Journal of Abnormal Psychology
2005, Vol. 114, No. 4, 551–556

0021-843X/05/$12.00 DOI: 10.1037/0021-843X.114.4.551

Dimensional Versus Categorical Classification of Mental Disorders in the Fifth Edition of the *Diagnostic and Statistical Manual of Mental Disorders* and Beyond: Comment on the Special Section

Timothy A. Brown and David H. Barlow
Boston University

The value of including dimensional elements in the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*) has been recognized for decades. Nevertheless, no proposals have been made for introducing dimensional classification in the diagnostic system in a valid and feasible manner. As an initial step in this endeavor, the authors suggest introducing dimensional severity ratings to the extant diagnostic categories and criteria sets. Although not without difficulties, this would begin to determine the feasibility of dimensional classification and would address some limitations of the purely categorical approach (e.g., failure to capture individual differences in disorder severity, and clinically significant features subsumed by other disorders or falling below conventional *DSM* thresholds). The utility of incorporating broader dimensions of temperament and personality in diagnostic systems beyond the fifth edition of the *DSM* is also discussed.

Keywords: diagnostic classification of mental disorders, dimensional versus categorical assessment of psychopathology, risk factors for mental disorders, comorbidity of mental disorders, *Diagnostic and Statistical Manual of Mental Disorders*

The purpose of this special section is to promote the development of dimensional models of diagnostic classification and approaches to research on psychopathology. Although each article in this series highlights the importance of incorporating dimensional elements in the formal nosology (the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* [*DSM–V*] and beyond), little was said about how this might be accomplished in future editions of the *DSM* (but see Widiger, Costa, & McCrae, 2002; cited in Widiger & Samuel, 2005). This reflects a longstanding predicament. For nearly 30 years, researchers have acknowledged the potential utility of such systems (e.g., Barlow, 1988; Kendell, 1975; Maser & Cloninger, 1990; Widiger, 1992). Over this considerable time span, however, no strong proposals have emerged with regard to exactly how dimensional classification could be introduced in the *DSM*. Investigators involved in the preparation of *DSM–IV* considered and rejected the adoption of a dimensional classification in part because "there is yet no agreement on the choice of the optimal dimensions to be used for classification purposes" (American Psychiatric Association [APA], 1994, p. xxii). With the possible exception of the Axis II disorders (Widiger & Samuel, 2005), *DSM–V* may be bound for the same destiny.

As was the case over a decade ago when the *DSM–IV* was published, a tremendous amount of empirical and conceptual groundwork is needed to understand how dimensional classification can be validly and practically realized in the *DSM*. Core dimensions must be identified, measured, and validated. As most researchers would concur that the current categorical model of classification should not be abandoned entirely (e.g., Brown & Barlow, 2002; Krueger, Markon, Patrick, & Iacono, 2005; Widiger & Samuel, 2005), it must also be determined at what level dimensional elements are best incorporated into the diagnostic system (e.g., dimensional severity ratings for the existing *DSM* disorder constructs; dimensional assessment of higher order constructs, not currently recognized by the *DSM*, that reflect putative risk factors for families of disorders and account for their high rate of comorbidity; cf. externalization: antisocial behavior and substance use disorders, Krueger et al., 2005; trait negative affect–neuroticism: anxiety and mood disorders, Brown, Chorpita, & Barlow, 1998; Clark, 2005; Watson, 2005). An equally daunting challenge is the development of a dimensional assessment system that is widely agreed upon by *DSM* investigators and that can be practically and reliably implemented by both clinicians and researchers.

The limitations of a purely categorical approach to diagnostic classification are widely documented. For example, in our work with the anxiety and unipolar mood disorders, we have encountered many problems with *DSM*'s categorical diagnostic system (see Brown & Barlow, 2002, for a review). A diagnostic reliability study of the *DSM–IV* anxiety and mood disorders (Brown, Di Nardo, Lehman, & Campbell, 2001) found that for many categories (e.g., social phobia, obsessive–compulsive disorder [OCD]), diagnostic disagreements less often involved boundary issues with other formal disorders but were primarily due to problems in defining and applying a categorical threshold on the number, severity, or duration of symptoms. This threshold problem is manifested in various ways. It can be seen in diagnostic disagree-

Timothy A. Brown and David H. Barlow, Center for Anxiety and Related Disorders, Boston University.

Correspondence concerning this article should be addressed to Timothy A. Brown, Center for Anxiety and Related Disorders, Boston University, 648 Beacon Street, 6th Floor, Boston, MA 02215-2013. E-mail: tabrown@bu.edu

ments where both raters concur that the key features of a disorder are present but disagree as to whether these features cause sufficient interference or distress to satisfy the *DSM–IV* threshold for a clinical disorder (common with social phobia and specific phobia). The problem is also evident in the high rates of disagreements involving "not otherwise specified" (NOS) diagnoses (both raters agree on the presence of clinically significant features of the disorder, but one rater does not assign a formal anxiety or mood disorder diagnosis because of subthreshold patient report of the number or duration of symptoms; common with generalized anxiety disorder [GAD] and major depressive disorder [MDD]). A similar problem is at the root of diagnostic disagreements involving MDD versus dysthymia (core features of clinically significant depression are observed by both raters, but disagreement occurs with regard to the severity or duration of these symptoms). Finally, whereas dimensional ratings of the severity of MDD symptoms are reliable ($r = .74$), the *DSM–IV* categorical severity specifiers of this disorder are not (e.g., $\kappa = .30$; Brown, Di Nardo, et al., 2001). The same pattern of results was obtained for the severity specifiers of panic and agoraphobia in *DSM–III–R* (APA, 1987; despite favorable reliability at the dimensional level), a finding that contributed to the removal of these specifiers in the *DSM–IV* definitions of panic disorder and agoraphobia (Di Nardo, Moras, Barlow, Rapee, & Brown, 1993).

In addition to introducing measurement error (cf. MacCallum, Zhang, & Preacher, 2003), imposing categories on dimensional phenomena leads to a substantial loss of potentially valuable clinical information. As noted by Widiger and Samuel (2005), the *DSM* does not provide adequate coverage for clinically significant symptom presentations that fail to meet criteria for formal diagnostic categories (as reflected by the high rate in which NOS diagnoses are assigned as current and lifetime conditions; cf. Brown, Campbell, Lehman, Grisham, & Mancill, 2001). Moreover, the *DSM* does not provide a sufficient mechanism to record the severity of disorders (e.g., the severity of depression rather than the presence–absence of comorbid mood disorder per se may be more relevant to the prediction of the treatment outcome or natural course of a principal anxiety disorder). Salient information is also lost by adherence to the *DSM*'s elaborate set of hierarchical exclusions and differential diagnostic decision rules. These rules are in place to foster diagnostic reliability and to avoid excessive comorbidity (i.e., two separate categories should not be assigned if one diagnosis is deemed to be "due to" the coexisting disorder that occupies a higher position in the hierarchy). For example, *DSM–IV* criteria indicate that GAD should not be diagnosed if its features occur exclusively during the course of a mood disorder. Adherence to diagnostic rules of this nature leads to considerable information loss and misleading findings about the overlap of various disorders. In a large-scale comorbidity study of *DSM–IV* anxiety and mood disorders (Brown, Campbell, et al., 2001), the comorbidity rate of GAD and dysthymia was found to be a mere 5% when *DSM–IV* diagnostic rules were strictly followed. This result is strongly at odds with the wealth of evidence indicating substantial phenotypic and genetic overlap of GAD and the mood disorders (e.g., Brown et al., 1998; Kendler, Neale, Kessler, Heath, & Eaves, 1992; Starcevic, 1995). Indeed, when this diagnostic hierarchy rule was ignored, the GAD–dysthymia comorbidity estimate rose to 90%.

Artifacts of *DSM–IV* differential diagnostic rules are evident in many other comorbid patterns (Brown, Campbell, et al., 2001). For instance, the presence of panic disorder with agoraphobia (PDA) was associated with decreased relative risk of conditions such as social phobia and specific phobia. Rather than reflecting a true lack of association between these conditions (indeed, one would predict considerable phenotypic overlap of these disorders; e.g., situational avoidance; cf. Watson, 2005), such findings are a byproduct of *DSM–IV* differential diagnostic guidelines (i.e., features of social or specific fear and avoidance were often judged to be better accounted for and thereby subsumed under the PDA diagnosis). As is discussed further in this commentary, this reflects a key limitation of studies that rely on binary diagnostic indicators as the units of analysis (e.g., latent structural analyses reviewed by Krueger et al., 2005, and Watson, 2005). While curtailing undue comorbidity, *DSM*'s differential diagnostic guidelines also forfeit potentially salient clinical information. As previously noted, strict adherence to *DSM–IV* criteria does not acknowledge the common situation where clinically significant GAD co-occurs with a mood disorder or post-traumatic stress disorder (PTSD). Nonetheless, such symptoms are relevant to the overall severity of the clinical presentation and may have strong implications for treatment planning, untreated course, and so forth. Dimensional systems have the potential to more richly convey clinical presentations.

Unless the notion of having a separate classification system for research is entertained, it seems unreasonable to propose a purely dimensional *DSM*. Indeed, clinical utility is a compelling argument for retaining categorical distinctions in the nosology. The question then becomes how and at what level should dimensional elements be incorporated in the *DSM*. Should the nosology be predominately categorical (e.g., where dimensional ratings are added to the existing categories) or predominately dimensional (e.g., where *DSM* entails a multidimensional assessment of functioning, in which categorical diagnostic labels are subsequently imposed on the basis of quantitative algorithms; e.g., Widiger & Samuel, 2005)? The least drastic option would be to introduce dimensional severity ratings to the extant diagnostic categories and/or the constituent symptom criteria (akin to the methods in the Anxiety Disorders Interview Schedule for the *DSM–IV*; cf. Brown et al., 1998). This alternative would also be the most practical because the categorical system would remain intact and the dimensional rating system could be regarded as optional in settings where its implementation is less feasible (e.g., primary care). This strategy is reasonably straightforward and would address some of the chief complaints with the current system, such as its failure to convey disorder severity as well as other clinically significant features that are either subsumed by other disorders (e.g., GAD in mood disorders and PTSD) or fall just below conventional thresholds because of a *DSM* technicality (e.g., subclinical or NOS diagnoses where the clinical presentation is a symptom or two short of a formal disorder). Because dimensional ratings would simply be added to the current diagnostic categories, this approach would have several other advantages, including (a) its basis on a preexisting and widely studied set of constructs (i.e., the *DSM–IV* disorders; cf. APA, 1994, p. xxii) and (b) the ability to retain functional analytic and temporal (duration) aspects of diagnosis that are difficult to capture in a purely psychometric approach (see below). Moreover, adding severity ratings to existing categories would provide a standardized assessment system that fosters across-site comparability in the study of dimensional models of psychopathology. This approach might be a prudent first step that

would assist in determining the feasibility of more ambitious dimensional systems (e.g., quantifying higher order dimensions).

However, at this most basic level, a dimensional approach would not resolve many of the key problems in current classification, such as poor reliability and high comorbidity. For instance, "difference in patient report" (i.e., patient gives different information to independent interviewers in response to inquiries about the presence, severity, or duration of symptoms) is a very common source of diagnostic unreliability (Brown, Di Nardo, et al., 2001) that would be equally germane to dimensional clinical assessment. The fact that quantitative rating systems already present in the *DSM* (Axis V Global Assessment of Functioning [GAF] ratings) have been found to be rather unreliable may not bode well for an expanded dimensional classification system (e.g., Bates, Lyons, & Shaw, 2002; di Nardo et al., 1993). Similarly, because the various disorder categories would remain relatively unchanged, a dimensional system of this nature would not address the problems of unfavorable diagnostic reliability and high diagnostic comorbidity. This could be viewed as a limitation of the latent structural analyses reviewed in some articles in this series (e.g., Krueger et al., 2005; Watson, 2005). Although such analyses have implications on how the *DSM* might be reorganized (e.g., substance use and antisocial behavior disorders belong in the same chapter to better reflect their comorbidity and shared etiology; Krueger et al., 2005), the same number of disorders would exist to define these second-order dimensions (e.g., externalizing) and hence the same degree of diagnostic unreliability and comorbidity would also exist. If diagnostic classification was focused at a higher level, then such problems might be ameliorated.

Indeed, at the other extreme, it has been suggested that the *DSM* be driven by dimensions corresponding to broader biologically and environmentally based constructs of temperament and personality (e.g., neuroticism–negative affectivity; Clark, 2005). This proposal is in keeping with the theories and evidence that the overlap in families of disorders (e.g., comorbidity and symptom overlap in anxiety and mood disorders, externalizing disorders, etc.) can be accounted for by the fact that such disorders emerge from shared biologic–genetic and psychosocial diatheses (e.g., Andrews, 1996; Barlow, 2002; Brown, 1996; Kendler et al., 1992; Krueger et al., 2005). In this framework, the *DSM* disorders represent different manifestations of these core vulnerabilities; such variability stems from the influence of other, more specific etiologic agents (e.g., environmentally based psychological vulnerabilities, other genetic or biological influences). Of course, this is a far more challenging endeavor because broader behavioral phenotypes not currently recognized by the *DSM* must be identified, validated, and measured in a manner that is feasible in clinical practice and research.

However, there is mounting evidence attesting to the importance of attending to these broader dimensions. For instance, Brown, Antony, and Barlow (1995) examined the course of comorbid diagnoses in a sample of 126 patients who underwent cognitive–behavioral treatment for PDA. A significant pre- to posttreatment decline in overall comorbidity was noted (40% to 17%, respectively). At a 2-year follow-up, however, comorbidity had returned to a level (30%) that was no longer significantly different from pretreatment. This was the case despite the fact that, in the aggregate, patients maintained or improved upon gains for PDA across the follow-up interval, indicating considerable independence between panic disorder symptoms and overall comorbidity. Although speculative, these findings could be interpreted in accord with the notion that cognitive–behavioral treatment was generally effective in addressing the symptoms and maintaining processes of PDA but did not result in substantial reductions in general predispositional features (e.g., neuroticism) leaving patients vulnerable to the emergence or persistence of other disorders. In a more direct evaluation of this issue, Kasch, Rottenberg, Arnow, and Gotlib (2002) examined the temporal stability (8 months) and predictive utility of self-reported levels of behavioral inhibition and behavioral activation in 41 individuals with MDD (most of whom received treatment during the follow-up interval). Time 1 behavioral activation levels predicted poorer clinical outcome of MDD at the 8-month reassessment. Moreover, the behavioral inhibition and activation scales were remarkably stable over time and clinical state. In fact, whereas over a third of depressed participants were classified as no longer depressed at the 8-month follow-up, behavioral inhibition and activation scores displayed the same high level of temporal stability in this group (e.g., Time 1 and Time 2 behavioral inhibition Ms = 24.2 and 23.3, respectively) as in a subgroup of participants who were depressed at both assessment points (e.g., Time 1 and Time 2 behavioral inhibition Ms = 24.0 and 23.9, respectively). Such findings prompt concerns that current psychosocial treatments have become overly specialized because they focus on disorder-specific features (e.g., fear of panic in PDA), neglecting broader dimensions that are more germane to favorable long-term outcomes.

If higher order dimensions were incorporated, the *DSM* would no longer be a tertiary, descriptive instrument. Because these dimensions represent general vulnerabilities, the *DSM* would become highly relevant to the primary and secondary prevention of mental disorders. As has been the case for new diagnostic categories, the introduction of broader trait constructs in the *DSM* would result in a proliferation of empirical inquiry in this domain (e.g., development of interventions directly targeting these higher order features). As Krueger et al. (2005) noted, the illusory boundary between Axis I and Axis II would be obliterated. The question would arise as to the importance of lower order disorder constructs, in tandem with the caution that the future nosology should not become overly reductionistic. For example, although the covariance of social phobia and MDD might be fully explained by shared vulnerability dimensions (Brown et al., 1998), differentiation may still be important because these constructs provide more information about treatment planning, risk of complications (suicidality), and so forth than does knowing an individual's standing on the continuum of these less specific psychopathological dimensions. If such differentiation were retained, the issue of weak disorder construct boundaries would remain at the lower order level, although the *DSM* would possess a hierarchical framework to account for this overlap.[1]

Each of the articles in this special section have illustrated or proposed research strategies for developing dimensional models

[1] It might be argued the *DSM* should also incorporate more disorder-specific aspects of vulnerability. This endeavor may often be hindered by criteria contamination between lower order vulnerability dimensions and disorder constructs (e.g., anxiety sensitivity vs. "fear of fear" in panic disorder), although this problem also exists to some degree for higher order dimensions (e.g., GAD vs. neuroticism).

of classification and psychopathology. Accordingly, we conclude this commentary by offering a few additional research recommendations.

Researchers should avoid relying too strongly on extant psychometric studies to guide decisions about how disorders should be defined, organized, or subtyped. Although dimensional classification would render the *DSM* more psychometric in nature, this endeavor should not be driven solely by psychometrics. For instance, the strengths and limitations of using *DSM* disorders as the binary units of analysis in latent structural investigations of comorbidity should be noted (cf. Krueger et al., 2005; Watson, 2005).[2] One advantage of this approach is its ability to readily accommodate lifetime data (conversely, dimensional approaches are primarily cross-sectional in nature). The approach is also preferred when the intent is to examine the behavior of disorder categories as defined by the existing nosology (*DSM–IV*). For instance, analyses conducted at the diagnostic level have revealed rates and patterns of comorbidity that have suggested poor discriminant validity or a possible reorganization of some disorders (e.g., GAD is more closely aligned with unipolar mood disorders than with other anxiety disorders; Brown, Campbell, et al., 2001; Watson, 2005). Similarly, the studies reviewed by Krueger et al. (2005) have shown that some extant *DSM* diagnoses act as indicators of an underlying dimension (Externalization). These findings challenge the purely categorical orientation of *DSM*, a conclusion that is made all the more persuasive by the fact that *DSM* disorders served as the binary units of analysis (cf. dimensional approaches where latent factors, defined by dimensional symptom-based indicators, are created to serve as proxies for the *DSM* disorder constructs; cf. Brown et al., 1998).

However, in addition to the problems previously raised (e.g., increased measurement error, concealed individual differences in severity), a key limitation of this methodology is that by using *DSM* disorders as the units of analysis, the researcher is bound to the diagnostic system they are attempting to evaluate (Brown, 1996). Earlier in this commentary, we showed that the diagnostic system can produce spurious patterns and rates of disorder comorbidity and does not adequately recognize subthreshold clinically significant entities (Brown, Campbell, et al., 2001). Yet, these comorbidity tables are transformed into input matrices for the latent structural analyses. Accordingly, the resulting latent structure can be largely a function of the diagnostic system in place at the time (as well as a function of the scope of disorders included in the analysis). A more compelling strategy for examining the discriminant validity of disorder constructs (and their relations to higher order dimensions of vulnerability) would be to analyze dimensional indicators of disorder features assessed without consideration of *DSM* diagnostic rules.

However, a dimensional approach to latent structural analysis is also not without caveats. Cross-sectionally obtained self-report measures usually do not convey the functional relationships among disorder features. For example, although a questionnaire may contain an item to assess "situational avoidance of subways" with the intent of quantifying specific phobia (fear of enclosed places), this symptom may be secondary to a variety of other disorder processes (e.g., fear of unexpected panic in PDA, fear of contamination in OCD, perception of increased likelihood of social interaction in social phobia; avoidance of situational reminders of crime victimization in PTSD). Although associated symptoms are most susceptible (e.g., situational fear/avoidance in all anxiety disorders; nonspecific somatic symptoms in GAD, mood disorders, and PTSD), key features are not immune to this problem (e.g., fear of negative social evaluation secondary to PDA). Because these relationships are typically obfuscated by cross-sectional self-report measurement, misleading latent structures may occur (cf. dissociation of agoraphobia and panic disorder; Watson, 2005). Clinical ratings better capture such relationships but may be influenced by the extant nosology (e.g., *DSM* diagnostic decision rules). Monomethod, latent structural evaluations of *DSM* disorder criteria sets have other salient limitations. For example, such analyses have often suggested that the latent structure of PTSD is defined by four factors (e.g., Intrusions, Avoidance, Dysphoria, Hyperarousal; Simms, Watson, & Doebbling, 2002). However, some of these factors are defined by only two symptoms (e.g., Avoidance, Hyperarousal). In addition to the low stability and indeterminacy of factors defined by two indicators, it is difficult to determine whether such factors are substantively meaningful or stem from artifacts of the criteria set or measurement approach (e.g., method effects arising from similarly phrased criteria; cf. Criteria C1 and C2). In future research, we recommend that researchers consider the underutilized multitrait–multimethod measurement model approach, incorporating background and distal variables that serve as clinical validators of substantive dimensions (see Brown, in press; Marsh & Grayson, 1995). This framework would allow the researcher to build equally important aspects of risk and psychopathology (e.g., neurobiology–genetics, treatment response) into the psychometric endeavor.

If the *DSM* is to incorporate broader dimensions of risk, then future studies of latent structure should directly assess and model these dimensions. In most studies conducted to date (e.g., Krueger,

[2] Nonetheless, structural equation modeling (SEM; e.g., confirmatory factor analysis [CFA]) provides another useful modeling framework in situations where the researcher is working with categorical outcomes such as binary diagnoses (cf. latent trait analysis; Krueger et al., 2005). With proper statistical estimators and procedures (e.g., robust weighted least squares, maximum likelihood with numerical integration), structural equation modeling/confirmatory factor analysis [SEM/CFA] can readily accommodate categorical outcomes. If so desired, the measurement parameters (and standard errors) of an SEM/CFA solution (e.g., factor loadings, intercepts) can be easily converted into the parameters of a two-parameter logistic item response theory (IRT) model (i.e., item discrimination and location parameters; cf. latent trait models, Krueger et al., 2005) or item probabilities. In addition to providing a wider array of goodness-of-fit information relative to IRT, the SEM/CFA framework offers expanded modeling possibilities, such as the ability to (a) embed the measurement model into a larger structural solution (e.g., include background variables or distal outcomes represented by continuous or categorical single indicators, or latent variables, that serve as clinical validators of the dimensions revealed by the measurement model); (b) model direct effects of covariates on the latent factors (cf. population heterogeneity) in addition to direct effects of covariates on measurement model indicators (cf. differential item functioning); (c) easily evaluate multidimensional models (i.e., measurement models with two or more latent factors); and (d) incorporate an error theory (e.g., measurement error covariances). Interested readers are referred to the following references for more information and applied examples of these approaches: Brown (in press); MacIntosh and Hashim (2003); Meade and Lautenschlager (2004); Muthén (1988); Muthén and Asparouhov (2002).

1999), the *DSM* disorders have served as binary indicators of these broader risk dimensions (i.e., *DSM* disorders within a selected group are moderately correlated at roughly the same magnitude, a single factor of "risk" will adequately account for these relationships). This approach is not fully concordant with the notion that these dimensions act as vulnerability factors that predate clinical psychopathology (and seem to persist after successful treatment; Kasch et al., 2002). Direct measurement of vulnerability (e.g., behavioral manifestations of genetic or psychosocial risk) would lead to more persuasive evaluations of the role of these dimensions in explaining the covariance, etiology, and course of disorders (Brown et al., 1998) and would provide invaluable information about how (and whether) these elements should be incorporated into the *DSM*.

Finally, we concur with the concerns raised in two of the articles (Krueger et al., 2005; Widiger & Samuel, 2005) regarding the contribution of the taxometric method to informing the issue of categorical versus dimensional diagnostic classification. As noted by Krueger et al. (2005), the subjective aspect of taxometric analysis often promotes disagreements among experts as to whether taxometric plots support a categorical or dimensional structural inference. This method is also limited by its inability to focus on more than one diagnostic construct (boundary) at a time (cf. latent class and latent profile analysis, factor mixture models; Muthén, 2002). Taxometric requirements with regard to sampling and indicator selection are often quite restrictive. Unselected samples (e.g., college students) are less appropriate because of low base rates or sizes of putative taxa and positive skew of indicators (J. Ruscio, Ruscio, & Keane, 2004; non-normality is typically the case when clinical psychopathology measures are administered to nonclinical samples). Whereas clinical samples seem more appropriate, indicators of disorder constructs may be less apt to pass suitability tests in these samples (e.g., high nuisance covariance and poor validity due to the influence of general distress, comorbidity, *DSM* criteria overlap), especially if a single assessment modality is relied on (shared method variance). This problem may be particularly germane to self-report measures (e.g., questionnaires), which are likely to be more prone to the influence of general distress and less likely to account for the functional relationships among disorder features. However, salient concerns have also been raised with regard to clinical ratings indicators; namely, their suspected propensity to yield pseudotaxonicity stemming from rater bias (Beauchaine & Waters, 2003). As with any latent structural approach, the conclusion of taxonicity (presence of a latent class) should be supported by clinical validation (cf. Watson, 2003). Nonetheless, as noted by some investigators (e.g., Widiger & Samuel, 2005), strong evidence of taxonicity does not negate the practical and informational value of dimensional classification (e.g., more reliable measurement of individual differences in disorder severity among class members) or vice versa.

In summary, we are on the cusp of a second revolution in nosology. The first revolution clearly occurred in 1980, with the publication of *DSM–III* (APA, 1980), which aspired to an objective, theoretical, and very descriptive system of nosology. Following on the important earlier work of the Washington University group (Feighner et al., 1972), the *DSM–III* and its successors became, by far, the most widely used system of nosology for mental disorders in the world (Maser, Kaelber, & Weise, 1991). Now the *DSM–V* process has begun, and a clear consensus has emerged that we must move beyond description and back to a consideration of etiologic theory to achieve the second revolution in nosology (e.g., Charney et al., 2002; Kupfer, First, & Regier, 2002; Phillips, First, & Pincus, 2003), but this approach is so radically different that we are very unlikely to achieve consensus in time for the publication of the *DSM–V*. Witness, for example, the proposals for a pathophysiologically based classification system emerging from tremendous advances in the understanding of neuroscience over the past decade (e.g., Charney et al., 2002; Phillips et al., 2003), compared with the equally spectacular advances in our understanding of biologically based, but behaviorally anchored, temperaments emerging from the laboratories of experimental psychopathology as exemplified by the articles in this special series. It seems clear that these new and sophisticated spectrum approaches will ultimately win the day, but which one? And how will they be integrated? For *DSM–V*, researchers will likely settle for the beginnings of an integration of dimensional elements into the (prototypical) categorical system that is *DSM–IV* (as suggested above). At the same time, it seems crucially important at this stage to take the first steps toward mapping out a *DSM* driven by dimensions based on constructs of temperament, personality, and genetics that are much broader in scope than have been articulated to date. In other words, to drive the next stage of research with *DSM–VI* in mind, it is time to go beyond simply describing what the broad biological and behavioral dimensions may be (and how they may related) by actually articulating in some detail how such a system would look and how it would work. Descriptions of such a system from pathophysiological, behavioral, and temperamental standpoints, or some combination, could then begin to be evaluated for fundamental issues of validity, a process that is likely to take decades. This process will not start, however, until someone takes the plunge and proposes an initial hypothetical nosological system.

In the meantime, there is much work to be done to integrate dimensional ratings, including perhaps incorporating some indexes of behavioral temperaments into the current prototypical categorical system to provide a richer and more satisfying approach. This would be the first step, as noted above, toward assessing the feasibility of more ambitious nosological systems based on dimensions. Thus, the articles in this series advance us both toward *DSM–V* and, more importantly, toward *DSM–VI*.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

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

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andrews, G. (1996). Comorbidity in neurotic disorders: The similarities are more important than the differences. In R. M. Rapee (Ed.), *Current controversies in the anxiety disorders* (pp. 3–20). New York: Guilford Press.

Barlow, D. H. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic.* New York: Guilford Press.

Barlow, D. H. (2002). *Anxiety and its disorders: The nature and treatment of anxiety and panic* (2nd ed.). New York: Guilford Press.

Bates, L. W., Lyons, J. A., & Shaw, J. B. (2002). Effects of brief training on application of the global assessment of functioning scale. *Psychological Reports, 91,* 999–1006.

Beauchaine, T. P., & Waters, E. (2003). Pseudotaxonicity in MAMBAC and MAXCOV analyses of rating-scale data: Turning continua into classes by manipulating observers' expectations. *Psychological Methods, 8,* 3–15.

Brown, T. A. (1996). Validity of the *DSM–III–R* and *DSM–IV* classification systems for anxiety disorders. In R. M. Rapee (Ed.), *Current controversies in the anxiety disorders* (pp. 21–45). New York: Guilford Press.

Brown, T. A. (in press). *Confirmatory factor analysis for applied research.* New York: Guilford Press.

Brown, T. A., Antony, M. M., & Barlow, D. H. (1995). Diagnostic comorbidity in panic disorder: Effect on treatment outcome and course of comorbid diagnoses following treatment. *Journal of Consulting and Clinical Psychology, 63,* 408–418.

Brown, T. A., & Barlow, D. H. (2002). Classification of anxiety and mood disorders. In D. H. Barlow (Ed.), *Anxiety and its disorders: The nature and treatment of anxiety and panic* (2nd ed., pp. 292–327). New York: Guilford Press.

Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and lifetime comorbidity of the *DSM–IV* anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology, 110,* 585–599.

Brown, T. A., Chorpita, B. F., & Barlow, D. H. (1998). Structural relationships among dimensions of the *DSM–IV* anxiety and mood disorders and dimensions of negative affect, positive affect, and autonomic arousal. *Journal of Abnormal Psychology, 107,* 179–192.

Brown, T. A., Di Nardo, P. A., Lehman, C. L., & Campbell, L. A. (2001). Reliability of *DSM–IV* anxiety and mood disorders: Implications for the classification of emotional disorders. *Journal of Abnormal Psychology, 110,* 49–58.

Charney, D. S., Barlow, D. H., Botteron, K., Cohen, J. D., Goldman, D., Gur, R. E., et al. (2002). Neuroscience research agenda to guide development of a pathophysiologically based classification system. In D. J. Kupfer, M. B. First, & D. A. Regier (Eds.), *A research agenda for DSM–IV* (pp. 31–83). Washington, DC: American Psychiatric Association.

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

di Nardo, P. A., Moras, K., Barlow, D. H., Rapee, R. M., & Brown, T. A. (1993). Reliability of *DSM–III–R* anxiety disorder categories: Using the Anxiety Disorders Interview Schedule—Revised (ADIS-R). *Archives of General Psychiatry, 50,* 251–256.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry, 26,* 57–63.

Kasch, K. L., Rottenberg, J., Arnow, B. A., & Gotlib, I. H. (2002). Behavioral activation and inhibition systems and the severity and course of depression. *Journal of Abnormal Psychology, 111,* 589–597.

Kendell, R. E. (1975). *The role of diagnosis in psychiatry.* Oxford: Blackwell.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992). Major depression and generalized anxiety disorder: Same genes, (partly) different environments? *Archives of General Psychiatry, 49,* 716–722.

Krueger, R. F. (1999). The structure of common mental disorders. *Archives of General Psychiatry, 56,* 921–926.

Krueger, R. F., Markon, K. E., Patrick, C. J., & Iacono, W. G. (2005). Externalizing psychopathology in adulthood: A dimensional-spectrum conceptualization and its implications for *DSM–V. Journal of Abnormal Psychology, 114,* 537–550.

Kupfer, D. A., First, M. B., & Regier, D. A. (2002). *A research agenda for DSM-V.* Washington, DC: American Psychiatric Association.

MacCallum, R. C., Zhang, S., & Preacher, K. J. (2003). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7,* 19–40.

MacIntosh, R., & Hashim, S. (2003). Variance estimation for converting MIMIC model parameters to IRT parameters in DIF analysis. *Applied Psychological Measurement, 27,* 372–379.

Marsh, H. W., & Grayson, D. (1995). Latent variable models of multitrait–multimethod data. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 177–198). Thousand Oaks, CA: Sage.

Maser, J. D., & Cloninger, C. R. (Eds.). (1990). *Comorbidity of mood and anxiety disorders.* Washington, DC: American Psychiatric Press.

Maser, J. D., Kaelber, C., & Weise, R. E. (1991). International use and attitudes toward *DSM–III* and *DSM–III–R*: Growing consensus in psychiatric classification. *Journal of Abnormal Psychology, 100,* 271–279.

Meade, A. W., & Lautenschlager, G. J. (2004). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods, 7,* 361–388.

Muthén, B. O. (1988). Some uses of structural equation modeling in validity studies: Extending IRT to external variables. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 213–238). Hillsdale, NJ: Erlbaum.

Muthén, B. O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika, 29,* 81–117.

Muthén, B., & Asparouhov, T. (2002). *Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus.* Unpublished manuscript (Available at http://www.statmodel.com/mplus/examples/webnote.html#web4).

Phillips, K. A., First, M. B., & Pincus, H. A. (2003). *Advancing DSM: Dilemmas in psychiatric diagnosis.* Washington, DC: American Psychiatric Association.

Ruscio, J., Ruscio, A. M., & Keane, T. M. (2004). Using taxometric analysis to distinguish a small latent taxon from a latent dimension with positively skewed indicators: The case of involuntary defeat syndrome. *Journal of Abnormal Psychology, 113,* 145–154.

Simms, L. J., Watson, D., & Doebbling, B. N. (2002). Confirmatory factor analyses of posttraumatic stress symptoms in deployed and nondeployed veterans of the Gulf War. *Journal of Abnormal Psychology, 111,* 637–647.

Starcevic, V. (1995). Pathological worry in major depression: A preliminary report. *Behaviour Research and Therapy, 33,* 55–56.

Watson, D. (2003). Investigating the construct validity of the dissociation taxon: Stability analyses of normal and pathological dissociation. *Journal of Abnormal Psychology, 112,* 298–305.

Watson, D. (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for *DSM–V, Journal of Abnormal Psychology, 114,* 522–536.

Widiger, T. A. (1992). Categorical versus dimensional classification: Implications from and for research. *Journal of Personality Disorders, 6,* 287–300.

Widiger, T. A., Costa, P. T., & McCrae, R. R. (2002). A proposal for Axis II: Diagnosing personality disorders using the five-factor model. In P. T. Costa & T. A. Widiger (Eds.), *Personality disorders and the five-factor model of personality* (2nd ed., pp. 431–456). Washington, DC: American Psychological Association.

Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the *Diagnostic and Statistical Manual of Mental Disorders—Fifth Edition. Journal of Abnormal Psychology, 114,* 494–504.

Received July 8, 2005
Accepted July 8, 2005 ■