DIAGNOSTIC METHODS AND ISSUES

ROGER D. DAVIS MS
THEODORE MILLON PH.D.

ANXIETY AND RELATED DISORDERS

Diagnostic Methods and Issues

ROGER D. DAVIS, MS. and THEODORE MILLON, PhD

e-Book 2015 International Psychotherapy Institute

From *Anxiety and Related Disorders* edited by Benjamin Wolman & George Stricker

Orig. Publisher: John Wiley & Sons

Table of Contents

Authors

Roger D. Davis, MS

University Fellow Department of Psychology
University of Miami
Miami, FL

Theodore Millon, PhD

Editor. *Journal of Personality Disorders*
Professor of Psychology
University of Miami
Coral Gables, FL

Diagnostic Methods and Issues

ROGER D. DAVIS, MS. and THEODORE MILLON, PhD

Scores of instruments are available to diagnose and assess anxiety. More are published each year, so that no one can be acquainted with all of them—new and old instruments; instruments that are empirically derived; instruments that claim a theoretical basis; instruments that assess a particular content domain well; instruments that measure many domains of content; instruments that look good, but have little validational evidence; instruments that possess a substantial research base, but were constructed in the looser nosologic climate of a bygone day—and these are just a few of the possibilities that come to mind.

When confronted with so many alternatives, a database is needed, a list of like kinds of things, systematically arranged, information regarding each systematically presented, together with a set of guidelines for making a selection intelligently. Unfortunately, no database such as this could be confined to a single chapter. Consequently, only the more widely used interview techniques, self-report measures, and clinician rating scales will be considered here. Instruments specific to various anxiety disorders are

presented in tabular form near the end of the chapter. Despite this caution, we hope this chapter will be sufficient to inform the clinician or researcher about what exists in the way of anxiety instruments and point him or her in the direction of original sources.

CLASSIFICATION ISSUES: A CONTEXT FOR DIAGNOSIS AND ASSESSMENT

Before reviewing diagnostic methods and instrumentation, it would be wise to examine certain assumptions that justify their use and interpretation. What is a taxonomy of mental disorders and why is it needed?

A taxonomy is a way of grouping together like kinds of things (Millon, 1991). Essentially, a taxonomy reflects the belief that the items classified fall into more or iess discrete categories. In psychopathology, taxonomies are usually referred to as nosologies. When explicitly articulated, groups and their clinical attributes form a diagnostic system. The presence or absence of such attributes can then be systematically inquired in order to determine group membership, a procedure known as differential diagnosis. The Diagnostic and Statistical Manual (DSM) consists of categories of psychopathology sanctioned by the American Psychiatric Association (e.g., DSM-III-R, 1987), of which the anxiety disorders constitute one particular species.

A nosology serves certain functional ends. From a research perspective,

it provides a means of organizing the dynamic body of knowledge that undergirds a science, allowing the history of substantive questions to be probed and gaps in scientific knowledge to be discerned. From a clinical perspective, a nosology provides a means of organizing clinical phenomena. By abstracting across presentations, a nosology formalizes certain clinical commonalities, relieving the clinician of the burden of dealing with each patient sui generis. At a minimum, persons within a taxon should be more alike than those selected across taxons. The converse is trivial, but also true. Patients from different groups should be more diverse than patients selected within the same group. Similarity-dissimilarity, then, is ultimately the organizing principle on which a nosology is constructed.

Unfortunately, similarity is a fuzzy notion. Exactly how, in what way, are persons who receive the same diagnosis alike? Two levels of similarity must be distinguished, manifest and latent. Patients whose psychopathologies are similar at a manifest level give presentations that look alike. In an empirical nosology, these patients are classified together. The latent level, however, deals with genotypic similarity. Taxons are formed on the basis of theoretical or etiologic commonalities. Patients possessing such commonalities are classified together, regardless of how the pathology is manifest.

Which kind of similarity forms the better basis for a nosology? Table 14.1 presents possible agreements and disagreements between latent and

manifest similarity for two patients.

As can be seen in the table, for any two patients, four possibilities exist. First, going clockwise, two presentations that appear similar, may be similar. In this case, etiologically identical pathways have produced manifestly similar results. Second, two presentations that appear similar, may be different. Here diverse etiologic pathways have produced manifestly similar results difficult to tease apart. Third, two presentations that appear different, may be different. In this case, different manifest characteristics legitimately depict the output from different etiologic pathways. Fourth, two presentations that appear different, may be similar. Here the interaction of identical pathology with individual differences produces diverse presentations.

TABLE 14.1. Matches and Mismatches for Latent and Manifest Levels of Similarity

		Similar at a Latent Level?	
		Yes	No
Similar at a Manifest Level?	Yes	I. Things that appear similar, are in fact similar.	II. Things that appear similar are in fact different. (Nosologically problematic)
	No	IV. Things that appear different are in fact similar. (Nosologically problematic)	III. Things that appear different are in fact different.

Table 14.1 resembles other tables used to present the logic of diagnostic efficiency statistics—true positives, false negatives, and so on (e.g., Baldessarini, Finklestein, & Arana, 1983). Whereas the latter compares the

apparent or obtained diagnosis with the so-called “true” diagnosis for a single subject, illustrating the diagnostic dilemma associated with imperfect predictors in ignorance of the “true” state of nature. Table 14.1 represents the nosologic dilemma, whereby multiple taxons must be established for multiple subjects, the “true” taxonic membership of each being unknown.

We have been speaking as if both manifest and latent levels were known and knowable. In fact, only a manifest level is ever observed. Latent structures and functions, traits and taxons, are, by definition, inferred. Nosologically, this presents tremendous difficulties for psychopathology, the implication being that the pursuit of patterns of covariation among imperfect predictors will lead us, in the second quadrant, to establish one taxon where two, or perhaps many, are needed, and, in the fourth quadrant, to establish two, or perhaps several, taxons where only one in fact exists.

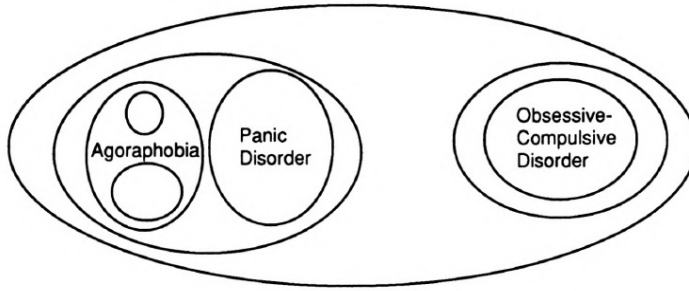
Such is the “bet” of pure empiricism: That things which look alike, are alike, and conversely, that things which do not resemble each other, are in fact different. That no cases of mismatched manifest and latent similarity exist. Since the “true” state of nature remains unknown, it is impossible to determine how often these mismatches actually occur. Perhaps they occur only a few times in the entire DSM, and perhaps the current nosology is rife with them. In any case, since diagnostic labels ideally represent a shorthand means of communicating a theory about the patient’s pathology, one wonders

about the worth of the current diagnostic agenda.

From the standpoint of what is required of a science, this empirical approach to the anxiety disorders is deficient in at least two ways. First, the current nosology lacks any integrative theoretical schema to explain why the psychopathology of anxiety takes the form of these particular disorders rather than others. Second, lack of an undergirding theoretical schema retards the progress of research attempting to illuminate specific mechanisms of particular disorders. Indeed, given the logic of Table 14.1, it can be seen that empirical methods alone are an insufficient basis on which to determine either the nature or the number of anxiety disorders.

The "number" side of this "nature-number" dilemma is intimately related to another contentious problem, the level of abstraction at which a nosology should be articulated. Are more or fewer taxons generally desirable? Pragmatically, the taxons represented must ensure adequate coverage. Scientifically, however, it does not appear possible to constrain the number of taxons that constitute a taxonomy on empirical grounds alone (see Figure 14.1). Cluster analytic methods exemplify this problem. As Blashfield (1980) notes, no sure method exists for solving the "number of clusters" problem. Yet, "Why does nature express itself in these taxons rather than others?" is a legitimate, if not fundamental, scientific question.

Figure 14.1.



A pseudotopographic set diagram illustrating the level of abstraction problem in empirical taxonomies. Smaller circles represent fewer patients, more severe pathology, and the presence of additional clinical attributes. How many taxons are there? One? Three? Or more?

As with nature and nurture, nature and number seem to be intimately connected. Within the anxiety disorders, such a relationship can perhaps be seen in the transition from DSM-II to DSM-III. Frances, Widiger, and Fyer (1990, p. 43), commenting on the issue of "splitting" versus "lumping" and its influence on comorbidity, described the DSM-III-R as a "splitter's dream and a lumper's nightmare" (p. 43). As these authors noted,

> A possible example of the splitting issue is the distinction among the DSM-III anxiety disorders of panic disorder, agoraphobia, and generalized anxiety disorder (GAD). Panic Disorder was included in DSM-III in large part because of Klein's research demonstrating that imipramine blocks recurrent panic attacks but has no apparent effect on associated anxiety or nonassociated phobic anxiety. GAD was created to cover the domain of DSM-II anxiety neurosis not covered by panic. There is now considerable interest in assessing the comorbidity of panic disorder with a variety of syndromes, including GAD. Data supporting the validity of the panic

> disorder diagnosis are extensive, and the diagnosis provides valid and useful information. However, much of the research on the comorbidity of panic and agoraphobia may be due to their mapping a common, overlapping domain of psychopathology, (p. 43)

Apparently, patients are not compelled to hug the taxonic high ground when presenting their clinical pictures. Nonprototypal and overlapping cases illustrate that many patients are quite at home in the taxonic "borderlands" as well.

Even more taxonomically distressing, nature-number issues are not confined to a single branch of the nosologic hierarchy. The relationship between anxiety and depression has long been a contentious issue. Many researchers have approached this problem by examining the relationship between anxiety and depression inventories. Dobson (1985), found that for 16 studies reviewed, anxiety-anxiety, depression-depression, and anxiety-depression scale correlations were all significant, ranging from .61 to .69. Clark and Watson (1991) examined approximately 400 self-report and clinician-rated studies using mood, symptom, and syndrome measures. Disappointing convergent and discriminant validity patterns were again found for most instruments, which, together with factorial evidence, led these authors to argue that "Anxiety and depressed syndromes share a significant nonspecific component that encompasses general affective distress and other common symptoms, whereas these syndromes are distinguished by physiological hyperarousal (specific to anxiety) versus the absence of positive

affect (specific to depression) (p. 331)." A mixed mood disorder, which these authors advocate adding to the nosology, would be represented by symptoms concerned with the absence of positive affectivity (e.g., apathy, hopelessness) and those related to an approximately equal level of negative affectivity. On other fronts, Maser and Cloninger (1990) offer an edited book that authoritatively and comprehensively treats the comorbidity of mood and anxiety disorders from a wide variety of perspectives.

Our purpose is not to propose a solution to the taxonomic quandaries that concern the anxiety disorders. Nevertheless, much can be learned by comparing ideals and actualities. To place the above issues in highest relief, we contrast the current diagnostic agenda with an "ideal" nosology: In a mature clinical science, theory, nosology, instrumentation, and intervention form a conceptually unified whole (Millon, 1990). The critical element that lends this structure cohesiveness is that its undergirding concepts posses systematic import (Hempel, 1965), that is, that these concepts are more or less invariably associated with a large number of other characteristics relevant to prediction in the subject domain. This explanatory and heuristic power suggests a nonarbitrary taxonomic organization within which major nosolgic categories can be grouped and differentiated. Because such a nosology makes theoretical and etiological statements about category members, the assignment of persons to groups is an explanatory rather than merely descriptive affair, illuminating mechanisms of pathology and

suggesting intervention strategies.

The current nosology is not yet at a mature, theoretically-driven stage. The current state of psychopathologic nosology and diagnosis resembles that of medicine a century ago. Concepts remain overwhelmingly descriptive. Yet, by definition, a clinical science must be applied to individual cases. How can the individual case be approached in a scientific, rather than descriptive, fashion? In short, the professional must "bootstrap" within a single subject, that is, develop a theory or theories of the individual and the individual's pathology which, in an iterative process of inference and hypothesis testing, unify available data and resist falsification sufficiently to justify their use in the construction of intervention strategies. Such theoretical development requires instrumentation.

THE TRIPLE RESPONSE MODEL: A CONTENT x METHOD APPROACH

Whether one is dealing with taxonomic matters or with individual cases, the utility of a theory is a function of both its simplicity and scope. The former is simply the number of theoretical constructs required to account for the phenomena of the subject domain. An evaluation of the latter, however, assumes that the range of things to be explained is known in advance. In the hard sciences, such as physics, the extent of the subject domain is easily discernable, in part because these constructs are assumed to possess some

form of physical existence. As one moves from harder to more weakly organized sciences, however, linear causal models give way to feedback (and possibly feedforward) processes that appear to operate not only horizontally, within a given "level" of organization, whether psychological, biological, or physical, but vertically, across organizational levels as well. As a result, psychological concepts are more often multireferential constructs that "float" above the level of data and resist unequivocal quantification by any one particular measurement technique.

It has proven difficult to say exactly what anxiety is, and to distinguish it from what are merely its correlates. Perhaps anxiety is primarily the result of unconscious conflicts striving for expression, of an inability to escape from situations in which one experiences a lack of self-efficacy, of overgeneralization of the anxiety response, of ruminative thoughts, of irrational primary and secondary appraisal, or of daily hassles coupled with neuroendocrine imbalances, and so on. More likely, however, anxiety is as often the result of some or all of these as it is of any one. In lieu of a comprehensive theory of anxiety pathology, we could at least ask for some way of ordering the constellation of causes and correlates that have been associated with the construct.

The triple-response concept represents such an approach, an approach which addresses the multireferential nature of the anxiety construct by

grouping clinical phenomena into several distinct content areas. As explained by Eiffert and Wilson (1991), Lang (1968) introduced the triple-response model for emotional behaviors, arguing that emotional behaviors are mediated by partially independent brain centers. These centers control three systems or modalities: motor, physiological, and verbal-cognitive. The motor component consists of observable aspects of the pathology, such as the degree of avoidance or rate of panic attacks. The physiological modality consists of muscle tension, heart rate, respiration, perspiration, hormonal fluctuations, and other somatic aspects of the pathology. The verbal-cognitive modality consists of verbal reports of anxiety or fear, as well as thoughts that occur before or during such episodes as panic attacks, compulsive rituals, and so on.

Whether the division of the organism into content areas is more pedagogic or substantive, as Lang (1968) apparently believed, the triple response concept has had a number of beneficial effects on anxiety research (Eiffert & Wilson, 1991). Theoretically, it has been helpful in suggesting connections between the behavioral, physiological, and cognitive modalities across a variety of emotional problems and disorders. Methodologically, it has led to the use of multiple assessment instruments within and across methods and modalities. Clinically, it has pointed to the multireferential nature of the anxiety construct and consequent need for comprehensive assessment, and thereby led to a greater integration of assessment and treatment through

consideration of each construct domain in the formulation of intervention strategies.

However, Eiffert and Wilson (1991) also argue that what is measured, the various content areas of the triple response paradigm, has often been confounded with the way of measuring it, that is, method of assessment. These authors recommend a "matrix model" to clearly distinguish method and content, and the division of the "verbal-cognitive-subjective" mode into a cognitive mode concerned with such things as information processing and cognitive styles, and an affective mode concerned with the phenomenology or subjective report of various mood states (see Table 14.2). Unfortunately, many of the cells in the matrix require some kind of special equipment, a certain kind of expertise, or an extended period of time to implement, any of which may not be readily available to clinicians.

TABLE 14.2. Content x Methods-of-Assessment Matrix

	Method of Assessment		
Content Area Assessed	Self-Report Verbal or Nonverbal	Observation	Instrument or Apparatus
Motoric	Mobility inventory, daily activity log	BAT	Pedometer, activity meter
Physiological	Body sensations questionnaire	Perspiring, blushing	GSR, EMG

Cognitive	Attributional style questionnaire	Response latency	STROOP test
Affective	Adjective checklists, mood visual analog scale	Facial expression	Not available

Source: Adapted from Eifert & Wilson (1991).

Although the content by methods-of-assessment matrix does not offer a theory of the anxiety disorders, it does argue for the insufficiency of any single method for measuring the four content areas. Given the influence of method variance (Campbell & Fiske, 1959), such a model suggests an interesting empirical critique of the DSM anxiety disorders and their associated clinical interviews. Although the DSM anxiety disorders criteria do get at verbal, cognitive, physiological, and affective criteria (though not consistently), they rely almost exclusively on the self-report method (though perhaps elicited by interview) for differential diagnostic purposes. To the extent that the patterns of covariation or comorbidity that suggested and shaped current diagnostic boundaries are artifactually influenced by method variance, the current constellation of anxiety disorders is specific to the self-report method. If variance across methods was taken into account, might diagnostic boundaries shift radically? Might entirely new taxons reveal their existence?

CATEGORICAL APPROACHES TO ANXIETY: STRUCTURED AND SEMI-STRUCTURED INTERVIEWS

Although a number of other formulations are possible, clinical conditions have traditionally been thought of categorically. Categorical systems provide a single label for a constellation of clinical attributes that ideally covary with such tenacity that they seem to characterize a discrete diagnostic entity. Thus categorical systems restore unity to the patient's pathology (Millon, 1991) and often suggest aspects of pathology that might otherwise have gone unobserved.

Eight anxiety disorders are recognized in the DSM-III-R: Panic Disorder with and without agoraphobia, agoraphobia without history of panic disorder, social phobia, simple phobia, obsessive compulsive disorder, post-traumatic stress disorder, generalized anxiety disorder, and anxiety disorder not otherwise specified. DSM-IV will likely include another, hybrid category, anxious depression. Descriptions and diagnostic and differential diagnostic criteria for these disorders can be found in the DSM, as can associated features, age at onset, course, impairment, complications, predisposing factors, prevalence, sex ratio, and familial pattern, when such information exists and is believed to be reliable.

Historically, clinical interviews have consisted mainly of a comparatively nondirective history-taking and mental status examination (Wiens, 1990). In the early days of psychopathology, such informality was not problematic: The elements of clinical science—theory, taxonomy,

instrumentation, and intervention—were largely unintegrated. Prior to the introduction of psychotropic medication, interventions were often the same, regardless of diagnosis. In such cases, diagnostic errors were meaningless.

Psychiatric diagnoses have also been notoriously unreliable (e.g., Matarazzo, 1990). This served as a considerable barrier to the development of psychopathology as an integrated science. To the extent that each clinician serves as his own criterion, clinical judgment is no better than opinion. Clinically, such a state of affairs communicates nothing about a patient's disorder and is worthless as a basis for intervention. Experimentally, unreliable diagnosis limits group homogeneity, introducing noise into research designs.

This began to change with the introduction of DSM-III and the explicit formulation of diagnostic criteria, from which diagnostic interviews could be developed. Structured and semi-structured interviews increase diagnostic reliability by providing the interviewer with a highly formalized set of questions. These questions internalize diagnostic criteria and standardize the encounter between interviewer and client. Generally, the degree of formality required depends upon the level of expertise of the interviewer. Structured interviews are usually intended to be administered by trained laypersons and therefore tend to be more highly rigorous. Semi-structured interviews, in contrast, rely more heavily on the clinical judgement of the professional and

can afford to be more open-ended.

Structured interviews have largely accomplished their intended purpose. The introduction of reliable structured interviews and the adoption of modern diagnostic models, beginning with DSM-III with its multiaxial taxonomy, polythetic categories, and field trials, together with the subsequent explosion of psychiatric research, have together produced a clinical science whose elements are now more coupled. Far from being irrelevant, diagnostic errors now often result in wasted time by clinicians and patients, wasted money by patients and third-party payers, mismedication, and possibility of legal entanglements. Arriving at a correct (or at least consensual) diagnosis is more important than ever before. To this end, the purpose of the modern clinical interview is to obtain a detailed history and statement of current symptoms which can serve as a competent basis for diagnosis and intervention.

Schedule for Affective Disorders and Schizophrenia (SADS)

The SADS (Endicott & Spitzer, 1978) was designed to make diagnoses in accordance with the Research Diagnostic Criteria (RDC; Spitzer, Endicott, & Robins, 1978). Since its inception, the SADS has grown into a family of instruments that includes the regular form, a lifetime form (SADS-L), a follow-up or change form (SADS-C), and more recently, a lifetime anxiety version

(SADS-LA). The SADS-LA (Fyer, Endicott, Mannuzza, & Klein, 1985) was developed from the SADS-L explicitly for the investigation of anxiety disorders, including separation anxiety disorder and adjustment disorder with anxious mood. All diagnoses covered in the lifetime version, such as major depression and substance use, frequently comorbid with the anxiety disorders, are also included. RDC, DSM-III, and DSM-III-R diagnoses can be generated.

One notable feature of this instrument is its lifetime sequential approach to assessment (Mannuzza, Fyer, Klein, & Endicott, 1986). Rather than simply ask whether the particular symptoms of a given disorder have ever been present, the SADS-LA seeks to provide a comprehensive portrait of the onset of symptoms, syndromes, and a variety of life events ranging from marriage, to career change, to the death of a child. These are recorded on a Life Chart Digital Coding Form. The eventual result is a computer-generated Life Chart, effectively a history of all relevant psychopathology, and its beginning and ending in relation to prominent life events. As a result, a great variety of investigations can be supported, for example, lifetime as well as cross-sectional comorbidity, and the sequencing of symptoms and life events in the development of disorder, for a variety of diagnostic criteria.

Test-retest reliability for the SAD-LA appears strong. An investigation by Mannuzza, Fyer, Martin, et al. (1989) using highly trained interviewers

showed lifetime kappas ranging from .60 for general anxiety disorder to .90 for agoraphobia. Simple phobia showed poorer reliability, which the authors attribute to imprecision of the DSM-III-R impairment and distress criterion for this disorder. Agreement was generally better for current episodes rather than past ones, especially for obsessive-compulsive disorder (.91 vs. .58) and social phobia (.68 vs. .33). No significant reliability differences were found between RDC and DSM-III-R criteria for any disorder. One limitation of the interview is its somewhat lengthy administration time, one and one-half to two and one-half hours.

The Structured Clinical Interview for the DSM-III-R—Patient Version (SCID-P)

The SCID-P (Spitzer, Williams, & Gibbon, 1988) is a semi-structured interview designed for use by trained clinicians and mental health professionals with psychiatric patients. Nonpatient (SCID-NP) and Axis II (SCID-II) versions are also available. While the entire SCID can be administered at intake, modules for each major diagnostic group are provided, allowing the clinician to "confirm and document a suspected DSM-III-R diagnosis" (Spitzer, Williams, Gibbon, & First, 1990, p. 1). Moreover, the authors encourage the adaptation of the inverview for specific purposes. Thus, a familiar or favored interview technique or scale can be synthesized with the advantages of a structured interview, and customized for particular studies.

Eleven modules make up the SCID-P, including an Overview Module and the Summary Score Sheet. During the overview, the interviewer records basic demographic information such as age, sex, education, and work history. Treatment history and a description of the current illness are also elicited. The authors note that upon completing the overview, the interviewer should possess sufficient information to justify a "tentative differential diagnosis." Current and lifetime diagnoses are generated, with the exception of, in the anxiety module, generalized anxiety disorder. These diagnoses are then recorded on the Summary Score Sheet, current diagnoses as present or absent, lifetime diagnoses as present, absent, or subthreshold. Administration time generally runs 60 to 90 minutes. The SCID-P is sufficient to meet the needs of most clinicians or researchers, but it does not (nor is it intended to) characterize the subject's psychopathology as richly as the Life Chart of the SADS-LA. Other versions of the SCID available include the SCID-UP, developed for the detailed study of Panic and Generalized Anxiety Disorders, and the SCID-NP-V for posttraumatic stress disorder.

Anxiety Disorders Interview Schedule—Revised (ADIS-R)

Di Nardo, O'Brien, Barlow, Waddell, and Blanchard (1983, p. 1070) state that the ADIS was developed "for three major purposes: to permit differential diagnosis among the DSM-III anxiety disorder categories, to provide sufficient information to rule out psychosis, substance abuse, and major affective

disorders, and to provide data beyond basic information required for establishing diagnostic criteria." The Hamilton Anxiety Scale and Hamilton Depression Scale are embedded in the interview by content area. Since the goal of the interview is the comprehensive description of anxiety pathology, skip-outs are infrequent.

Di Nardo and Barlow (1990) reported ADIS kappa coefficients of .905 for social phobia, .854 for agoraphobia with panic, .825 for obsessive-compulsive disorder, .651 for panic disorder, .571 for generalized anxiety disorder, and .558 for simple phobia. Blanchard, Gerardi, Kolb, and Barlow (1986) reported a kappa of .857 for the presence or absence of PTSD using "expert opinion" as the criterion diagnosis.

Critique of the Categorical Approach and Diagnostic Interviews

Although diagnostic interviews have greatly improved the reliability of diagnosis, a number of problems remain. Reliability is no substitute for validity. Diagnostic categories should reflect some underlying reality. Whether the current scheme of anxiety disorders accomplishes this goal is by no means certain. The standardized format of interview techniques imbues diagnostic categories with a measurement precision one would expect only from taxons which indeed exist. As we have seen, however, manifest similarity can mask genotypic heterogeneity. The reliability of diagnostic

interviews promotes the masquerade of such “composite” taxons as singular disorders.

At a practical level, questions remain about the internal structure of the interviews themselves. Although interdiagnostician reliability at the diagnostic level may be fairly high, reliability at the symptom level has less often been examined. Since a polythetic model requires an individual meet only a subset of diagnostic criteria, the interrater reliability of the diagnosis itself is surely greater than the reliability of individual interview questions intended to assess specific symptoms or content areas. Yet a reliable symptom picture is exactly what is needed to inform personalized intervention strategies.

DIMENSIONAL MEASURES: SELF REPORT

Many self-report instruments are available. Those which deal with anxiety as a unidimensional construct mainly include items associated with generalized anxiety and panic attacks. For the most part, these instruments were constructed prior to DSM-III, before the DSM-II category anxiety neurosis was split into generalized anxiety disorder, panic, and agoraphobia. Such instruments might be called syndrome measures, because they assess anxiety as it cuts across many different disorders. Items related to obsessional or compulsive content and phobias are fewer, when present.

Consequently, the specificity of these scales to any one anxiety disorder is probably limited.

The Beck Anxiety Inventory (BAI)

The BDI (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961; Beck & Steer, 1987) is a well-known and extensively used instrument (for a review, see Beck, Steer, & Garbin, 1988). The BAI (Beck, Epstein, Brown, & Steer, 1988) is a more recent development and has not yet been well researched. The authors state that "The BAI was developed to address the need for an instrument that would reliably discriminate anxiety from depression while displaying convergent validity" (p. 893).

The scale consists of 21 items, each scored 0 (Not at all) to 3 (Severely—I could barely stand it) according to the degree the respondent has been "bothered" by the particular symptom within the past week. The final score ranges from 0 to 63. A factor analysis reported by the authors yielded two factors: (1) somatic symptoms and (2) subjective anxiety and panic symptoms. The scale is weighted toward the first factor. The authors report high internal consistency (alpha = .92) and a BAI-BDI correlation of only .48.

State-Trait Anxiety Inventory (STAI)

One can distinguish between an enduring tendency to feel or behave in

a particular way, and the way one feels now, that is, between traits and states, a distinction which underlies the STAI (Spielberger & Rickman, 1990). The STAI is composed of two scales, A-Trait and A-State, each 20 items long. The A-Trait scale regards frequency. Subjects are requested to report how they generally feel. The A-State scale regards intensity. Subjects are instructed to report how they currently feel. The STAI was originally published as Form X in 1970. This was revised in 1983 to address such problems as response biases and discriminant validity in relation to depression. A children's version is also available.

Validity information concerning the STAI comes not only from "experimental" evidence (state-anxiety manipulating paradigms, such as test taking or relaxation training), but from an examination of its psychometric characteristics as well. The internal consistency of both the trait and state scales is high, approximately .90. Test-retest reliability of the A-trait scale is also high, while test-retest reliability of the A-state scale is much lower over long intervals. That these psychometric characteristics conform to such theoretical expectations speaks well of the scale.

Spielberger has recently applied the trait-state distinction in another inventory, the State-Trait Anger Expression Inventory (STAXI; Spielberger, 1988). The STAXI yields measures of State-Anger, Trait-Anger, and Anger Expression. The latter concerns whether anger is expressed (Anger-Out) or

suppressed (Anger-In), a distinction no doubt having important behavioral health and psychodynamic implications.

Minnesota Multiphasic Personality Inventory-2

The MMPI, originally published in 1943, has long been a staple for mental health professionals. Unfortunately, construction by the empirical keying approach, which neglects content and internal consistency concerns, rendered interpretation of individual scales problematic, ultimately resulting in the use of two- and three-point codes profiles and a variety of MMPI "cookbooks."

In 1989, the MMPI-2 was published with the addition of fifteen content scales, developed using "multi-stage, multi-method procedures that combined rational and statistical methods" (Butcher, Graham, Williams, & Ben-Porath, 1990, p. 26). These can be interpreted instead of or in addition to the heterogeneous clinical scales. The content scales include the ANX (23 items) anxiety scale, the OBS obsessiveness (16 items) scale, the FRS (23 items) fears scale, and the DEP (33 items) depression scale. Regarding the ANX and OBS scales the authors state:

> High scorers on ANX report general symptoms of anxiety including tension, somatic problems (i.e., heart pounding and shortness of breath), sleep difficulties, worries, and poor concentration. They fear losing their minds, find life a strain, and have difficulty making decisions. They appear to be aware of these symptoms and problems, and admit to having them.

> High scorers on OBS have tremendous difficulty making decisions and are likely to ruminate excessively about issues and problems, causing others to become impatient. Having to make changes distresses them, and they may report some compulsive behaviors like counting or saving unimportant things. They are excessive worriers who frequently become overwhelmed by their own thoughts, (p. 36)

Although the content scales are relatively new, scale descriptions and reliability statistics are promising. Butcher et al. (1990) report internal consistencies ranging from .82 to .90 for the ANX scale, .78 to .84 for the OBS scale, and .71 to .84 for the FRS scale, for male and female subjects from psychiatric, alcoholic, and military samples. As might be expected, the ANX and OBS scales show moderate intercorrelation, .66 and .72 for males and females respectively. Much of this can be explained as item overlap. Two shared items heavily influence the correlation due to the relative shortness of the OBS scale. Without these items the correlation drops to .44 and .52 for males and females, respectively.

Millon Clinical Multiaxial Inventory-II (MCMI-II)

Compared to the MMPI, the MCMI (Millon, 1987) is a short (less than 200 items) instrument intended to coordinate with the multiaxial format of the DSM. The personality scales of the MCMI are grounded in a three-polarity metapsychology derived from evolutionary theory (Millon, 1990). The MCMI-II contains several subscales intended to screen for typical Axis I disorders, such as anxiety, dysthymia, and alcohol dependence. Millon (1987) reported a

sensitivity of .69 and a positive predictive power of .74 for the anxiety scale.

Symptom Checklist-90-R

The SCL-90-R (Derogatis, 1977) is "a multidimensional self-report symptom inventory designed to measure symptomatic psychological distress" (Derogatis, 1982, p. 277). Ninety items are rated from "Not at all" to "Extremely" on a 5-point scale in relation to symptom severity. From these 90 items, nine primary clinical-rational symptom dimensions are formed: somatization (12 items), obsessive-compulsive (10 items), interpersonal sensitivity (9 items), depression (13 items), anxiety (10 items), hostility (6 items), phobic anxiety (7 items), paranoid ideation (6 items), and psychoticism (10 items). In addition, three global severity indexes are derived, a global severity index (GSI), a positive symptom total (PST), and a positive symptom distress index (PSDI). The global indexes provide different ways of looking at the severity of symptoms in conjunction with the number of symptoms. The PST is simply the number of symptoms reported. The PSDI, however, is described as "a pure intensity measure, adjusted for the number of symptoms present" (Derogatis, 1982, p. 278), while the GSI reflects both the intensity of distress and number of reported symptoms. Norms for psychiatric outpatients, inpatients, nonpatient adolescents, and nonpatient normals are available for both sexes. Symptoms are generally assessed during the past week, however, the time frame is flexible for research purposes.

Administration time runs 15 to 20 minutes, and a microcomputer scoring program is available.

Although the SCL-90-R is a popular and widely used instrument, some researchers have questioned its internal structure. Wetzler (1989) noted that while Derogatis, Lipman, Covi, and Rickels (1972) and Prusoff and Klerman (1974) reported that the SCL-90 clearly discriminated depressed and anxious patients, later studies (e.g., Angst & Dobler-Mikola, 1985; Clark & Friedman, 1983) failed to confirm this finding. Other researchers have questioned the independence and composition of its scales. Cyr, McKenna-Foley, and Peacock (1985) opened a literature review of the factor structure of the SCL-90-R and related variants by stating that "Many studies provide evidence of the poor item consistency among factors across studies, the low frequency of factor replication based on postulated dimensions, and questionable factorial constancy across various criteria for the several forms of the Symptom Checklist." These authors advised using the SCL-90-R as a measure of general distress only.

Profile of Mood States (POMS)

The POMS (McNair, Lorr, & Droppleman, 1971) is a 65-item affect adjective checklist developed through factor analytic research with both psychiatric patients and normals. Six primary mood dimensions are assessed:

tension-anxiety, depression-dejection, confusion-bewilderment, anger-hostility, vigor-activity, and fatigue-inertia. Each item is rated on a five-point scale from "Not at all" to "Extremely." The time frame of the instrument is "the past week including today." College and outpatient norms for males and females are provided.

Moods are by definition more time-limited than are symptoms and traits. Reliabilities therefore are constrained by the transient nature of the construct. One month test-retest correlations of from .61 to .69 (McNair & Lorr, 1964) and 20-day correlations (McNair et al., 1971) of from .65 to .74 have been reported. Internal consistencies for the POMS dimensions range from acceptable, .74, to high, .92. The POMS anxiety and depression scales, though moderately intercorrelated, appear to possess higher convergent and discriminant validities than the Multiple Affect Adjective Checklist anxiety and depression measures, in part because each item is scaled rather than simply checked (Clark & Watson, 1991).

DIMENSIONAL MEASURES: CLINICIAN RATINGS

Numerous clinician-rated instruments are available to assess anxiety. Two will be considered here, the Hamilton Anxiety Rating Scale (HRSA) and the Zung Anxiety Scale (ZAS).

The Hamilton Rating Scale for Anxiety (HRSA)

Hamilton introduced the HRSA in 1959 to assess the severity of clinical anxiety in patients diagnoses as suffering from anxiety neurosis.

Apparently, several slightly different versions of the original scale are in use. Generally, the HRSA consists of approximately 90 symptoms of anxiety grouped rationally under 13 to 15 categories, including anxious mood, tension, fears, insomnia, cognitive symptoms, depressed mood, somatic symptoms (muscular), somatic symptoms (sensory), cardiovascular symptoms, respiratory symptoms, gastrointestinal symptoms, genitourinary symptoms, autonomic symptoms, and behavior-at-interview. These items are scored 0 (not present) to 4 (very severe) depending on severity in the last one week, and then totaled to obtain a global severity rating. Sheehan and Harnett-Sheehan (1990, p. 91) notes that "scores above 18 are usually considered abnormal."

Hamilton (1959) reported a correlation of .89 between independent ratings of the same interview. Two orthogonal factors, a general severity factor, and a bipolar psychic vs. somatic factor were also reported. Maier, Buller, Philipp, and Heuser (1988) found joint-rater interview reliabilities of .74, .73, and .70 for the total score, the psychic factor, and the somatic factor, respectively.

Since its publication, the HRSA has become one of the most widely used

rating scales for anxiety. Nearly every study regarding the effectiveness of potential anxiolytics includes it. Both Hamilton scales are included on the ADIS. Nevertheless, the HRSA has its limitations. The HRSA is not intended to be used with patients whose anxiety is associated with other psychiatric disorders. Although the HRSA covers a wide variety of symptoms, these were grouped into variables on rational grounds, so that there is no guarantee that such groupings in fact form natural clusters. The reliability of some items (respiratory symptoms, behavior at interview, and autonomic symptoms) appears low, less than .30 (Maier et al., 1988), and it is unclear how enduring severe symptoms and severe symptoms of brief duration (e.g., panic attacks) should be weighted when appraising the patient's condition. Nor is the HRSA especially useful for the differential diagnosis of anxiety disorders. No obsessional or panic-specific items are included. In a patient sample studied by Di Nardo and Barlow (1990), no significant difference between HRSA scores for panic disorder, generalized anxiety disorder, agoraphobia, and obsessive-compulsive disorder were found.

A number of investigators have taken up the task of refining the HRSA. Snaith, Baugh, Clayden, Husain, and Sipple (1982) recently developed the Clinical Anxiety Scale (CAS) from an item analysis of the HRSA. The CAS consists of six variables scored on a five-point scale. Exact scoring instructions are given. The authors state (p. 520) that the CAS is largely confined to "psychic anxiety and tension in the somatic musculature."

The Zung Inventories

Zung (1971) developed the Self-Rating Anxiety Scale (SAS) and the clinician-rated Anxiety Status Inventory (ASI). Both are 20-item instruments based on DSM-II description of anxiety neurosis and other anxiety symptoms described by authoritative psychiatry texts of the time. Five items assess affective symptoms; fifteen assess somatic complaints. The ASI deals with the severity of each symptom, while the SAS deals with the frequency of each symptom.

ASI items are scored 1 to 4. Clinicians are encouraged to use all available information in assigning severity values, including intensity ("How bad was it?"), duration ("How long did it last?"), and frequency ("How much of the time did you feel that way?") (Zung, 1971, p. 373). These are assumed to co-vary so that, for example, a "2" corresponds to "Mild in intensity or duration, present some of the time in frequency." An Interview Guide is presented in order to facilitate administration and coverage, however, the clinician is allowed to ask additional questions and probe for details. The time frame is arbitrarily limited to one week.

Each SAS item corresponds to a similar item in the ASI. Items are again scored 1 to 4, from "None or a little of the time" to "Most or all of the time," and the time frame is again set at one week. Five items are scored opposite to the other fifteen to discourage response biases.

Little data is available regarding the reliability and validity of the ASI and SAS. Zung (1971) reported correlations between the ASI and SAS of .66, but only .30 between the SAS and TMAS and .33 between the ASI and TMAS. Within an anxiety disordered group the ASI-SAS correlation rose to .74. Split-half correlations of .83 (ASI) and .71 (SAS) were reported. Jegede (1977) studied the characteristics of the SAS in a group of normals and a group of Nigerian outpatients. Alpha's of .69 for the normal group and .81 for the patients were reported. Item 17 ("My hands are usually dry and warm") was negatively correlated with the remaining items in both samples, suggesting that scale performance might be improved by deleting this item.

DISORDER SPECIFIC SCALES AND DIAGNOSTIC EFFICIENCY STATISTICS

Anxiety is what is common to the anxiety disorders. Logically, then, unidimensional scales of anxiety should possess only limited specificity with regard to the disorders themselves. Fortunately, given the development of many scales specific to the anxiety disorders and aspects of these disorders, the possibility of using these comparatively short scales in place of reliable, but tedious, structured interviews becomes an issue of great clinical import. The anxiety disorders are considered in detail elsewhere in this text. Nevertheless, Tables 14.4 through 14.7 list instruments relevant to the anxiety disorders.

What are diagnostic efficiency statistics and in what ways are scores on disorder specific scales diagnostic? Adoption of the polythetic model in DSM-III approached a paradigm shift in the conception of mental disorders. By this model, no single criterion is necessary or sufficient for the diagnosis of disorder. Instead, only some number of diagnostic criteria must be met, say four or five. The polythetic model recognizes the natural heterogeneity among patients which exists even within a single diagnostic taxon. In terms of the medical model of mental illness, which finds its greatest applicability in the Axis I disorders, we might say that, ideally, this heterogeneity derives from the interaction of individual differences and an underlying disease entity or process, so that variability is a natural characteristic in the manifestation of pathology.

Since all clinical attributes are to some extent the result of a disease by individual differences interaction, that is, fallible rather than unequivocal predictors, the probability of possessing the disorder given any one predictor or set of predictors becomes of interest. Indeed, this is the diagnostician's dilemma: What is the positive predictive power (PPP), the probability of disorder, given the symptom or clinical picture? Presumably, when a certain number of critical predictors or symptoms are present, the probability of disorder is deemed sufficiently high to justify intervention, and a diagnosis is made. Although the positive predictive power of diverse sets of diagnostic criteria taken, say, four at a time, need not be equal (Widiger, Hurt, Frances,

Clarkin, & Gilmore, 1984), such is the justification of diagnostic thresholds.

Positive predictive power is part of a larger family of diagnostic efficiency statistics which includes sensitivity, specificity, and negative predictive power (NPP). Sensitivity is the proportion of all patients who possess the symptom of interest, while specificity is equal to the proportion of patients without the disorder who do not possess the symptom of interest. Negative predictive power is the proportion of patients without the symptom of interest, and without the disorder. These statistics are easily summarized in tabular form (see Table 14.3).

TABLE 14.3. Diagnostic Efficiency Statistics

Diagnosis Given the Symptom:	True (or Criterion) Diagnosis	
	Possess Disorder	Lack Disorder
Positive	True-Positives (a)	False-Positives (b)
Negative	False-Negatives (c)	True-Negatives (d)
Total	Total with Disorder (a + c)	Total without Disorder (b + d)

Base rate = (a + c)/(a + c + b + d)
Sensitivity = a/(a + c)Specificity = d/(b + d)
Positive predictive power = a/(a + b)
Negative predictive power = d/(c + d)

TABLE14.4. Instruments for Use with Post-Traumatic Stress Disorder

Clinician-Administered PTSD Scale—Form I	Blake et al. (1990)
Crime-Related Post-Traumatic Stress Scale (within SCL-90-R)	Saunders, Arata, & Kilpatrick (1990)
Impact of Events Scale	Horowitz, Wilner, & Alverez (1979)
Incident Report Interview	Kilpatrick et al. (1987)
Keane MMPI Subscale	Keane, Malloy, & Fairbank (1984)
Mississippi Scale for Combat Related PTSD	Keane, Cadell, & Taylor (1988)
Mississippi Scale for Combat Related PTSD (Short Form)	Hyer, Davis, Boudewyns, & Woods (1991)
Penn Inventory for PTSD	Hammerburg (1992)
The PTSD Interview	Watson et al. (1991)
Rape Aftermath Symptom Test	Kilpatrick (1988)
Sexual Experiences Survey	Koss & Gidycz (1985)
Vietnam Stress Inventory	Wilson & Krauss (1984)

TABLE 14.5. Instruments for Use with Agoraphobia and Panic Disorder

Agoraphobia Scale	Ost (1990)

Agoraphobic Cognitions Questionnaire	Chambless, Caputo, Bright, & Gallagher (1984)
Anxiety Sensitivity Index	Reiss, Peterson, Gursky, & McNally (1986)
Body Sensations Questionnaire	Chambless et al. (1984)
Dyadic Adjustment Scale	Spainer (1979)
Fear Survey Schedule (FSS-III)	Wolpe & Lang (1964)
Fear Questionnaire	Marks & Mathews (1979)
Locke-Wallace Marital Adjustment Scale	Locke & Wallace (1959)
Marital Satisfaction Inventory	Synder, Wills, & Keiser (1981)
Mobility Inventory	Chambless, Caputo, Jasin, Gracely, & Williams (1985)
Panic Attack Questionnaire	Norton, Dorward. & Cox (1986)
Panic Attack Symptom Questionnaire	Clum, Broyles, Borden, & Watkins (1990)
Panic Attack Cognitions Questionnaire	Clum et al. (1990)
Sheehan Panic Attack and Anticipatory Anxiety Scale	Sheehan (1983)

TABLE 14.6. Instruments for Use with Obsessive-Compulsive Disorder

Compulsive Activity Checklist	Marks, Hallam, Connolly, & Philpott (1977)
Hamburg Obsession/Compulsion Inventory-Short Form	Klepsch, Zaworka, Hand, Lunenschloss, & Jauernig (1991)
Leyton Obsessional Inventory (Card Sort)	Cooper (1970)
Leyton Obsessional Inventory (Paper & Pencil)	Kazarian, Evans, & Lefave (1977)
Lynfield Obsessional Compulsive Questionnaire	Allen (1977)
Maudsley Obsessional-Compulsive Questionnaire	Hodgson & Rachman (1977)
Padua Inventory	Sanavio (1988)
Yale-Brown Obsessive Compulsive Scale	Goodman et al. (1989a,b)

TABLE 14.7. Instruments for Use with Social Phobia

Embarrassibility Scale	Modigliani (1968)
Fear of Negative Evaluation Scale	Watson & Friend (1969)
Fear of Negative Evaluation Scale (Brief Version)	Leary (1983a)
Fear Questionnaire	Marks & Mathews (1979)
Fear Survey Schedule (FSS-III)	Wolpe & Lang (1964)

Interaction Anxiousness Scale	Leary (1983b)
Shyness Scale	Cheek & Buss (1981)
Situation Questionnaire	Rehm & Marston (1968)
Social Anxiety History Questionnaire	Turner, Beidel, Dancu, & Keys (1986)
Social Anxiety Questionnaire	Arkowitz, Lichtenstein, McGovern. & Hines(1975)
Social Anxiety Inventory	Richardson & Tasto (1976)
Social Avoidance and Distress Scale	Watson & Friend (1969)
Social Interaction Self-Statement Test	Glass, Merluzzi, Biever, & Larsen (1982)
Social Performance Survey Schedule	Lowe & Cautela (1978)
Social Phobia and Anxiety Inventory	Beidel, Turner, Stanley, & Dancu (1989)
Social Reticence Scale (SRS)	Jones & Russell (1982)
SRS—Revised	Jones, Briggs, & Smith (1986)
Social Situations Questionnaire	Bryant & Trower (1974)
Stanford Shyness Survey	Zimbardo (1977)
Survey of Heterosexual Interactions	Twentyman & McFall (1975)

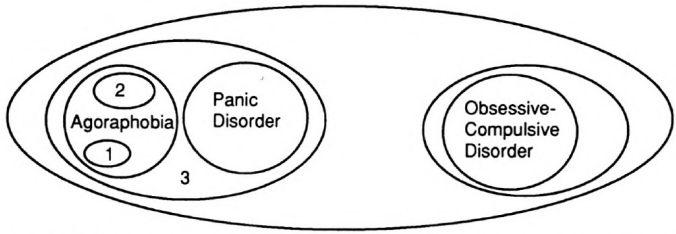
Source: Glass & Arnkoff (1989), brief review of most of these instruments.

Unfortunately, diagnostic efficiency statistics often sometimes seem to work counterintuitively or at odds with one another. For example, it is quite possible for sensitivity to be low, yet PPP, high. This occurs when very few patients with a given disorder possess a particular symptom, but when they do, it is an extremely good predictor. Moreover, optimal diagnostic cutting scores vary with the prevalence (base) rate of a disorder (see especially Baldessarini et al., 1983). When base rates across clinical settings are substantially different from development conditions, optimalcutting scores can vary widely. Although some authors have considered this factor when providing instrument validation data (e.g., Keane, Caddell, & Taylor, 1988), on the whole such information is lacking.

We will not examine the all of the vissitudes of diagnostic efficiency statistics here (see Baldessarini et al., 1983; Meehl & Rosen, 1955; Widiger et al., 1984), but only note that some of these can be grasped intuitively if connected to the level-of-abstraction issue. Consider Figure 14.2. For polemical purposes, assume that, as in the earlier example, agoraphobia and panic disorder indeed map a common domain of psychopathology, and that, as shown, agoraphobia contains two smaller "sub-taxons," and that they have as their nearest neighbor obsessive-compulsive disorder. Assume further that all those within a set also possess the given attribute. Attribute 1 represents

the case of low sensitivity and high PPP. In regard to attribute 2, more patients possess the symptom, which also predicts disorder. Thus attribute 2 possesses moderate sensitivity and high PPP. For agoraphobia, attribute 3 possesses high sensitivity, in that all those with agoraphobia also possess attribute 3. Unfortunately, attribute 3 is also common to panic disorder, which, on the basis of this attribute alone, leads to a large number of false positives and thus possesses poor differential diagnostic positive predictive power relative to either agoraphobia or panic disorder. Nevertheless, the attribute holds high differential power relative to obsessive-compulsive disorder.

Figure 14.2.



A pseudotopographic set diagram for illustrating the relationship between sensitivity and positive predictive power. Smaller circles are assumed to represent fewer patients, severe pathology, and the presence of additional clinical attributes. Attribute 1 possesses high PPP for agoraphobia, but poor sensitivity. Attribute 3 possesses high sensitivity, but poor PPP due to a large number of false positives.

SUMMARY

The current scheme of anxiety disorders is essentially an empirical creation, lacking an undergirding theoretical schema. Disputes concerning the nature and number of the disorders themselves continue both within (e.g., the validity of GAD) and across (e.g., relationship between anxiety and depression) branches of the diagnostic hierarchy. Fortunately, a means exists of organizing both the clinical phenomena related to anxiety and methods for its investigation through the content by methods-of-assessment matrix. Beyond suggesting the possibility of new nosologic taxons, the use of multimethod approaches across a variety of content areas should be quite helpful in developing a unified theory of the patient's pathology.

Several diagnostic interviews have proven highly reliable, but to revisit theoretical shortcomings, reliability is no substitute for validity, that is, for a theoretical basis. Numerous dimensional measures of anxiety exist, in both clinician-rated and self-report formats. However, anxiety is what is common to the anxiety disorders. Logically, then, a unidimensional anxiety scale can possess only limited differential diagnostic utility. That is not to say that such scales are useless, only that they will probably not be sufficient to answer which anxiety disorders are present, to what degree, and in what way.

Fortunately, the specification and refinement of diagnostic criteria in DSM-III and later editions, together with the development of structured interviews, have led to the emergence of a variety of scales intended to assess

specific anxiety disorders. The Mobility Inventory (Chambless, Caputo, Jasin, Gracely, & Williams, 1985), for example, is intended to assess self-reported agoraphobic avoidance. Since these scales are much more tightly focused than, say, the HRSA, they should be particularly helpful in suggesting the presence or absence of symptoms, illuminating areas of intervention, and measuring therapeutic progress. While an exclusively empirical approach to psychopathology cannot be recommended, in general these scales should offer a considerable improvement in predictive power and specificity over unidimensional anxiety scales, and should be invaluable in determining the future directions of this area both clinically and experimentally.

REFERENCES

Allen, J. J. (1977). The measurement of obsessionality: First validation studies of the Lynfield obsessional/compulsive questionnaire. *Journal of International Medical Research, 5(5),* 12-15.

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

Angst, J., & Dobler-Mikola, A. (1985). The Zurich study: VI. A continuum from depression to anxiety disorders? *European Archives of Psychiatry and Neurological Science, 235,* 179-186.

Arkowitz, H., Lichtenstein, K., McGovern, K., & Hines. P. (1975). Behavioral assessment of social competence in males. *Behavior Therapy, 6,* 3-13. Baldessarini. R. J., Finklestein, S., & Arana, G. W. (1983). The predictive power of diagnostic tests and the effect of prevalence of illness. *Archives of General Psychiatry, 40,* 569-573.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical

anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56,* 893-897.

Beck, A. T., & Steer, R. A. (1987). *Manual for the revised Beck Depression Inventory.* San Antonio, TX: The Psychological Corporation.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8,* 77-100.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561-571.

Beidel, D. C., Turner, S. M., Stanley, M. A., & Dancu, A. V. (1989). The Social Phobia and Anxiety Inventory: Concurrent and external validity. *Behavior Therapy, 20, AM-421.*

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Klauminzer, G., Charney, R., & Keane, T. M. (1990). A clinician rating scale for assessing current and lifetime PTSD: The CAPS-1. *The Behavior Therapist, 13,* 187-188.

Blanchard, E. B., Gerardi, R. J., Kolb, L. C., & Barlow, D. H. (1986). The utility of the Anxiety Disorders Interview Schedule (ADIS) in the diagnosis of Post-Traumatic Stress Disorder (PTSD) in Vietnam veterans. *Behaviour Research and Therapy, 24,* 577-580.

Blashfield, R. K. (1980). Propositions regarding the use of cluster analysis in clinical research. *Journal of Consulting and Clinical Psychology, 48,* 456-459.

Bryant, B., & Trower, P. E. (1974). Social difficulty in a student sample. *British Journal of Educational Psychology, 44,* 13-21.

Butcher, J. N., Graham, J. R., Williams, C. L., & Ben-Porath, Y. S. (1990). *Development and Use of the MMPI-2 Content Scales.* Minneapolis: University of Minnesota Press.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81-105.

Chambless, D. L., Caputo, C., Bright, P., & Gallagher, R. (1984). Assessment of fear in agoraphobics: The Body Sensations Questionnaire and the Agoraphobic Cognitions Questionnaire. *Journal of Consulting and Clinical Psychology, 62.* 1090-1097.

Chambless, D. L., Caputo, G. C., Jasin, S. E., Gracely, E. J., & Williams, C. (1985). The Mobility Inventory for Agoraphobia. *Behaviour Research and Therapy, 23,* 35-44.

Cheek, J. M., & Buss, A. H. (1981). Shyness and sociability. *Journal of Personality and Social Psychology, 41,* 330-339.

Clark, A., & Friedman, M. J. (1983). Factor structure and discriminant validity of the SCL-90 in a veteran psychiatric population. *Journal of Personality Assessment, 47,* 396-404.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100,* 316-336.

Clum, G. A., Broyles, S., Borden, J., & Watkins, P. L. (1990). Validity and reliability of the Panic Attack Symptoms and Cognitions Questionnaire. *Journal of Psychopathology and Behavioral Assessment, 12,* 233-247.

Cooper, J. E. (1970). The Leyton Obsessional Inventory. *Psychological Medicine, 1,* 48-64.

Cyr, J. J., McKenna-Foley, J. M., & Peacock, E. (1985). Factor structure of the SCL-90-R: Is there one? *Journal of Personality Assessment, 49,* 571-578.

Derogatis, L. R. (1977). *SCL-90-R administration, scoring, and procedures manual,* vol. 1. Baltimore: Clinical Psychometric Research.

Derogatis, L. R. (1982). Self-report measures of stress. In L. Goldberger & S. Breznitz (Eds.), *Handbook of stress: Theoretical and clinical aspects* (pp. 270-294 New York: The Free Press).

Derogatis, L. R., Lipman, R. S., Covi, L., & Rickels, K. (1972). Factorial invariance of symptom dimensions in anxious and depressive neuroses. *Archives of General Psychiatry, 27,* 659-665.

Di Nardo, P. A., & Barlow, D. H. (1990). Syndrome and symptom co-occurrence in the anxiety disorders, In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of Mood and Anxiety Disorders* (pp. 205-230). Washington, DC: American Psychiatric Press.

Di Nardo, P. A., O'Brien, G. T., Barlow, D. H., Waddell, M. T., & Blanchard, E. B. (1983). Reliability of DSM-III anxiety disorder categories using a new structured interview. *Archives of General Psychiatry, 40,* 1070-1074.

Dobson, K. S. (1985). The relationship between anxiety and depression. *Clinical Psychology Review, 5,* 307-324.

Eifert, G., & Wilson, P. H. (1991). The triple response approach to assessment: A conceptual and methodological reappraisal. *Behaviour Research and Therapy, 29,* 283-292.

Endicott, J., & Spitzer, R. L. (1978). A diagnostic interview: The Schedule for Affective Disorders and Schizophrenia. *Archives of General Psychiatry, 27,* 678-687.

Fyer, A. J., Endicott, J., Mannuzza, S., & Klein, D. F. (1985). *Schedule for Affective Disorders and Schizophrenia-Lifetime Version* (modified for the study of anxiety disorders). New York: Anxiety Disorders Clinic, New York State Psychiatric Institute.

Frances, A., Widiger, T., & Flyer, M. R. (1990). The influence of classification methods on comorbidity. In J. D. Maser & C. R. Cloninger (Eds.), *Comorbidity of Mood and Anxiety Disorders* (pp. 41-60). Washington, DC: American Psychiatric Press.

Glass, C. R., & Arnkoff, D. B. (1989). Behavioral assessment of social anxiety and social phobia. *Clinical Psychology Review, 9,* 75-90.

Glass, C. R., Merluzzi, T. V., Biever, J. L., & Larsen, K. H. (1982). Cognitive assessment of social anxiety: Development and validation of a self-statement questionnaire. *Cognitive Therapy and Research, 6,* 37-55.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heninger, G. R., & Charney, D. S. (1989a). The Yale-Brown Obsessive-Compulsive Scale: I. Development, use, and reliability. *Archives of General Psychiatry. 46,* 1006-1011.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Delgado, P., Heninger, G. R., & Charney, D. S. (1989b). The Yale-Brown Obsessive-Compulsive Scale: II. Validity. *Archives of General Psychiatry, 46,* 1012-1016.

Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology, 32,* 50-55.

Hammerburg, M. (1992). Penn Inventory for Posttraumatic Stress Disorder: Psychometric properties. *Psychological Assessment, 4,* 67-76.

Hempel, C. G. (1965). *Aspects of scientific explanation.* New York: Free Press.

Hodgson, R. J., & Rachman, S. (1977). Obsessional-compulsive complaints. *Behaviour Research and Therapy, 15,* 389-395.

Horowitz, M., Wilner, N., & Alverez, W. (1979). Impact of Events Scale: A measure of subjective distress. *Psychosomatic Medicine, 41,* 209-218.

Hyer, L., Davis, H., Boudewyns, P. A., & Woods, M. G. (1991). A short form of the Mississippi Scale for Combat-Related PTSD. *Journal of Clinical Psychology, 47,* 510-518.

Jegede, R. O. (1977). Psychometric attributes of the Self-Rating Anxiety Scale. *Psychological Reports, 40,* 303-306.

Jones, W. H., Briggs, S. R., & Smith, T. G. (1986). Shyness: Conceptualization and measurement. *Journal of Personality and Social Psychology, 51,* 629-639.

Jones, W. H., & Russell, D. (1982). The Social Reticence Scale: An objective measure of shyness. *Journal of Personality Assessment, 46,* 629-631.

Kazarian, S. S., Evans, D. R., & Lefave, K. (1977). Modification and factorial analysis of the Leyton Obsessional Inventory. *Journal of Clinical Psychology, 33,* 422-425.

Keane, T. M., Cadell, J. M., Taylor, K. L. (1988). Mississippi Scale for Combat-Related Posttraumatic Stress Disorder: Three studies in reliability and validity. *Journal of Consulting and Clinical Psychology, 56,* 85-90.

Keane, T. M., Malloy, P. F., & Fairbank, J. A. (1984). Empirical development of an MMPI subscale for the assessment of combat-related posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology, 52,* 888-891.

Kilpatrick, D. G. (1988). Rape Aftermath Symptom Test. In M. Herson & A. S. Bellack (Eds.), *Dictionary of Behavioral Assessment Techniques* (pp. 366-367). New York: Pergamon Press.

Kilpatrick, D. G., Veronen, L. J., Saunders, B. E., Best, C. L., Amick-McMullen, A., & Paduhovich, J. (1987). *The psychological impact of crime: A study of randomly surveyed crime victims.* (Final Report, Grant No. 84-IJ-CX-0039). Washington, DC: National Institute of Justice.

Klepsch, R., Zaworka, W., Hand, I., Lunenschloss, K., & Jauernig, G. (1991). Derivation and validation of the Hamburg Obsession/Compulsion Inventory—Short Form (HOCI-S): First results. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 196-201.

Koss, M. P., & Gidycz, C. A. (1985). Sexual Experiences Survey: Reliability and Validity. *Journal of Consulting and Clinical Psychology, 53,* 422-423.

Lang, P. J. (1968). Fear reduction and fear behavior: Problems in treating a construct. In J. M. Schlien (Ed.), *Research in psychotherapy* (Vol. III.) (pp. 90-102). Washington, DC: APA.

Leary, M. R. (1983a). A brief version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin, 9,* 371-376.

Leary, M. R. (1983b). Social anxiousness: The construct and its measurement. *Journal of Personality Assessment, 47, 66-75.*

Locke, H. J., & Wallace, K. M. (1959). Marital-adjustment and prediction tests: Their reliability and validity. *Marriage and Family Living. 21,* 251-255.

Lowe, M. R., & Cautela, J. R. (1978). A self-report measure of social skill. *Behavior Therapy. 9,* 535-544.

Maier, W., Buller, R., Philipp. M., & Heuser, I. (1988). The Hamilton Anxiety Scale: Reliability, validity, and sensitivity to change in anxiety and depressive disorders. *Journal of Affective Disorders, 14,* 61-68.

Mannuzza, S., Fyer. A., Klein, D., and Endicott, J. (1986). Schedule for Affective Disorders and Schizophrenia-Lifetime Version modified for the study of anxiety disorders (SAD-LA): Rationale and conceptual development. *Journal of Psychiatric Research. 20,* 317-325.

Mannuzza, S., Fyer, A. J., Martin, L. Y., Gallops, M. S., Endicott, J., Gorman, J., Liebowitz. M. R., & Klein, D. F. (1989). Reliability of anxiety assessment: I. Diagnostic agreement. *Archives of General Psychiatry, 46,* 1093-1101.

Marks, I. M., & Mathews, A. M. (1979). Brief standard self-rating for phobic patients. *Behaviour Research and Therapy, 17,* 263-267.

Marks, I., Hallam, R., Connolly, S., & Philpott, R. (1977). *Nursing in behavioural psychotherapy: An advanced clinical role for nurses.* London and Tonbridge: Whitefriars Press.

Maser, J. D., & Cloninger, C. R. (Eds.), (1990). *Comorbidity of Mood and Anxiety Disorders.* Washington. DC: American Psychiatric Press.

Matarazzo, J. D. (1990). Psychological assessment versus psychological testing: Validation from Binet to the school, clinic, and courtroom. *American Psychologist, 45,* 999-1017.

McNair, D. M., & Lorr, M. (1964). An analysis of mood in neurotics. *Journal of Abnormal and Social Psychology, 69,* 620-627.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1971). Profile of Mood States. San Diego: EdiTS/Educational and Industrial Testing Service.

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin, 52,* 194-216.

Millon, T. (1987). *Manual for the Millon Clinical Multiaxial Inventory-11.* Minneapolis: National Computer Systems.

Millon, T. (1990). *Toward a new personology.* New York: Wiley.

Millon, T. (1991). Classification in psychopathology: Rationale, alternatives, and standards. *Journal of Abnormal Psychology. 100.* 245-261.

Modigliani, A. (1968). Embarrassment and embarrassibility. *Sociometry, 31,* 313-326.

Norton, G. R., Dorward, J., & Cox, B. J. (1986). Factors associated with panic attacks in nonclinical subjects. *Behavior Therapy, 17,* 239-252.

Ost, L-G. (1990). The Agoraphobia Scale: An evaluation of its reliability and validity. *Behaviour Research and Therapy, 28,* 323-329.

Prusoff, B., & Klerman, G. L. (1974). Differentiating depressed from anxious neurotic outpatients: Use of discriminant function analysis for separation of neurotic affective states. *Archives of General Psychiatry, 30,* 302-309.

Rehm, L. P., & Marston, A. R. (1968). Reduction of social anxiety through modification of self-reinforcement. *Journal of Consulting and Clinical Psychology, 32,* 565-574.

Reiss, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency, and the prediction of fearfulness. *Behaviour Research and Therapy, 24,* 1-8.

Sanavio, E. (1988). Obsessions and compulsions: The Padua Inventory. *Behaviour Research and Therapy, 26,* 169-177.

Saunders, B. E., Arata, C. M., & Kilpatrick, D. G. (1990). Development of a Crime-Related Post-Traumatic Stress Disorder scale for women within the Symptom Checklist-90-Revised. *Journal of Traumatic Stress, 3,* 439-448.

Sheehan, D. V. (1983). *The anxiety disease.* New York: Scribner's.

Sheehan, D. V., & Harnett-Sheehan, K. (1990). Psychometric Assessment of Anxiety Disorders. In N. Sartorius, V. Andreoli, G. Cassano, L. Eisenberg, P. Kielholz, P. Pancheri, & G. Racagni (Eds.), *Anxiety: Psychobiological and clinical perspectives* (pp. 85-100). New

York: Hemisphere.

Snaith, R. P., Baugh, S. J., Clayden, A. D., Husain, A., & Sipple, M. A. (1982). The Clinical Anxiety Scale: An instrument derived from the Hamilton Anxiety Scale. *British Journal of Psychiatry, 141,* 518-523.

Spanier, G. B. (1979). The measurement of marital quality. *Journal of Sex and Marital Therapy, 5,* 288-300.

Spielberger, C. D. (1988). *State-Trait Anger Expression Inventory manual.* Odessa, FL: Psychological Assessment Resources.

Spielberger, C. D., & Rickman, R. L. (1990). Assessment of state and trait anxiety. In N. Sartorius, V. Andreoli, G. Cassano, L. Eisenberg, P. Kielholz, P. Pancheri, & G. Racagni (Eds.), *Anxiety: Psychobiological and clinical perspectives* (pp. 69-84). New York: Hemisphere.

Spitzer, R. L., Endicott, J., and Robins, E. (1978). Research Diagnostic Criteria: Rationale and reliability. *Archives of General Psychiatry, 35,* 773-782.

Spitzer, R. L., Williams, J. B. W., & Gibbon, M. (1988). *Structured Clinical Interview for DSM-III-R.* New York: Biometric Research Department, New York State Psychiatric Institute.

Spitzer, R. L., Williams, J. B. W., Gibbon, M., & First, M. B. (1990). *User's guide for the Structured Clinical Interview for DSM-III-R.* Washington, DC: American Psychiatric Press.

Synder, D. K., Wills, R. M., & Keiser, T. W. (1981). Empirical validation of the Marital Satisfaction Inventory: An actuarial approach. *Journal of Consulting and Clinical Psychology, 49,* 262-268.

Turner, S. M., Beidel, D. C., Dancu, C. V., & Keys, D. J. (1986). Psychopathology of social phobia and comparison to avoidant personality disorder. *Journal of Abnormal Psychology, 95,* 389-394.

Twentyman, C. T., McFall, R. M. (1975). Behavioral training of social skills in shy males. *Journal of Consulting and Clinical Psychology, 43,* 384-395.

Watson, C. G., Juba, M. P., Manifold, V., Kucala, T., & Anderson, P. E. D. (1991). The PTSD interview: Rationale, description, reliability, and concurrent validity of a DSM-Ill-based technique. *Journal of Clinical Psychology, 47,* 179-188.

Watson, D, & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology, 47,* 525-535.

Widiger, T. A., Hurt, S. W., Frances, A. F., Clarkin, J. F., & Gilmore, M. (1984). Diagnostic efficiency and DSM-III. *Archives of General Psychiatry, 41,* 1005-1012.

Wiens, A. N. (1990). Structured clinical interviews for adults. In G. Goldstein & M. Herson (Eds.), *Handbook of Psychological Assessment* (2nd ed.). (pp. 324-341). New York: Pergamon.

Wilson, C. G., & Krauss, G. E. (1984, September). *The Vietnam Era Stress Inventory: A scale to measure war stress and Post-traumatic Stress Disorder among Vietnam veterans.* Paper presented at the 3rd National Conference on Post-traumatic Stress Disorder, Baltimore.

Wolpe, J., & Lang, P. J. (1964). A fear survey schedule for use in behavior therapy. *Behavior Research and Therapy, 2,* 27-30.

Zimbardo, P. G. (1977). *Shyness: What it is and what to do about it.* Reading, MA: Addison-Wesley.

Zung, W. W. K. (1971). A rating instrument for anxiety disorders. *Psvchosmatics, 12,* 371-379.