RK. 2012. Do clinicians' taxonomies of mental disorders match the hierarchical structure of the DSM? *J. Clin. Psychol.* 68:620–30

Ford MR, Widiger TA. 1989. Sex bias in the diagnosis of histrionic and antisocial personality disorders. *J. Consult. Clin. Psychol.* 57:301–5

Garb HN. 1998. *Studying the Clinician: Judgment Research and Psychological Assessment*. Washington DC: Am. Psychol. Assoc.

Ghaemi SN. 2003. *The Concepts of Psychiatry: A Pluralistic Approach to the Mind and Mental Illness*. Baltimore, MD: Johns Hopkins Univ. Press

Gould S. 1991. Exaptation: a crucial tool for evolutionary analysis. *J. Soc. Issues* 47:43–65

Grilo CM, Sanislow CA, Gunderson JG, Pagano ME, Yen S, et al. 2004. Two-year stability and change of schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. *J. Consult. Clin. Psychol.* 72:767–75

Grilo CM, Sanislow CA, Shea MT, Skodol AE, Stout RL, et al. 2005. Two-year prospective naturalistic study of remission from major depressive disorder as a function of personality disorder co-morbidity. *J. Consult. Clin. Psychol.* 73:78–85

Grob G. 1991. Origins of DSM-I: a study in appearance and reality. *Am. J Psychiatry* 148:421–31

Gunderson JG, Shea MT, Skodol AE, McGlashan TH, Morey LC, et al. 2000. The collaborative longitudinal personality disorders study: development, aims, design, and sample characteristics. *J. Pers. Dis.* 14:300–15

Hamilton S, Rothbart M, Dawes R. 1986. Sex bias, diagnosis, and DSM-III. *Sex Roles* 15:269–74

Haslam N. 2002. Kinds of kinds: a conceptual taxonomy of psychiatric categories. *Philos. Psychiatry Psychol.* 9:203–17

Haslam N, Holland E, Kuppens P. 2012. Categories versus dimensions in personality and psychopathology: a quantitative review of taxometric research. *Psychol. Med.* 42:903–20

Henry K, Cohen C. 1983. The role of the labeling process in diagnosing borderline personality disorder. *Am. J. Psychiatry* 140:1527–29

Horowitz LM, Post DL, French RS, Wallis KD, Siegelman EY. 1981. The prototype as a construct in abnormal psychology: 2. Clarifying disagreement in psychiatric judgments. *J. Abnorm. Psychol.* 90:575–85

Houts A. 2000. Fifty years of psychiatric nomenclature: reflections on the 1943 War Department Technical Bulletin, Medical 203. *J. Clin. Psychol.* 56:935–67

Insel T. 2013. Transforming diagnosis. *My Blog, Tom Insel, M.D., NIMH Director*. June 24. **http://www.nimh. nih.gov/about/director/2013/transforming-diagnosis.shtml?utm_source=rss_readers&utm_ medium=rss&utm_campaign=rss_full**

Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, et al. 2010. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am. J. Psychiatry* 167:748–51

Jones K. 2012. Dimensional and cross-cutting assessment in the DSM-5. *J. Counsel. Dev.* 90:481–87

Keeley JW, Blashfield RK. 2010. Clinicians' conceptualizations of comorbid cases: a test of additive versus non-additive models. *J. Clin. Psychol.* 66:1121–30

Keeley JW, DeLao CS, Kirk CL. 2013. The commutative property in comorbid diagnosis: Does A + B = B + A? *Clin. Psychol. Sci.* 1:16–29

Kendell RE, Cooper JE, Gourlay AJ, Copeland JRM, Sharpe L, Gurland BJ. 1971. Diagnostic criteria of American and British psychiatrists. *Arch. Gen. Psychiatry* 25:123–30

Kendell R, Jablensky A. 2003. Distinguishing between the validity and the utility of psychiatric diagnosis. *Am. J. Psychiatry* 160:2–12

Kessler R. 2002. The categorical versus dimensional assessment controversy in the sociology of mental illness. *J. Health Soc. Behav.* 43:171–88

Kim NS, Ahn W. 2002. Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *J. Exp. Psychol.: Gen.* 131:451–76

Kirmayer L, Young A. 1999. Culture and context in the evolutionary concept of mental disorder. *J. Abnorm. Psychol.* 108:446–52

Kraemer H, Noda A, O'Hara R. 2004. Categorical versus dimensional approaches to diagnosis: methodological challenges. *J. Psychiatr. Res.* 38:17–25

Krueger RF, Eaton NR, Derringer J, Markon KE, Watson D, Skodol AE. 2011. Personality in DSM-5: helping delineate personality disorder content and framing the metastructure. *J. Pers. Assess.* 93:325–31

Kupfer DJ, First MD, Regier DA. 2002. *A Research Agenda for DSM-V*. Washington, DC: Am. Psychiatric Publ.

Kupfer DJ, Regier DA. 2011. Neuroscience, clinical evidence, and the future of psychiatric classification in DSM-5. *Am. J. Psychiatry* 168:672–74

Lenzenweger MF, Clarkin JF, Yeomans FE, Kernberg OF, Levy KN. 2008. Refining the borderline personality disorder phenotype through finite mixture modeling: implications for classification. *J. Pers. Dis.* 22:313–31

Lilienfeld SO, Marino L. 1995. Mental disorder as a Roschian concept: a critique of Wakefield's "harmful dysfunction" analysis. *J. Abnorm. Psychol.* 104:411–20

Lilienfeld SO, Marino L. 1999. Essentialism revisited: evolutionary theory and the concept of mental disorder. *J. Abnorm. Psychol.* 108:400–11

Livesley WJ. 1986. Trait and behavioral prototypes of personality disorder. *Am. J. Psychiatry* 143:728–32

Lopez M, Compton W, Grant B, Breiling J. 2007. Dimensional approaches in diagnostic classification: a critical appraisal. *Int. J. Methods Psychiatr. Res.* 16(Suppl. 1):S6–7

Lorr M, ed. 1966. *Explorations in Typing Psychotics*. New York: Pergamon

Lorr M, Klett CJ, McNair DM. 1963. *Syndromes of Psychosis*. New York: Pergamon

Luhrmann TM. 2000. *Of Two Minds: The Growing Disorder in American Psychiatry*. New York: Alfred A. Knopf

Markon KE, Krueger RF, Watson D. 2005. Delineating the structure of normal and abnormal personality: an integrative hierarchical approach. *J. Pers. Soc. Psychol.* 88:139–57

Matarazzo JD. 1978. The interview: its reliability and validity in psychiatric diagnosis. In *Clinical Diagnosis of Mental Disorders: A Handbook*, ed. BB Wolman, pp. 899–932. New York: Plenum

Meehl PE. 1995. Bootstraps taxometrics: solving the classification problem in psychopathology. *Am. Psychol.* 50:266–75

Moore TV. 1930. The empirical determination of certain syndromes underlying praecox and manic-depressive psychoses. *Am. J. Psychiatry* 86:719–38

Moore TV. 1933. The essential psychoses and their fundamental syndromes. *Stud. Psychol. Psychiatry* 3:1–28

Morey LC, Hopwood CJ, Markowitz JC, Gunderson JG, Grilo CM, et al. 2012. Comparison of alternative models for personality disorders, II: 6-, 8- and 10-year follow-up. *Psychol. Med.* 42:1705–13

Morey LC, Ochoa ES. 1989. An investigation of adherence to diagnostic criteria: clinical diagnosis of the DSM-III personality disorders. *J. Pers. Disord.* 3:180–92

Muthen LK, Muthen BO. 2010. *Mplus User's Guide*. Los Angeles, CA: Muthen & Muthen. 6th ed.

Narrow W, Kuhl E. 2011. Dimensional approaches to psychiatric diagnosis in DSM-5. *J. Mental Health Policy* 14:197–200

Overall JE. 1963. A configural analysis of psychiatric diagnostic stereotypes. *Behav. Sci.* 8:211–19

Overall JE, Woodward J. 1975. Conceptual validity of a phenomenological classification of psychiatric patients. *J. Psychiatr. Res.* 12:215–30

Paris J. 2004. Gender differences in personality traits and disorders. *Curr. Psychiatry Rep.* 6:71–74

Pereda N, Forero CG. 2012. Contribution of criterion A2 to PTSD screening in the presence of traumatic events. *J. Trauma. Stress* 25:587–91

Picardi A, Viroli C, Tarsitani L, Miglio R, de Girolamo G, et al. 2012. Heterogeneity and symptom structure of schizophrenia. *Psychiatry Res.* 198:386–94

Popper K. 1985. Falsificationism versus conventionalism. In *Popper Selections*, ed. D Miller. Princeton, NJ: Princeton Univ. Press. (Original work published 1934)

Reed GM. 2010. Toward ICD-11: improving the clinical utility of WHO's international classification of mental disorders. *Prof. Psychol. Res. Pr.* 41:457–64

Reed GM, Roberts M, Keeley J, Hoopell C, Matsumoto C, et al. 2013. Mental health professionals' natural taxonomies of mental disorders: implications for clinical utility of ICD-11 and DSM-5. *J. Clin. Psychol.* 69(12):1191–212

Regier D. 2007. Dimensional approaches to psychiatric classification: refining the research agenda for DSM-V: an introduction. *Int. J. Methods Psychiatr. Res.* 16 (Suppl. 1):S1–5

Remington R. 2012. Neurocognitive diagnostic challenges and the DSM-5: perspectives from the front lines of clinical practice. *Issues Mental Health Nurs.* 33:626–29

Roberts M, Medina-Mora M, Reed G, Keeley J, Sharan P, et al. 2012. A global clinicians' map of mental disorders to improve ICD-11. *Int. Rev. Psychiatry* 24:578–90

Rosenhan DL. 1973. On being sane in insane places. *Science* 179:250–58

Ruscio J, Haslam N, Ruscio AM. 2006. *Introduction to the Taxometric Method*. Mahwah, NJ: Erlbaum

Rush AJ, Pincus HA, First MB, Blacker D, Endicott J, et al., eds. 2000. *Handbook of Psychiatric Measures*. Washington, DC: Am. Psychiatr. Publ.

Sadler J. 1999. Horsefeathers: a commentary on "evolutionary versus prototype analyses of the concept of disorder." *J. Abnorm. Psychol.* 108:433–37

Sadler J. 2005. *Values and Psychiatric Diagnosis*. New York: Oxford Univ. Press

Skodol AE, Gunderson JG, McGlashan TH, Dyck IR, Stout RL, et al. 2002. Functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder. *Am. J. Psychiatry* 159:276–83

Skodol A, Pagano M, Bender D, Shea M, Gunderson J, et al. 2005. Stability of functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder over two years. *Psychol. Med.* 35:443–51

Slater L. 2004. *Opening Skinner's Box: Great Psychological Experiments of the Twentieth Century*. New York: W.W. Norton

Smoller J, Gardner-Schuster E, Misiaszek M. 2008. Genetics of anxiety: Would the genome recognize the DSM? *Depress. Anxiety* 25:368–77

Spitzer RL, Endicott J, Robins E. 1975. Clinical criteria for psychiatric diagnosis and DSM-III. *Am. J. Psychiatry* 132:1187–92

Spitzer RL, Fleiss JL. 1974. A re-analysis of the reliability of psychiatric diagnosis. *Br. J. Psychiatry* 125:341–47

Spitzer RL, Williams JBW, Gibbon M, First MB. 1990. *Structured Clinical Interview for DSM-III-R—Patient Edition*. Washington, DC: Am. Psychiatr. Publ.

Stein D, Craske M, Friedman M, Phillips K. 2011. Meta-structure issues for the DSM-5: How do anxiety disorders, obsessive-compulsive and related disorders, post-traumatic disorders, and dissociative disorders fit together? *Curr. Psychiatry Rep.* 13:248–50

Stengel E. 1959. Classification of mental disorders. *Bull. World Health Organ.* 21:601–63

Thurstone LL. 1934. The vectors of the mind. *Psychol. Rev.* 41:1–32

Trull TJ, Durrett C. 2005. Categorical and dimensional models of personality disorders. *Annu. Rev. Clin. Psychol.* 1:355–80

Wakefield JC. 1992a. The concept of mental disorder: on the boundary between biological facts and social values. *Am. Psychol.* 47:373–88

Wakefield JC. 1992b. Disorder as harmful dysfunction: a conceptual critique of DSM–III–R's definition of mental disorder. *Psychol. Rev.* 99:232–47

Wakefield JC. 1999. Evolutionary versus prototype analyses of the concept of disorder. *J. Abnorm. Psychol.* 108:374–99

Warner R. 1978. The diagnosis of antisocial and hysterical personality disorders. *J. Nerv. Mental Dis.* 166:839–45

Watson D. 2005. Rethinking the mood and anxiety disorders: a quantitative hierarchical model for DSM-V. *J. Abnorm. Psychol* . 114:522–36

Watson D, Clark LA, Tellegen A. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.* 54:1063–70

Widiger TA. 1998. Invited essay: sex biases in the diagnosis of personality disorders. *J. Pers. Disord.* 12:95–118

Widiger TA. 2011. A shaky future for personality disorders. *J. Pers. Disord.* 2:54–67

Widiger TA, Samuel DB. 2005. Diagnostic categories or dimensions? A question for the *Diagnostic and Statistical Manual of Mental Disorders*–fifth edition. *J. Abnorm. Psychol.* 114:494–504

Widiger TA, Simonsen E. 2005. Alternative dimensional models of personality disorder: finding a common ground. *J. Pers. Disord.* 19:110–30

Widiger TA. Spitzer RL. 1991. Sex bias in the diagnosis of personality disorders: conceptual and methodological issues. *Clin. Psychol. Rev.* 11:1–22

Wright AG, Krueger RF, Hobbs MJ, Markon KE, Eaton NR, Slade T. 2013. The structure of psychopathology: toward an expanded quantitative empirical model. *J. Abnorm. Psychol.* 122:281–94

Zachar P. 2000. Psychiatric disorders are not natural kinds. *Philos. Psychiatry Psychol.* 7:167–82

Zimmerman M. 1988. Why are we rushing to publish DSM-IV? *Arch. Gen. Psychiatry* 45:1135–38

Zubin J. 1967. The classification of behavior disorders. *Annu. Rev. Psychol.* 18:373–406

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Clinical Psychology* articles may be
found at http://www.annualreviews.org/errata/clinpsy

New From Annual Reviews:

Annual Review of Organizational Psychology and Organizational Behavior

Volume 1 • March 2014 • Online & In Print • http://orgpsych.annualreviews.org

Editor: **Frederick P. Morgeson,** *The Eli Broad College of Business, Michigan State University*

The *Annual Review of Organizational Psychology and Organizational Behavior* is devoted to publishing reviews of the industrial and organizational psychology, human resource management, and organizational behavior literature. Topics for review include motivation, selection, teams, training and development, leadership, job performance, strategic HR, cross-cultural issues, work attitudes, entrepreneurship, affect and emotion, organizational change and development, gender and diversity, statistics and research methodologies, and other emerging topics.

Complimentary online access to the first volume will be available until March 2015.

Access this and all other Annual Reviews journals via your institution at www.annualreviews.org.

ANNUAL REVIEWS

It's about time. Your time. It's time well spent.

New From Annual Reviews:

Annual Review of Statistics and Its Application

Volume 1 • Online January 2014 • http://statistics.annualreviews.org

Editor: **Stephen E. Fienberg,** *Carnegie Mellon University*

Associate Editors: **Nancy Reid,** *University of Toronto*

Stephen M. Stigler, *University of Chicago*

The *Annual Review of Statistics and Its Application* aims to inform statisticians and quantitative methodologists, as well as all scientists and users of statistics about major methodological advances and the computational tools that allow for their implementation. It will include developments in the field of statistics, including theoretical statistical underpinnings of new methodology, as well as developments in specific application domains such as biostatistics and bioinformatics, economics, machine learning, psychology, sociology, and aspects of the physical sciences.

Complimentary online access to the first volume will be available until January 2015.

Access this and all other Annual Reviews journals via your institution at www.annualreviews.org.