CONCURRENT VALIDITY OF DEPRESSION DIAGNOSES VERSUS CONTINUOUS SYNDROME DEPRESSION SCORES

By

Shellene M. Ball

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Psychology

May, 2006

Nashville, Tennessee

Approved:

Professor Bahr Weiss

Professor David A. Cole

ACKNOWLEDGEMENTS

I would like to thank my advisor and mentor, Bahr Weiss, for taking me in as his student, having confidence in my potential, and providing the nurturing guidance and support any novice needs throughout this process. I hope to emulate the same generous spirit toward others and their scholarship as he has toward me and mine.

I also would like to thank my husband, Lynn, and daughter, Colby. I feel honored to share my life with these incredible people. Their unwavering patience, support, and their many sacrifices have allowed me the latitude to choose my path.

I would like to acknowledge the support of NIH training grant, T32-MH070329-01. This funding has given me the opportunity to transition from quantitative to clinical psychology and to pursue this and other research endeavors.

Finally, I would like to thank Steve Hollon and Irene Elkin for their generous support in responding to our requests for TDCRP data and other study-related information. I also would like to thank David Cole for reviewing and providing helpful feedback on this manuscript.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

There has been a long-standing tension between categorical and dimensional systems of psychopathology that reflects fundamental differences in the systems' conceptualization of psychopathology as well as the empirical realities of their utility and limitations. In a conference convened by NIMH in 1990 (Jensen, Koretz, Locke, Schneider, Radke-Yarrow, Richters, & Rumsey, 1993) to set a progressive research agenda in the area of child and adolescent mental health, it was concluded that work comparing categorical and dimensional systems of psychopathology was essential in order to gain a better understanding of these systems and their utility in the research and treatment of child and adolescent psychopathology. More recently, an increasing amount of attention has been given to the evaluation of categorical and dimensional systems of psychopathology as planning for the DSM-V progresses; for example, an entire section in a recent volume of the *Journal of Abnormal Psychology* was devoted to summarizing research in this area and weighing the pros and cons of implementing dimensional systems for certain diagnoses in the DSM-V (Krueger, Watson, & Barlow, 2005). Thus, while there has been a long running and continuing discussion in the literature, it is clear that additional work needs to be done in this area before consensus can be reached on optimal practices for clinicians and researchers.

Research that has focused on the comparative utility of categorical and dimensional systems of psychopathology typically falls into one of two types of research:

(a) that considering the statistical / methodological implications of dichotomizing a continuous measure (e.g., MacCallum, Zhang, Preacher & Rucker, 2002), which is a fundamental aspect of diagnoses; and (b) that investigating applied and substantive aspects of categorical versus continuous measures of psychopathology (e.g., Ferdinand, Heijmens Visser, Hoogerheide, van der Ende, Kasius, Koot, & Verhulst, 2004). Before reviewing these literatures, however, it is important first to consider some of the basic assumptions and issues inherent to the comparison of nosological and dimensional systems.

Categorical and dimensional measures utilize two different broad conceptualizations of psychopathology: (a) psychopathology as a class qualitatively distinct from 'normality;' and (b) psychopathology as continuous syndromes that vary on a continuum from 'normality.' Psychiatric diagnostic systems based on the first conceptualization of psychopathology include DSM-IV (American Psychiatric Association, 1994) and ICD-10 (World Health Organization, 1992); measures designed to provide an assessment within these frameworks include structured clinical interviews such as SCID, SADS, and K-SADS (First, Spitzer, Gibbon, & Williams, 2002; Endicott & Spitzer, 1978; Chambers, Puig-Antich, Hirsch, Ambrosini, & Tabrizi, 1985). Assessment systems based on the second conceptualization involve ratings of people made on continuous syndromes, and include such measures as the Child Behavior Checklist and the Personality Assessment Inventory (Achenbach & Edelbrock, 1983; Morey, 1991).

The applied goals of these two types of systems generally are different. The first seeks to determine an individual's membership or non-membership in classes of

psychopathology, often for the purposes of selection for treatment or inclusion in a research study, whereas the second seeks to quantify an individual's level of psychopathology along a given dimension(s). In addition to the practical differences of these two approaches, two distinct psychometric models underlie these two approaches to measurement of psychopathology. Continuous syndrome measures imply observed variation among individuals reflects (to the extent that the measure is reliable and valid) true score variation along a quantitative trait or traits. Thus, differences between individuals reflect differences in the degree or level of the latent construct of psychopathology. Categorical measures imply a different psychometric model, based on the assumption of qualitative differences between individuals with and without psychopathology. The psychometric model underlying the categorical measurement of psychopathology treats variation within the class along a continuous measure of psychopathology as error variance.

As the above discussion suggests, continuous and categorical systems both have positive and negative aspects, depending on the uses to which they are applied. Diagnoses derived from categorical systems have been valued historically for their ease of communication among professionals, and for their necessity when categorical decisions must be made (Cantwell, 1996; Kraemer, Noda, & O'Hara, 2004). However, although classifying individuals is useful and appealing for the above reasons, there is no consensus on how best to construct categories. Most diagnostic systems are based on what ultimately are arbitrary cut-points (i.e., the number and specific symptoms that must be endorsed to receive a diagnosis) for determining membership in classes of psychopathology. For example, the DSM-IV diagnostic criteria for affective disorders

reflect consensus of appointed committee members rather than empirically derived optimal cut-points (Ferdinand et al., 2004). This often can be problematic since, for instance, individuals with sub-clinical levels of symptoms may benefit from treatment (McDermott & Weiss, 1995; Hinshaw, Lahey, & Hart, 1993), and individuals with sub-threshold symptom levels often are not fundamentally different from individuals who have passed the diagnostic threshold (e.g., Fergusson, Horwood, Ridder, & Beautrais, 2005).

However, even "empirical" derivation of optimal cut-points is not problem-free, as there are numerous procedures used to classify individuals or objects and little consensus on best practice (Fergusson & Horwood, 1995). Some examples of the various methods used to empirically classify individuals or objects include ROC analysis, factor analysis, cluster analysis, bimodality, multiple modality, admixture analysis, latent class analysis, latent growth trajectory analysis, discontinuous regression against an external variable, and taxometrics / coherent cut kinetics (see Kraemer et al., 2004; Beauchaine, 2003; and Trull, Widiger, & Guthrie, 1990 for brief summaries of these techniques).

Although the DSM-IV model and the measures assessing it have been criticized for such apparent flaws, there are some potential advantages relative to continuous measures of psychopathology. Diagnoses via the DSM-IV or other systems allow for the inclusion of certain simultaneous criteria that may be critical for differentiating psychopathology from normality, such as: (a) a required time duration for symptoms, (b) whether or not there is impaired psychosocial functioning and subjective distress in regards to the symptoms, and (c) presence of key symptoms (i.e. sadness or anhedonia in MDD). Such criteria as these have seldom if ever been incorporated into continuous

dimensional systems, yet may reflect important underlying realities of psychopathology that differentiate different forms of psychopathology from each other as well as from normality. Conversely, although continuous syndrome approaches to the measurement of psychopathology do not lend themselves to making categorical decisions, they do have distinct advantages, the foremost of which is that their adherence to the theoretical conceptualization of psychopathology as continuous syndromes may reflect reality (Clark, Watson, & Reynolds, 1995). And it should be noted that some forms of psychopathology may best be conceptualized and assessed categorically, whereas others should be conceptualized and assessed continuously.

Literature Review

In this next section, the literature comparing categorical and dimensional systems for assessing psychopathology is reviewed. As mentioned above, this review will focus on two types of studies that have evaluated the relative merits of categorical versus continuous approaches to the assessment of psychopathology: (a) studies addressing statistical implications of dichotomizing a continuous measure; and (b) studies addressing applied and/or substantive issues associated with categorical versus continuous measurement of psychopathology. Both types of studies are relevant to research comparing nosological and dimensional systems because both methodological and substantive factors need to be considered to answer comparative questions.

Statistical Studies

One of the most consistent conclusions reached in the statistical literature has been that substantial negative effects arise when a continuous measure is converted to a categorical measure (which is usually achieved via dichotomization) (Cohen, 1983; MacCallum et al., 2002; Maxwell & Delaney, 1993). This directly relates to the issue of diagnostic systems, as all diagnostic systems for psychopathology (e.g. DSM, ICD) involve dichotomization of a continuous rating scale of some form to construct a categorical diagnosis.

These issues have perhaps been best summarized by MacCallum et al. (2002), who conducted a comprehensive review of studies in this area. They also conducted additional simulation work to illustrate the negative consequences of dichotomization, reviewed several prominent psychological journals to determine the pervasiveness of dichotomization, and examined the logic behind reasons researchers give for dichotomizing continuous variables. The results of their simulation studies suggest that dichotomization generally results in reduced statistical power, as demonstrated in earlier studies (e.g., Cohen, 1983), but also may result in spurious significant main and interaction effects when one or more variables are dichotomized. Furthermore, the simulation studies also indicated that the effects of dichotomization are particularly problematic when sample sizes and population correlations are small. Dichotomization under these circumstances more often results in spuriously inflated relations among variables. Thus, in any given study with dichotomized variables, it is difficult to determine what effect dichotomization may have had on observed relationships.

Although papers have repeatedly warned against the use of dichotomization (e.g., Cohen, 1983; Maxwell & Delaney, 1993), MacCallum et al. (2002) found that the practice of dichotomization still was fairly common in their survey of several prominent psychological journals. To better understand the reasons behind the practice of dichotomization, individual studies including dichotomized variables were reviewed and justifications for dichotomization were noted. The most commonly cited justifications for dichotomization included: (a) considering observed relations between dichotomized variables and other variables as a conservative estimate of the true relations among variables, (b) lack of information about appropriate analytic methods for investigating relations among continuous variables, (c) dichotomizing when higher correlations are observed between dichotomized and other variables, (d) dichotomizing to simplify findings, (e) dichotomizing to represent classes of individuals, and (f) dichotomizing to increase the reliability of a measure (MacCallum et al., 2002). The simulation evidence previously mentioned calls into question the logic behind each of these justifications.

These six justifications are relevant to this discussion because they reflect common errors in judgment made by researchers who use diagnoses to predict outcomes. Justifications (a) and (c) above are based on the assumption that observed correlations are unbiased estimates of true relations among variables. The simulated results of MacCallum et al. (2002) and Maxwell and Delaney (1993) indicate that this assumption carries an increased risk for making inappropriate inferences. Researchers who compare categorical and continuous measures by observing the relations these measures have with outcome variables of interest assume that the observed relations between the categorical

and outcome measures are unbiased. This assumption is faulty and inferences made based on this type of comparison are tenuous at best.

Justification (e), dichotomizing to represent distinctive classes of individuals, is most relevant to the current discussion because nosological systems dichotomize for the purpose of distinguishing between classes of individuals with and without psychopathology. There are many inherent problems to this practice according to MacCallum and colleagues (2002). Forming classes based on dichotomization assumes the class distinction is more important than individual differences, that dichotomization is a good representation of the underlying latent class structure, and that the chosen cut-point accurately reflects the base rates of the classes. The current nosological systems (i.e. DSM-IV, ICD-10) are not based on empirical evidence supporting adherence to these assumptions and are not likely to be good representations of a theorized latent class structure.

Given the problems associated with dichotomization, MacCallum et al. (2002) conclude that there are few if any circumstances where dichotomization of a continuous variable is justified. Two exceptions to this general rule were noted: (a) When results from taxometric analyses suggest there are underlying categories and a distinct optimal cut-point for dichotomization; and (b) when count variables are extremely skewed. These exceptions were noted with the caveat that individual differences information will be lost in both cases and that lost information still may have more utility than class distinction depending on the question.

This work by MacCallum et al. (2002) and others (e.g., Cohen, 1983; Maxwell & Delaney, 1993) on the effects of dichotomization of continuous measures has obvious

implications for the utility of diagnostic vs. syndromal measures. As noted by Kraemer et al. (2004), hypothesis testing should be conducted using continuous measures of psychopathology rather than categorical diagnostic measures. However, it should be noted that diagnoses are more than a mere dichotomization of a continuous scale. As mentioned above, diagnoses via the DSM-IV system often also include other criteria beyond crossing a severity / frequency symptom threshold, such as impaired functioning, presence of key symptoms, and a specified duration of symptoms (American Psychiatric Association, 1994). These "simultaneous criteria" can be thought of as representing statistical interactions wherein the implications of crossing the symptom severity / frequency threshold – vis-à-vis the hypothesized latent diagnostic construct – differ as a function of the status of the simultaneous criteria. To the extent that the simultaneous criteria in diagnostic systems actually do increment the prediction of outcomes, it may be important to consider collecting and including them in addition to continuous measures of psychopathology. However, to the best of our knowledge there has been no research in this area.

Substantive Studies

In addition to these statistical studies evaluating the effects of dichotomized measures, researchers have compared the magnitude of relations between categorical and continuous measures of psychopathology and other relevant variables (e.g., Fergusson & Horwood, 1995; Jensen, Watanabe, Richters, Roper, Hibbs, Salzberg, & Liu, 1996). The rationale is that comparing the concurrent and predictive validity of diagnostic and syndrome measures of psychopathology against various theoretically-relevant criteria, it

will be possible to determine the relative validity of these two approaches to conceptualizing psychopathology. These studies assessing concurrent and predictive validity build upon the descriptive work of others who used analytic methods to determine whether the structure of phenomena fit a categorical or dimensional model by assessing relations among constructs, which is a fundamental and ultimate goal of science. This work is particularly relevant to understanding and comparing both the diagnostic and dimensional systems in current use and their relations with relevant outcome variables.

Fergusson and Horwood (1995) studied the predictive validity of categorical versus dimensional measures of ODD, ADHD, and CD in a sample of 935 adolescents from a longitudinal study of a birth cohort in New Zealand (the Christchurch Health and Development Study; CHDS). Self and parent report of ODD, ADHD, and CD symptoms were collected when adolescents were 15 and outcome data were collected at age 16. In order to construct the dimensional measures assessing ODD, ADHD, and CD, latent trait models were fitted to the item sets defining each syndrome to insure unidimensionality of each item set. Then, scores on individual items were summed to form a scale score for each syndrome. Diagnoses were derived from the item sets by applying DSM-III-R criteria for ODD, ADHD, and CD. Outcome variables studied included self and parent report of various indices of delinquency: Daily cigarette smoking, cannabis use, alcohol problems, self-reported juvenile offending, and high school dropout.

Estimates of the relations between the categorical and dimensional measures of ODD, ADHD, and CD and outcome variables were calculated using phi coefficients. In each case, the point estimate of this relation was greater for the dimensional measure of

the construct. Although the point estimates of the relations were greater when using dimensional measures, statistical tests were not conducted to determine whether differences between phi coefficients based on diagnoses versus syndromes were significant. Fergusson and Horwood's (1995) study additionally has been criticized (Jensen, 1995) because (a) diagnoses were based on symptom counts that did not include simultaneous criteria included in the DSM model; (b) the external validators seemed theoretically more related to CD rather than ADHD or ODD; and (c) an epidemiological sample rather than a clinical sample was used to investigate psychopathology and related outcomes. Given the limitations to this study, its results must be considered as suggestive rather definitive regarding the hypothesis that continuous systems used to measure psychopathology are more appropriate than nosological systems for predicting outcomes.

A study comparing the concurrent validity of nosological and dimensional measures of psychopathology was conducted by Jensen et al. (1996). Data for this study were collected from 482 randomly selected military families in the Washington, D.C. area whose children ranged in age from 5 to 17. Jensen et al. (1996) conducted this comparison differently from that of Fergusson and Horwood (1995). Jensen et al. (1996) constructed categorical measures of psychopathology from information collected with a dimensional measure (i.e., the CBCL) and dimensional measures of psychopathology from information collected with a diagnostic measure (i.e., the DISC). They then compared the resulting categorical measures (CBCL- and DISC-derived) with one another, and the resulting dimensional measures (CBCL- and DISC-derived) with one another with respect to their relations with criterion factors. Comparison proceeded in this manner because generally it is assumed that dichotomization results in a loss of

statistical power to detect relations, yielding results that are inherently biased against dichotomous diagnostic systems. However, it has been demonstrated, as discussed above, that a loss of statistical power is not the only outcome that can occur when measures are dichotomized. The methods used in this study provide unique information about the utility of current nosological versus dimensional systems of psychopathology because they provide means for a fairer comparison of the two approaches to conceptualizing psychopathology.

Jensen et al. (1996) found that when nosological and dimensional systems were compared in this way, observed relations with external criteria were similar. Specifically, in almost all cases observed relations between dimensional measures of psychopathology (anxiety, depression, hyperactivity, oppositional / aggressive, delinquent / conduct disorder, and total symptoms) and the criterion factors (school dysfunction, need for and use of mental health services, family relationships, and psychosocial / developmental risk factors) did not differ statistically, regardless of the source from which dimensional information was derived (i.e. CBCL or DISC). Similarly, the observed relations between categorical measures of psychopathology and the outcomes (e.g. ADHD diagnosis, and total number of ADHD problems) outlined above also did not differ statistically in almost all cases, regardless of the source from which the categorical information was derived (i.e. CBCL or DISC). Statistical differences in concurrent validity coefficients across categorical versus dimensional measures were not tested, however, although point estimates of the validity coefficients were sometimes larger for dimensional measures and sometimes larger for categorical measures. One must keep in mind, however, the cautions issued against dichotomization previously discussed when interpreting these

results (MacCallum et al., 2002; Maxwell & Delaney, 1993). Taken together, these findings suggest that dimensional information can be derived from either a syndrome or diagnostic system (i.e., the CBCL or DISC, respectively) and used to predict concurrent outcomes equally well. However, as we have noted above, diagnoses represent more than dichotomization, and Jensen et al. (1996) did not include or test effects of simultaneous criteria.

Ferdinand et al. (2004) assessed the relative predictive validity of DISC diagnoses and CBCL scale scores in relation to (a) use of outpatient and inpatient mental health services, (b) parents' desire for professional help, (c) disciplinary problems in school, and (d) police and judicial contacts. Data for this study were collected from 96 families whose children, aged 6 to 12, were referred to a children's psychiatric hospital in Rotterdam. The 8 narrowband CBCL scales, the Internalizing scale, the Externalizing scale, and the Total Problems score were used as univariate predictors of the criterion measures. DISC / DSM-III-R diagnoses occurring in at least 10% of the sample also were used as univariate predictors of the criterion measures. In these analyses, many of the CBCL scale scores and DISC / DSM diagnoses were significant univariate predictors of the outcomes. All significant predictors from the first set of analyses were entered simultaneously to predict each outcome in a new set of regressions to determine how information from the CBCL and DISC could be optimally combined to maximize prediction of outcomes. These authors found (a) a significant interaction between the DISC Conduct Disorder diagnosis and the CBCL Delinquent Behavior scale in the prediction of disciplinary problems in school; (b) a significant main effect for the DISC Agoraphobia diagnosis in the prediction of outpatient treatment; (c) a significant main

effect for the DISC Oppositional Defiant Disorder diagnosis in the prediction of inpatient treatment; (d) a significant interaction effect between the DISC Generalized Anxiety Disorder and Oppositional Defiant Disorder diagnoses in the prediction of police / judicial contacts; and (e) a significant main effect for the DISC Psychosis Screen in the prediction of wish for professional help.

Results from the Ferdinand et al. (2004) study indicate that, in some cases, diagnoses and dimensional measures of psychopathology may increment each other in the prediction of relevant outcomes. However, the generalizability of the specific findings from this study may be limited because of the nature of the specialized referred sample of children and adolescents used. Specifically, the boys within Ferdinand et al.'s (2004) sample scored significantly higher on CBCL scales than those generally referred to receive mental health services. Ferdinand et al. noted that the children and adolescents participating in this study were recruited from a university based clinic, which usually receives referrals for cases that are more complex and severe than what is seen in typical referred populations. It also is notable that in this small, referred sample of children and adolescents problem domains were widely variable and outcome domains studied also were broad. Interestingly, significant predictors of these broad outcomes were limited to very specific problem domains.

Results from these substantive studies fill a gap in the literature because they provide information about categorical and continuous measures' of psychopathology relative importance in predicting outcome criteria. However, the results from these studies are mixed, making it difficult to draw conclusions. For example, Fergusson and Horwood (1995) concluded that dimensional measures were better predictors of later

outcomes while Ferdinand et al. (2004) found that sometimes diagnoses were better predictors, sometimes continuous syndromes were better predictors, and sometimes interactions between the two were better predictors depending on the outcome. Jensen et al.'s (1996) approach was somewhat different than the approach of the authors in the other two studies. Specifically, their approach was to rescale each type of measure as both a categorical and continuous measure and compare the corresponding dimensional and categorical measures. When they did this, they found that the DISC and CBCL performed approximately equally well in terms of the magnitude of relations with outcomes. They did not compare categorical and continuous measures against one another, however.

In addition, to date no studies have directly assessed the role that the simultaneous criteria may play in the relative utility of diagnostic versus syndromal systems of psychopathology. Given that it is well established that dichotomization of a continuous measure results in a loss of power, insofar as diagnostic measures may add unique information to syndrome measures it is likely that the simultaneous criteria may be the source of any unique information. There are other limitations to the current literature as well. First, it is difficult to know whether findings will replicate across samples and problem domains. Some of the existing studies in this area focus on broad problem domains and broad outcome domains whereas others focus on narrower domains. Research looking at broad problem and outcomes domains is problematic, because a wide range of factors could be hypothesized to affect broad outcomes. When specific problems are related to broad outcomes (i.e., Ferdinand et al., 2004), the results must be replicated in order to gauge their reliability and generalizability. Second, most of the

existing studies comparing concurrent and predictive validity of categorical and continuous measures of psychopathology have focused on child and adolescent psychopathology, perhaps because the CBCL provides a relatively comprehensive analogue to the categorical measures of DSM diagnoses. Nonetheless, little research has been conducted in this area with adults, and insofar as psychopathology may become more differentiated from normality with increased development, differences between diagnostic and syndromal measures may be more evident in adults. Third, in general, current studies have confounded questionnaire with level of assessment (i.e. categorical vs. dimensional). That is, we do not know if differences in predictive or concurrent validity coefficients are due to use of categorical vs. dimensional systems, or if these differences are reflective of different content across measures used.

The purpose of the present study is to build upon the existing literature comparing categorical diagnostic measures and continuous syndrome measures of psychopathology and their relations to relevant outcomes. It extends the current literature to adult psychopathology and addresses some of the limitations of existing studies. Specifically, the current study compares categorical diagnostic measures and continuous syndromes based on the same interview questions so as to unconfound the assessment instrument with the level of assessment (continuous vs. categorical). It focuses on a single class of psychopathology (Major Depressive Disorder) so that our outcome domains can be specifically and theoretically related to the disorder/syndrome in question. Perhaps most importantly, we directly assess effects of simultaneous criteria used in diagnostic systems to determine whether they increment prediction of relevant outcomes. Finally, we

include both continuous and categorical outcome measures, to most broadly assess the validity of the different systems.

In the present study, a treatment sample is used for two reasons. First, diagnoses often represent a critical selection and outcome variable in treatment studies, and use of a treatment sample will indicate whether this use is necessary or even appropriate. Second, use of a treatment sample allows for use of a true categorical correlate: whether the research participant (a) received treatment, or (b) was in the control group.

CHAPTER II

METHODS

Participants

Data for this study came from the NIMH Treatment for Depression Collaborative Research Program study (TDCRP; Elkin, Parloff, Hadley, & Autry, 1985). Participants were 239 adults (mean age = 35 years, SD = 8.54 years) diagnosed with depression at the baseline data collection period who subsequently were randomized into one of four treatment conditions (i.e., Cognitive Behavior Therapy / CBT, Interpersonal Therapy / IPT, Imipramine and Clinical Management / IMI-CM, or Placebo and Clinical Management / PLAC-CM) at one of three study sites (George Washington University, the University of Oklahoma, and the University of Pittsburgh). This sample was predominantly Caucasian (89%) and female (70%). Participants were interviewed at multiple time points during treatment, at treatment termination, and at follow-up intervals post-treatment. Because the sample was homogeneous in regards to a depression diagnosis at pre-treatment, our analyses focus on the data collected at termination and follow-up time points.

Measures

Several measures from the TDCRP dataset were used in the present study. The SADS-C data were used to generate diagnostic as well as syndrome measures of depression; in addition, the three simultaneous criteria (Duration of Symptoms, Functioning, and Key Symptoms) were derived from the SADS-C data. Three measures

from the TDCRP were used to generate our dependent variables: The Life Events Interview, the Dysfunctional Attitudes Scale, and the Social Network Form.

Schedule for Affective Disorders and Schizophrenia-Change form. Participants were interviewed with the Schedule for Affective Disorders and Schizophrenia – Change form (SADS-C), a widely used structured clinical interview, at each assessment time point after the baseline assessment (during treatment, at termination, and at follow-up time points) (Spitzer & Endicott, 1978). Both categorical and continuous measures of depression were constructed from data collected with the SADS-C. A categorical / diagnostic variable for MDD was constructed for each individual by determining whether he/she met all DSM-IV criteria for the disorder including the simultaneous criteria (i.e., duration of symptoms, presence of a key symptom, and functioning below 50 on the GAS). A continuous syndrome measure of MDD was constructed by calculating the mean score across SADS-C items assessing DSM-IV MDD symptoms.

Duration of Symptoms. The SADS-C assesses duration of current symptoms as occurring for less than one week, one week, or two weeks or more (with the latter criterion required to receive a diagnosis). A dichotomous duration variable was constructed wherein 1 equaled when symptoms had lasted two weeks or more per the DSM-IV criterion, and 0 equaled when symptoms had lasted less than two weeks. The original SADS-C duration variable also was used as an ordinal index of duration of symptoms.

Functioning. Both dichotomous (greater than 50 versus less than or equal to 50, following the DSM-IV criterion) and continuous GAS scores from the SADS-C were used as indices of functioning.

Key Symptoms. A dichotomous measure was constructed from relevant SADS-C questions to assess whether a key symptom (i.e., either sadness or anhedonia) of MDD was present at a given time point.

Life Events Interview. The Life Events Interview (Elkin et al., 1985) is a measure that assesses the occurrence of significant life events across several categories (e.g., illness of participant, death of someone close to participant, events affecting job of someone in participant's household, events affecting participant's finances), the number of events that have occurred in a given category, and the degree of subjective stress caused by the event. If a participant endorsed a life event as having occurred, they were asked to rate the degree of stress that the event caused on a scale from 1 (not at all stressful) to 5 (extremely stressful). A degree of stress variable was constructed by taking the mean of the degree of stress participants reported across life events and multiplying by the number of life events endorsed.

Dysfunctional Attitudes Scale. The Dysfunctional Attitudes Scale (DAS) is a 21 item measure that was constructed to assess dysfunctional beliefs and schemas (e.g., If a person asks for help, it is a sign of weakness.) thought to be related to depression (Weissman & Beck, 1978; Weissman, 1979). For each item, participants are presented with a stated belief and are asked to indicate the degree to which they agree with the statement. Agreement with the stated belief is rated on a 7-point scale (1 = totally disagree to 7 = totally agree). A total scale score is calculated by summing across all item responses for this scale.

Social Network Form. The Social Network Form is a measure that was developed specifically for the TDCRP protocol (Elkin et al., 1985). This measure assesses the

number of persons with whom participants had contact on a regular basis (at least weekly), the average number of contact hours per week with each person, and the degree of satisfaction the participant feels with regard to these relationships. Satisfaction is rated for each of 7 domains of relationships (relationships with persons within participant's household, with family members outside of household, with participant's children living outside of household, with in-laws, with neighbors regarded as close friends, with coworkers, and with other close friends) on a 5 point scale (1 = very dissatisfied to 5 = very satisfied). Two variables were constructed from this measure: (a) A sum of the weekly contact hours with other persons and (b) a variable that averaged satisfaction ratings across the personal relationships reported.

Treatment Status. A dichotomous variable was constructed to distinguish among participants assigned to treatment (i.e., psychotherapy or pharmacotherapy conditions) versus control conditions (i.e., placebo and clinical management condition).

CHAPTER III

RESULTS

Overview of Analyses

The goals of this study were to determine: (a) whether the categorical diagnostic variable significantly increments prediction of dependent variables over and above prediction by the continuous syndrome scores, and vice versa; and (b) whether the simultaneous diagnostic criteria (i.e., key symptoms; duration of symptoms; adaptive functioning) significantly increment prediction of these dependent variables over and above prediction by the continuous syndrome scores, in order to determine the utility (if any) of the various underlying components of the diagnosis.

We used several continuous dependent variables, including those derived from the Life Events Interview, the DAS, and the Social Network Form. As a categorical dependent variable, we used Treatment Status (i.e., whether the participant was in the control group [0] or in one of the three treatment groups [1]). We used this variable because it represents a true categorical variable; i.e., it was not obtained by dichotomizing a continuous variable. Treatment Status was treated as a dependent variable, as in a discriminant function analysis, in order to allow for assessment of more complex relations (e.g., the interaction between the continuous syndrome measure and the DAS) than would be feasible if Treatment Status were treated as the independent variable. It should be noted that analyses with the Treatment Status variable were conducted only

with data from the treatment termination assessment since treatment effects should grow weaker at time points further from treatment termination.

Descriptive Statistics

Descriptive statistics and bivariate correlations for the predictor and dependent variables are reported in Tables 1 through 4, by time point (i.e., termination, 6 month, 12 month, and 18 month follow up, respectively). Both the categorical diagnostic and the continuous depression scores correlated significantly with all dependent variables, with the exception of the Life Events Inventory Degree of Stress (LEI-DS) variable at the earlier time points. The relation between depression scores and degree of stress became stronger at time points further from treatment termination. Simultaneous diagnostic characteristics were significantly correlated with both depression scores as well as the dependent variables across time points with exception of the LEI-DS variable at the earlier time points. These relations also became stronger at time points further from treatment termination.

Table 1

Means, Standard Deviations, and Correlations for Predictor and Dependent Variables at Treatment Termination Assessment

Measure	Mean	SD	N	2	3	4	5	6	7	8	9
1. Diagnostic	0.10	0.31	155	.72**	.47**	.41**	-.57**	-.02	.26**	-.23**	-.13
2. Continuous	0.28	0.11	155	-	.72**	.46**	-.82**	.03	.48**	-.25**	-.27**
3. Key Symptoms	0.34	0.48	155		-	.32**	-.62**	-.02	.38**	-.16*	-.26**
4. Symptom Duration	1.91	0.90	154			-	-.35**	-.03	.24**	-.11	-.19*
5. Functioning	70.25	10.82	155				-	.12	-.46**	.24**	.39**
6. LEI – DS	7.75	5.47	114					-	-.16	.17	.07
7. DAS	113.29	36.01	154						-	-.25**	-.24**
8. SNF – TH	78.88	63.79	154							-	.11
9. SNF – MS	4.27	0.67	154								-

Note: Diagnostic = DSM-IV MDD diagnostic variable; Continuous = mean of scores on depression items; Key Symptoms = presence of either sadness or anhedonia; Symptom Duration = duration of current symptoms (1 = < 1 wk; 2 = 1 wk; 3 = 2 wks or more); Functioning = GAS rating; LEI – DS = Life Events Inventory degree of stress; DAS = Dysfunctional Attitude Scale Total Score; SNF – TH = Social Network Form, Total Contact Hours; SNF – MS = Social Network Form, mean satisfaction.

* $p < .05$, ** $p < .01$.

Table 2

Means, Standard Deviations, and Correlations for Predictor and Dependent Variables at 6 Month Follow-Up Assessment

Measure	Mean	SD	N	2	3	4	5	6	7	8	9
1. Diagnostic	0.17	0.37	192	.74**	.58**	.53**	-.56**	.18*	.32**	-.15*	-.22**
2. Continuous	0.30	0.13	192	-	.77**	.49**	-.84**	.09	.38**	-.21**	-.33**
3. Key Symptoms	0.38	0.49	192		-	.45**	-.72**	.08	.33**	-.14	-.30**
4. Symptom Duration	1.91	0.92	192			-	-.45**	-.06	.25**	-.08	-.22**
5. Functioning	69.18	11.96	192				-	-.02	-.41**	.18*	.36**
6. LEI – DS	8.72	6.26	148					-	.09	.11	.03
7. DAS	113.02	36.67	191						-	-.15*	-.24**
8. SNF – TH	81.50	70.18	187							-	.03
9. SNF – MS	4.16	0.66	188								-

Note: Diagnostic = DSM-IV MDD diagnostic variable; Continuous = mean of scores on depression items; Key Symptoms = presence of either sadness or anhedonia; Symptom Duration = duration of current symptoms (1 = < 1 wk; 2 = 1 wk; 3 = 2 wks or more); Functioning = GAS rating; LEI – DS = Life Events Inventory degree of stress; DAS = Dysfunctional Attitude Scale Total Score; SNF – TH = Social Network Form, Total Contact Hours; SNF – MS = Social Network Form, mean satisfaction.

* $p < .05$, ** $p < .01$.

Table 3

Means, Standard Deviations, and Correlations for Predictor and Dependent Variables at 12 Month Follow-Up Assessment

Measure	Mean	SD	N	2	3	4	5	6	7	8	9
1. Diagnostic	0.19	0.39	186	.70**	.56**	.51**	-.62**	.09	.36**	-.04	-.31**
2. Continuous	0.31	0.13	186	-	.72**	.51**	-.85**	.19*	.47**	-.16*	-.47**
3. Key Symptoms	0.42	0.50	186		-	.49**	-.70**	.12	.42**	-.06	-.39**
4. Symptom Duration	2.03	0.91	186			-	-.54**	-.06	.19*	-.06	-.20**
5. Functioning	63.33	12.01	186				-	-.12	-.48**	.24**	.54**
6. LEI – DS	9.03	6.36	150					-	.16	.11	-.04
7. DAS	112.58	36.15	187						-	-.15	-.34**
8. SNF – TH	71.98	60.16	183							-	.19**
9. SNF – MS	4.14	0.73	184								-

Note: Diagnostic = DSM-IV MDD diagnostic variable; Continuous = mean of scores on depression items; Key Symptoms = presence of either sadness or anhedonia; Symptom Duration = duration of current symptoms (1 = < 1 wk; 2 = 1 wk; 3 = 2 wks or more); Functioning = GAS rating; LEI – DS = Life Events Inventory degree of stress; DAS = Dysfunctional Attitude Scale Total Score; SNF – TH = Social Network Form, Total Contact Hours; SNF – MS = Social Network Form, mean satisfaction.

* $p < .05$, ** $p < .01$.

Table 4

Means, Standard Deviations, and Correlations for Predictor and Dependent Variables at 18 Month Follow-Up Assessment

Measure	Mean	SD	N	2	3	4	5	6	7	8	9
1. Diagnostic	0.12	0.32	191	.71**	.46**	.48**	-.53**	.20*	.19**	-.11	-.18*
2. Continuous	0.29	0.12	191	-	.74**	.50**	-.83**	.24**	.41**	-.20**	-.43**
3. Key Symptoms	0.38	0.49	191		-	.41**	-.66**	.20*	.31**	-.15*	-.37**
4. Symptom Duration	1.86	0.86	191			-	-.42**	.11	.15*	-.16*	-.13
5. Functioning	69.22	10.61	191				-	-.24**	-.39**	.23**	.49**
6. LEI – DS	10.09	8.07	149					-	.18*	.20*	-.07
7. DAS	111.33	33.87	191						-	-.24**	-.37**
8. SNF – TH	66.29	53.32	187							-	.18*
9. SNF – MS	4.16	0.70	189								-

Note: Diagnostic = DSM-IV MDD diagnostic variable; Continuous = mean of scores on depression items; Key Symptoms = presence of either sadness or anhedonia; Symptom Duration = duration of current symptoms (1 = < 1 wk; 2 = 1 wk; 3 = 2 wks or more); Functioning = GAS rating; LEI – DS = Life Events Inventory degree of stress; DAS = Dysfunctional Attitude Scale Total Score; SNF – TH = Social Network Form, Total Contact Hours; SNF – MS = Social Network Form, mean satisfaction.

* $p < .05$, ** $p < .01$.

1A. Effects on the Categorical Dependent Variable (Treatment Status) at the Treatment Termination Assessment

A series of categorical logit models (Agresti, 1996) were used to determine whether there were significant total and unique effects of the categorical diagnostic measure and continuous syndrome scores on Treatment Status (i.e., whether there were significant differences between participants in a treatment versus control group vis-à-vis the diagnostic or syndrome depression measures). Although total effects were not significant for either the diagnostic or continuous depression scores (i.e., when used as individual predictors), effects of both variables were significant when both variables were entered into the model. Table 5 summarizes these results.

A second series of categorical logit models were conducted to determine whether the simultaneous criteria significantly increased prediction over and above prediction by the continuous syndrome scores. In this series of analyses, both continuous syndrome scores and specific components of the categorical diagnostic measure (i.e., Key Symptoms, Symptom Duration, or GAS Scores) were used to predict Treatment Status. The interactions between continuous syndrome scores and the simultaneous diagnostic criteria also were included and were the effects of primary interest because a diagnosis is in essence an interaction between levels of symptoms and simultaneous diagnostic criteria.

As noted above, the total effect of continuous depression scores was not significant; however, when both continuous depression scores and Key Symptoms were entered into the model, the effects of both variables and their interaction were significant. Similar effects were found when Symptom Duration was added to the prediction model. When GAS scores were added to the model, the effects of continuous depression scores

and GAS scores were not significant; however, their interaction was significant. Table 5 summarizes these results.

Table 5
Effects on the Categorical Dependent Variable (Treatment Status) at the Treatment Termination Assessment

Model	*DF*	*Chi-Square*	*p*
Model 1			
Diagnostic Depression	1	0.89	.35
Model 2			
Continuous Depression	1	2.24	.13
Model 3			
Diagnostic Depression	1	6.83	.01
Continuous Depression	1	9.07	.003
Model 4			
Continuous Depression	1	4.84	.03
Key Symptoms	1	7.17	.01
Continuous * Key	1	8.48	.004
Model 5			
Continuous Depression	1	2.95	.09
GAS	1	3.34	.07
Continuous * GAS	1	3.93	.05
Model 6			
Continuous Depression	1	4.53	.03
Symptom Duration	1	4.60	.03
Continuous * Duration	1	3.85	.05

Note: GAS = SADS-C Global Assessment of Functioning Scale; Symptom Duration = SADS-C duration of symptoms variable

To understand the significant interaction effects found between continuous depression scores and simultaneous diagnostic characteristics in the prediction of Treatment Status, we conducted logit models regressing Treatment Status onto continuous depression scores separately (1) for groups endorsing Key Symptoms and for groups not endorsing Key Symptoms; (2) for groups with GAS scores less than 50 and for groups with GAS scores greater than or equal to 50; and (3) for groups with Symptom Duration less than two weeks and for groups with Symptom Duration greater than or equal to two weeks. We expected that continuous symptom scores would be more strongly related to Treatment Status when simultaneous diagnostic criteria were present

than when they were absent. Interestingly, we found the opposite pattern of relations: we found that continuous symptom scores were significantly and positively related to Treatment Status when simultaneous diagnostic characteristics were absent but not when they were present. Table 6 summarizes these results.

Table 6
Effects on the Categorical Dependent Variable (Treatment Status) at the Treatment Termination Assessment by Simultaneous Diagnostic Criterion Group

Model	*DF*	*Chi-Square*	*p*
Key Symptoms Absent			
Continuous Depression	1	9.15	.003
Key Symptoms Present			
Continuous Depression	1	0.44	.51
GAS ≥ 50			
Continuous Depression	1	4.24	.04
GAS < 50			
Continuous Depression	1	1.53	.22
Symptom Duration < 2 wks.			
Continuous Depression	1	5.63	.02
Symptom Duration ≥ 2 wks.			
Continuous Depression	1	0.46	.50

Note: Continuous Depression = continuous depression symptom scores; GAS = Global Assessment Score

1B. Effects on Continuous Dependent Variables at Treatment Termination Assessment

A series of linear regressions were performed to examine the differential prediction of the continuous dependent variables by the diagnostic versus continuous depression scores, at the treatment termination assessment. As can be seen in the summary of these results in Table 7, neither the diagnostic nor the continuous depression scores were significantly related to the Life Events Inventory – Degree of Stress (LEI-DS) dependent variable, either as total or unique effects (i.e., when entered alone or when entered simultaneously as predictors in the regression models). Total effects for both the diagnostic and continuous depression scores were significant for the Dysfunctional

Attitudes Scale (DAS) total score, but only the unique effect for the continuous depression scores was significant; i.e., the categorical diagnostic variable did not significantly increase the prediction over and above that of the continuous measure. For the Social Network Form – Total Hours dependent variable (SNF-TH), total effects for both the diagnostic and continuous depression scores were significant but unique effects were not; i.e., neither depression variable incremented the prediction of the other. For the Social Network Form – Mean Satisfaction dependent variable (SNF-MS), the diagnostic depression scores did not account for a significant amount of variability in this dependent variable, although the total and unique effects of the continuous depression scores were significant. In sum, unique effects for the diagnostic measure were non-significant for all of the four continuous measures; i.e., it did not increment prediction of the continuous syndrome measure.

Table 7
Effects of Diagnostic and Continuous Depression on the Continuous Dependent Variables at Treatment Termination Assessment

Outcome	*Predictor*		*F*	*R-Square*	*p*
LEI – DS	Diagnostic	df=1	0.03	.000	.86
	Continuous	df=1	0.12	.001	.73
	Model	df=2	0.26	.005	.77
		Diagnostic	0.41		.52
		Continuous	0.49		.49
DAS	Diagnostic	df=1	11.06	.07	.00
	Continuous	df=1	45.17	.23	.00
	Model	df=2	24.33	.24	.00
		Diagnostic	2.92		.09
		Continuous	35.12		.00
SNF – TH	Diagnostic	df=1	8.78	.05	.005
	Continuous	df=1	10.38	.06	.005
	Model	df=2	5.63	.07	.00
		Diagnostic	0.88		.35
		Continuous	2.40		.12
SNF – MS	Diagnostic	df=1	2.75	.02	.10
	Continuous	df=1	11.70	.07	.00
	Model	df=2	6.48	.08	.00
		Diagnostic	1.24		.27
		Continuous	10.04		.00

Note: Diagnostic = DSM-IV MDD diagnostic variable; Continuous = mean of scores on depression items; Both = simultaneous entrance of both categorical and continuous depression scores into the regression equation; LEI – DS = Life Events Inventory degree of stress; DAS = Dysfunctional Attitude Scale Total Score; SNF – TH = Social Network Form, Total Contact Hours; SNF – MS = Social Network Form, mean satisfaction.

Another series of linear regression analyses were conducted with data from this time point to examine whether the addition of the simultaneous diagnostic characteristics to continuous depression scores significantly increased the prediction of the continuous dependent variables. Table 8 summarizes the findings from this series of analyses.

Table 8
Effects of Continuous Depression Scores and Simultaneous Diagnostic Characteristics on Continuous Dependent Variables at Treatment Termination Assessment

Outcome	*Predictors*	*F*	*p*
LEI – DS	Continuous	0.62	.43
	Key Symptoms	0.56	.46
	Continuous	0.28	.60
	Symptom Duration	0.25	.62
	Continuous	6.62	.01
	GAS	8.13	.01
DAS	Continuous	16.81	.00
	Key Symptoms	0.49	.49
	Continuous	33.69	.00
	Symptom Duration	0.11	.74
	Continuous	6.55	.01
	GAS	2.36	.13
SNF – TH	Continuous	6.50	.01
	Key Symptoms	0.19	.66
	Continuous	8.51	.00
	Symptom Duration	0.01	.94
	Continuous	1.73	.19
	GAS	0.45	.50
SNF – MS	Continuous	2.30	.13
	Key Symptoms	1.46	.23
	Continuous	6.26	.01
	Symptom Duration	1.08	.30
	Continuous	1.23	.27
	GAS	15.24	.00

Note: LEI – DS = Life Events Inventory – Degree of Stress; DAS = Dysfunctional Attitudes Scale; SNF – TH = Social Network Form – Total Hours; and SNF – MS = Social Network Form – Mean Satisfaction

For all four dependent variables, unique effects (i.e., controlling for continuous depression scores) for (a) Key Symptoms or (b) Symptom Duration were not significant. However, for the LEI-DS and SNF-MS dependent variables, the unique effect of GAS scores was significant.

A final series of linear regression analyses were conducted at this time point to examine whether interaction effects between continuous depression scores and simultaneous diagnostic characteristics incremented prediction of continuous dependent variables at the termination assessment. Table 9 summarizes the findings from this series of analyses. As can be seen in Table 9, all interactions effects were non-significant.

Table 9
Interaction Effects (Controlling for Component Main Effects) on Continuous Dependent Variables at Treatment Termination Assessment

Outcome	*Predictors*	*F*	*p*
LEI – DS	Cont * Key Symptoms	2.47	.12
	Cont * Symptom Duration	0.47	.50
	Cont * GAS	1.26	.26
DAS	Cont * Key Symptoms	1.86	.17
	Cont * Symptom Duration	1.12	.29
	Cont * GAS	0.80	.37
SNF – TH	Cont * Key Symptoms	0.76	.39
	Cont * Symptom Duration	0.84	.36
	Cont * GAS	0.08	.77
SNF – MS	Cont * Key Symptoms	0.49	.48
	Cont * Symptom Duration	0.26	.61
	Cont * GAS	2.05	.15

Note: LEI – DS = Life Events Inventory – Degree of Stress; DAS = Dysfunctional Attitudes Scale; SNF – TH = Social Network Form – Total Hours; SNF – MS = Social Network Form – Mean Satisfaction; and Cont = continuous depression scores

2. Effects on Continuous Dependent Variables at 6 Month Follow Up Assessment

Paralleling the 1B analyses (above) conducted at the Treatment Termination assessment, another series of linear regressions were conducted to examine the differential prediction of the continuous dependent variables by the diagnostic versus continuous depression scores using data from the 6 month follow-up assessment. Table 10 summarizes the results of these analyses.

Table 10

Effects of Diagnostic and Continuous Depression on Continuous Dependent Variables at 6 Month Follow-Up Assessment

Outcome	*Predictor*		*F*	*R-Square*	*P*
LEI – DS	Diagnostic	df=1	4.76	.03	.03
	Continuous	df=1	1.25	.01	.27
	Model	df=2	2.72	.04	.07
		Diagnostic	4.16		.04
		Continuous	0.68		.41
DAS	Diagnostic	df=1	21.65	.10	.00
	Continuous	df=1	32.66	.15	.00
	Model	df=2	16.63	.15	.00
		Diagnostic	0.66		.42
		Continuous	10.52		.00
SNF – TH	Diagnostic	df=1	4.37	.02	.04
	Continuous	df=1	8.48	.04	.00
	Model	df=2	4.23	.04	.02
		Diagnostic	0.03		.87
		Continuous	4.01		.05
SNF – MS	Diagnostic	df=1	10.17	.05	.00
	Continuous	df=1	22.88	.11	.00
	Model	df=2	11.55	.11	.00
		Diagnostic	0.30		.59
		Continuous	12.31		.00

Note: Diagnostic = DSM-IV MDD diagnostic variable; Continuous = mean of scores on depression items; Both = simultaneous entrance of both categorical and continuous depression scores into the regression equation; LEI – DS = Life Events Inventory degree of stress; DAS = Dysfunctional Attitude Scale Total Score; SNF – TH = Social Network Form, Total Contact Hours; SNF – MS = Social Network Form, mean satisfaction.

As can be seen in the summary of results in Table 10, significant total and unique effects of the categorical diagnostic depression scores, but not for the continuous depression scores, were found for the dependent variable, LEI-DS. However, for the other dependent variables (DAS, SNF-TH, and SNF-MS), the total effects of both diagnostic and continuous depression scores were significant, but significant unique effects were found only for the continuous depression scores. In sum, the diagnostic scores incremented prediction of the continuous syndrome scores for one of four dependent variables, whereas the continuous syndrome scores incremented prediction of the diagnostic scores for three of four dependent variables.

A second series of linear regression analyses were performed with data from the six-month follow-up assessment to examine whether the addition of simultaneous diagnostic characteristics to continuous depression scores incremented prediction of the continuous dependent variables. Table 11 summarizes results from this series of analyses.

Table 11
Effects of Continuous Depression Scores and Simultaneous Diagnostic Characteristics at 6 Month Follow-Up Assessment

Outcome	*Predictors*	*F*	*p*
LEI – DS	Continuous	0.33	.57
	Key Symptoms	0.01	.91
	Continuous	3.35	.07
	Symptom Duration	2.61	.11
	Continuous	2.88	.09
	GAS	1.69	.20
DAS	Continuous	8.64	.00
	Key Symptoms	0.79	.37
	Continuous	20.17	.00
	Symptom Duration	0.89	.35
	Continuous	1.23	.27
	GAS	5.74	.02
SNF – TH	Continuous	4.77	.03
	Key Symptoms	0.21	.65
	Continuous	7.32	.01
	Symptom Duration	0.15	.70
	Continuous	1.90	.17
	GAS	0.05	.83
SNF – MS	Continuous	5.40	.02
	Key Symptoms	0.76	.38
	Continuous	13.76	.00
	Symptom Duration	0.67	.41
	Continuous	0.68	.41
	GAS	4.39	.04

Note: LEI – DS = Life Events Inventory – Degree of Stress; DAS = Dysfunctional Attitudes Scale; SNF – TH = Social Network Form – Total Hours; and SNF – MS = Social Network Form – Mean Satisfaction

When simultaneous diagnostic characteristics were added to continuous depression scores in the model predicting LEI-DS scores, no significant unique effects

were found for the simultaneous diagnostic characteristics. When simultaneous diagnostic characteristics were included with continuous depression scores in the regression models looking at the DAS dependent variable, only GAS scores showed a significant unique effect. Further, the unique effect of continuous depression scores was not significant when GAS scores were controlled for in this model. Unique effects of simultaneous diagnostic characteristics were not significant with respect to the dependent variable, SNF-TH, controlling for the effects of continuous depression scores. Significant unique effects of GAS scores and continuous depression scores were found in the model predicting the SNF-MS dependent variable. No unique effects were found for the simultaneous diagnostic characteristics, Key Symptoms and Symptom Duration, when controlling for continuous depression scores in the models predicting SNF-MS scores.

A final series of linear regression analyses were conducted at this time point to examine whether interaction effects between continuous depression scores and simultaneous diagnostic characteristics incremented prediction of continuous dependent variables at the 6 month follow up assessment. Table 12 summarizes the findings from this series of analyses. As can be seen in Table 12, one interaction effect, between continuous depression scores and Key Symptoms with respect to the dependent variable SNF-MS was significant.

Table 12
Interaction Effects (Controlling for Component Main Effects) on Continuous Dependent Variables at 6 Month Follow-Up Assessment

Outcome	*Predictors*	*F*	*p*
LEI – DS	Cont * Key Symptoms	1.92	.17
	Cont * Symptom Duration	1.10	.30
	Cont * GAS	1.06	.31
DAS	Cont * Key Symptoms	0.08	.78
	Cont * Symptom Duration	0.82	.37
	Cont * GAS	0.15	.70
SNF – TH	Cont * Key Symptoms	0.38	.54
	Cont * Symptom Duration	0.22	.64
	Cont * GAS	1.73	.19
SNF – MS	Cont * Key Symptoms	4.44	.04
	Cont * Symptom Duration	0.98	.32
	Cont * GAS	3.29	.07

Note: LEI – DS = Life Events Inventory – Degree of Stress; DAS = Dysfunctional Attitudes Scale; SNF – TH = Social Network Form – Total Hours; SNF – MS = Social Network Form – Mean Satisfaction; and Cont = continuous depression scores

3. Effects on Continuous Dependent Variables at 12 Month Follow Up Assessment

Results from a series of regression analyses performed to examine the differential prediction of the continuous dependent variables by the diagnostic versus continuous depression scores, using data from the 12 month follow-up assessment, are presented in Table 13.

Table 13

Effects of Diagnostic and Continuous Depression on Continuous Dependent Variables at 12 Month Follow-Up Assessment

Outcome	*Predictor*		*F*	*R-Square*	*P*
LEI – DS	Diagnostic	df=1	1.14	.01	.29
	Continuous	df=1	5.42	.04	.02
	Model	df=2	3.05	.04	.05
		Diagnostic	0.70		.40
		Continuous	4.93		.03
DAS	Diagnostic	df=1	27.22	.13	.00
	Continuous	df=1	51.77	.22	.00
	Model	df=2	26.04	.22	.00
		Diagnostic	0.47		.50
		Continuous	21.77		.00
SNF – TH	Diagnostic	df=1	.23	.00	.63
	Continuous	df=1	4.53	.02	.03
	Model	df=2	3.19	.03	.04
		Diagnostic	1.83		.18
		Continuous	6.14		.01
SNF – MS	Diagnostic	df=1	18.79	.09	.00
	Continuous	df=1	50.81	.22	.00
	Model	df=2	25.34	.22	.00
		Diagnostic	0.11		.74
		Continuous	28.97		.00

Note: Diagnostic = DSM-IV MDD diagnostic variable; Continuous = mean of scores on depression items; Both = simultaneous entrance of both categorical and continuous depression scores into the regression equation; LEI – DS = Life Events Inventory degree of stress; DAS = Dysfunctional Attitude Scale Total Score; SNF – TH = Social Network Form, Total Contact Hours; SNF – MS = Social Network Form, mean satisfaction.

As can be seen in the summary of results in Table 13, there was a significant total effect for continuous depression scores but not for diagnostic depression scores with respect to the dependent variables, LEI-DS and SNF-TH. Both continuous and diagnostic depression scores had significant total effects with respect to the dependent variables, DAS and SNF-MS. Significant unique effects were found for the continuous depression scores for all four dependent variables, but no unique effects were significant for the diagnostic depression scores.

Table 14 summarizes regression results when the prediction of continuous dependent variables was incremented with inclusion of the simultaneous diagnostic characteristics to the continuous depression scores, at the 12 month follow-up

assessment. As can be seen in Table 14, a significant unique effect was found for Symptom Duration on LEI-DS scores, controlling for continuous depression scores. Significant unique effects were found also for GAS scores with respect to the DAS, SNF-TH, and SNF-MS dependent variables, when controlling for continuous depression scores.

Table 14
Effects of Continuous Depression Scores and Simultaneous Diagnostic Characteristics on Continuous Dependent Variables at 12 Month Follow-Up Assessment

Outcome	*Predictors*	*F*	*p*
LEI – DS	Continuous	3.23	.07
	Key Symptoms	0.07	.80
	Continuous	11.00	.00
	Symptom Duration	5.91	.02
	Continuous	4.02	.05
	GAS	0.90	.35
DAS	Continuous	13.18	.00
	Key Symptoms	3.26	.07
	Continuous	44.50	.00
	Symptom Duration	1.01	.32
	Continuous	2.84	.09
	GAS	5.99	.02
SNF – TH	Continuous	5.14	.02
	Key Symptoms	1.21	.27
	Continuous	3.89	.05
	Symptom Duration	0.10	.75
	Continuous	1.40	.24
	GAS	7.46	.01
SNF – MS	Continuous	17.26	.00
	Key Symptoms	1.13	.29
	Continuous	42.22	.00
	Symptom Duration	0.62	.43
	Continuous	0.02	.88
	GAS	19.87	.00

Note: LEI – DS = Life Events Inventory – Degree of Stress; DAS = Dysfunctional Attitudes Scale; SNF – TH = Social Network Form – Total Hours; and SNF – MS = Social Network Form – Mean Satisfaction

A final series of linear regression analyses were conducted at this time point to examine whether interaction effects between continuous depression scores and

simultaneous diagnostic characteristics incremented prediction of the continuous dependent variables at the 12 month follow up assessment. Table 15 summarizes the findings from this series of analyses. As can be seen in Table 14, all interaction effects were non-significant.

Table 15
Interaction Effects (Controlling for Component Main Effects) on Continuous Dependent Variables at 12 Month Follow-Up Assessment

Outcome	*Predictors*	*F*	*p*
LEI – DS	Cont * Key Symptoms	0.34	.56
	Cont * Symptom Duration	0.21	.65
	Cont * GAS	0.01	.92
DAS	Cont * Key Symptoms	2.73	.10
	Cont * Symptom Duration	0.11	.74
	Cont * GAS	2.42	.12
SNF – TH	Cont * Key Symptoms	1.39	.24
	Cont * Symptom Duration	0.41	.52
	Cont * GAS	0.01	.92
SNF – MS	Cont * Key Symptoms	3.30	.07
	Cont * Symptom Duration	2.92	.09
	Cont * GAS	0.36	.55

Note: LEI – DS = Life Events Inventory – Degree of Stress; DAS = Dysfunctional Attitudes Scale; SNF – TH = Social Network Form – Total Hours; SNF – MS = Social Network Form – Mean Satisfaction; and Cont = continuous depression scores

4. Effects on Continuous Dependent Variables at 18 Month Follow Up Assessment

Table 16 summarizes results of a series of linear regressions examining the prediction of the continuous dependent variables by the diagnostic versus continuous depression scores. As can be seen in Table 16, significant total effects were found for diagnostic and continuous depression scores on the dependent variables LEI-DS, DAS, and SNF-MS; a significant total effect was found for continuous depression scores but not diagnostic depression scores on the fourth dependent variable, SNF-TH. Neither diagnostic nor continuous depression scores had significant unique effects on the

dependent variable, LEI-DS, both diagnostic and continuous depression scores had significant unique effects on the dependent variables DAS and SNF-MS, and only the continuous depression scores had significant unique effects on the SNF-TH dependent variable.

Table 16
Effects of Diagnostic and Continuous Depression on Continuous Dependent Variables at 18 Month Follow-Up Assessment

Outcome	*Predictor*		*F*	*R-Square*	*P*
LEI – DS	Diagnostic	df=1	6.22	.04	.01
	Continuous	df=1	8.69	.06	.00
	Model	df=2	4.47	.06	.01
		Diagnostic	0.28		.60
		Continuous	2.64		.11
DAS	Diagnostic	df=1	6.86	.04	.01
	Continuous	df=1	37.55	.17	.00
	Model	df=2	21.72	.19	.00
		Diagnostic	5.07		.03
		Continuous	35.32		.00
SNF – TH	Diagnostic	df=1	2.44	.01	.12
	Continuous	df=1	7.32	.04	.01
	Model	df=2	3.79	.04	.02
		Diagnostic	0.28		.60
		Continuous	5.08		.03
SNF – MS	Diagnostic	df=1	6.22	.03	.01
	Continuous	df=1	42.47	.19	.00
	Model	df=2	25.82	.22	.00
		Diagnostic	7.66		.01
		Continuous	44.00		.00

Note: Diagnostic = DSM-IV MDD diagnostic variable; Continuous = mean of scores on depression items; Both = simultaneous entrance of both categorical and continuous depression scores into the regression equation; LEI – DS = Life Events Inventory degree of stress; DAS = Dysfunctional Attitude Scale Total Score; SNF – TH = Social Network Form, Total Contact Hours; SNF – MS = Social Network Form, mean satisfaction.

Table 17 summarizes results of a series of regression analyses examining whether the addition of simultaneous diagnostic characteristics to continuous depression scores increments the prediction of the continuous dependent variables, using data from the 18 month follow-up assessment. As can be seen in Table 17, the only significant unique

effect of the simultaneous diagnostic characteristics was for GAS scores on the dependent variable, SNF-MS, controlling for continuous depression scores.

Table 17
Effects of Continuous Depression Scores and Simultaneous Diagnostic Characteristics on Continuous Dependent Variables at 18 Month Follow-Up Assessment

Outcome	*Predictors*	*F*	*p*
LEI – DS	Continuous	2.79	.10
	Key Symptoms	0.11	.74
	Continuous	7.07	.01
	Symptom Duration	0.15	.70
	Continuous	0.72	.40
	GAS	0.91	.34
DAS	Continuous	16.37	.00
	Key Symptoms	0.02	.90
	Continuous	33.65	.00
	Symptom Duration	1.01	.32
	Continuous	5.51	.02
	GAS	1.87	.17
SNF – TH	Continuous	2.90	.09
	Key Symptoms	0.03	.85
	Continuous	3.37	.07
	Symptom Duration	0.91	.34
	Continuous	0.04	.83
	GAS	2.61	.11
SNF – MS	Continuous	11.96	.00
	Key Symptoms	1.71	.19
	Continuous	41.03	.00
	Symptom Duration	2.29	.13
	Continuous	0.38	.54
	GAS	14.96	.00

Note: LEI – DS = Life Events Inventory – Degree of Stress; DAS = Dysfunctional Attitudes Scale; SNF – TH = Social Network Form – Total Hours; and SNF – MS = Social Network Form – Mean Satisfaction

A final series of linear regression analyses were conducted at this time point to examine whether interaction effects between continuous depression scores and simultaneous diagnostic characteristics incremented prediction of continuous dependent variables at the 18 month follow up assessment. Table 18 summarizes the findings from this series of analyses. As can be seen in Table 18, the only significant interaction effect

was between continuous depression scores and Duration, with respect to the dependent variable SNF-MS.

Table 18
Interaction Effects (Controlling for Component Main Effects) on Continuous Dependent Variables at 18 Month Follow-Up Assessment

Outcome	*Predictors*	*F*	*p*
LEI – DS	Cont * Key Symptoms	0.02	.89
	Cont * Symptom Duration	0.00	1.00
	Cont * GAS	2.42	.12
DAS	Cont * Key Symptoms	0.03	.87
	Cont * Symptom Duration	0.38	.54
	Cont * GAS	0.13	.71
SNF – TH	Cont * Key Symptoms	0.00	.97
	Cont * Symptom Duration	0.01	.94
	Cont * GAS	0.08	.78
SNF – MS	Cont * Key Symptoms	1.02	.31
	Cont * Symptom Duration	4.85	.03
	Cont * GAS	0.03	.87

Note: LEI – DS = Life Events Inventory – Degree of Stress; DAS = Dysfunctional Attitudes Scale; SNF – TH = Social Network Form – Total Hours; SNF – MS = Social Network Form – Mean Satisfaction; and Cont = continuous depression scores

5. Summary of Results

Several series of analyses were conducted to examine the incremental prediction of diagnostic versus continuous depression scores. Another series of analyses were conducted to determine what, if any, aspects of a diagnosis might add to continuous depression scores' prediction of categorical and continuous dependent variables. A final series of analyses were conducted to examine whether interaction effects among continuous depression scores and simultaneous diagnostic characteristics would increment prediction of continuous dependent variables.

In the first series of analyses, Treatment Status was used as the dependent variable in a series of logit models. Both the diagnostic and continuous depression scores significantly incremented prediction over and above the other predictor, indicating that at

least for these data, the diagnostic measure provided unique information above that contained in the continuous depression measure. To try to determine what aspect of the diagnosis might underlay this effect, we conducted several logit analyses wherein we added each simultaneous diagnostic characteristic to the continuous depression scores in the prediction of Treatment Status. Significant effects were found for Key Symptoms, Symptom Duration, and the interactions between these variables and continuous depression scores. This suggests that the unique effect of the diagnostic information may be a function of the Key Symptoms and Symptom Duration criteria.

A series of linear regressions were conducted at each time point to examine the differential prediction of continuous dependent variables by diagnostic versus continuous depression scores. Although, as expected, results varied somewhat across time points, we found that the continuous syndrome scores significantly incremented the diagnostic scores 12 times whereas the diagnostic scores significantly incremented the continuous syndrome scores only 3 times, out of 16 total tests. This suggests that, at least in regards to the prediction of these continuous dependent variables, the diagnostic variable contributes relatively little unique information.

Another series of linear regressions were conducted at each time point to examine what, if any, aspects of the diagnosis might increment prediction of the continuous dependent variables. As expected, these results also varied somewhat across time points. However, results showed that the GAS simultaneous criterion scores and Symptom Duration sometimes incremented prediction of LEI-DS scores when controlling for continuous depression scores. GAS scores and continuous depression scores were the best predictors of DAS scores. In two cases, GAS but not continuous depression scores

had a significant unique effect on DAS scores. And, in two cases, continuous depression scores but not GAS scores had a significant unique effect on DAS scores. Continuous depression scores had significant total and unique effects, controlling for simultaneous diagnostic characteristics, on SNF-TH scores in all but one instance. In one instance, GAS scores but not continuous depression scores had a significant unique effect on SNF-TH scores. GAS scores, but not continuous depression scores, had significant total and unique effects on SNF-MS scores in all except one instance. In that case, GAS and continuous depression scores both had significant total and unique effects on SNF-MS scores.

A final series of linear regression analyses examined whether interactions between continuous depression scores and simultaneous diagnostic characteristics incremented the prediction of continuous dependent variables, controlling for the main effects of continuous depression scores and simultaneous diagnostic characteristics. Out of 48 interaction tests, only two were significant. At the 6 month follow up assessment, there was a significant interaction effect between continuous depression scores and Key Symptoms on the dependent variable, SNF-MS, and at the 18 month follow up assessment, there was a significant interaction effect between continuous depression scores and Duration on the dependent variable, SNF-MS. Thus, overall, the interaction effects between the continuous syndrome scores and the simultaneous diagnostic criteria – which we have argued represent the potential source of incremental utility to diagnoses – added little to prediction of the continuous dependent variables.

CHAPTER IV

DISCUSSION

Historically, diagnoses have been valued for a number of reasons, the primary of which is the belief that they provide information that is unique – as well as meaningful – about psychopathology over and above information obtained in a single syndrome score. That is, diagnoses are believed to identify groups that are unique, relative to groups that might be identified by a syndrome score. Consequently, diagnoses have and continue to be frequently used (a) to identify and select samples (e.g., Elkin et al., 1985; Weertman, Arntz, Schouten, & Dreessen, 2005; Chard, 2005); (b) as outcomes in treatment studies (e.g., Chard, 2005; Bockting, Schene, Spinhoven, Koeter, Wouters, Huyser, Kamphuis, & The DELTA Study Group, 2005); (c) as predictors of outcomes in treatment studies (e.g., Weertman et al., 2005; Bockting et al., 2005); and (d) as the independent variable in descriptive psychopathology studies (e.g., Amir, Beard, Przeworski, 2005; Rohde, Lewinsohn, & Klein, 2005). Still, the utility and validity of diagnoses is not universally accepted, for a number of reasons. For example, it has been argued (e.g., Widiger & Samuel, 2005) that they are a product of arbitrary decision rules rather than reflecting inherent aspects of psychopathology, and hence do not provide unique meaningful information. In addition, it is well known (e.g., Cohen, 1983; MacCallum et al., 2002; Maxwell & Delaney, 1993) that dichotomizing a continuous measure (a fundamental part of establishing a diagnosis) is associated with a number of negative statistical consequences.

However, diagnoses are more complex than a simple dichotomization of a single continuous symptom measure. DSM diagnoses (APA, 1994), for example, often involve time duration, functional impairment, and key symptom criteria, in addition to a dichotomous symptom frequency cut point criterion. Thus, the utility and validity of diagnoses versus continuous symptom measures of psychopathology may be more complex than the analogue question, which is relatively settled, regarding the negative consequences of dichotomization. Although a number of studies have evaluated the utility and validity of certain aspects of diagnoses, to date there appear to have been no studies that have more fully investigated their utility and validity by assessing these other simultaneous criteria. Furthermore, some of the previous substantive studies evaluating the relative predictive utility of categorical diagnostic measures versus continuous syndrome measures of psychopathology have had inherent problems, such as using different instruments to generate the syndrome scores and diagnoses (e.g., CBCL and DISC, respectively; Fergusson & Horwood, 1995; Jensen et al., 1996; Ferdinand et al., 2004) and thus different questionnaire or interview items to derive the diagnoses and continuous syndrome measures, making it impossible to tell if differences in predictive utility are a function of the diagnoses versus the continuous syndrome scores or the measures' content.

In the present study, we sought to address a number of questions regarding the relative utility and validity of diagnoses versus continuous symptom measures of psychopathology. We improved upon the methodology of previous work by deriving diagnoses and continuous symptom scores from the same SADS-C questions to avoid the confounding effect of content with diagnoses vs. syndromes. Additionally, we separately

examined the utility of aspects of a diagnosis that distinguishes it from a simple dichotomized symptom score – the simultaneous diagnostic criteria – which also has not been done in previous work.

We had several specific analytic goals. First, we were interested in examining the predictive utility of depression diagnoses and continuous depression symptom scores to predict both categorical and continuous dependent variables. Second, we were interested in examining whether components of depression diagnoses (i.e., the simultaneous diagnostic criteria – Key Symptoms, Symptom Duration, and GAS Scores) added unique variance to the prediction of the dependent variables, above the continuous measure of symptoms. This would indicate whether there was any utility to the underlying aspects of diagnoses that differentiate them from a simple dichotomization. And third, because when they are used in a diagnosis these simultaneous criteria represent interactions (i.e., all three simultaneous criteria must be met, or the diagnosis cannot be given), we were interested in determining if there were unique effects for the interactions between the continuous symptom measure and the simultaneous diagnostic criteria in predicting dependent variables. In this study, we focused on depression because (a) it is one of the most common forms of psychopathology; and (b) diagnoses are frequently used in the depression research literature.

Overall, although there clearly are limitations to this conclusion (discussed below), our results indicate that depression diagnoses contain relatively little predictive utility over and above continuous symptom scores of depression, at least in regards to a number of key dependent variables. Similarly, our results suggest that two of the three simultaneous diagnostic criteria, Key Symptoms and Symptom Duration, also have little

incremental predictive utility. We did find that Level of Functioning (based on the GAS) as a main effect has some predictive utility. In fact, in many instances, we found significant unique effects for GAS scores but not for continuous depressive symptom scores in the prediction of the dependent variables. However, in the tests of the interactions among continuous symptom scores and simultaneous diagnostic criteria, which were the tests most directly relevant to the primary purpose of this study, we found few significant effects.

Taken together, these findings suggest that the components of depression diagnoses that historically have been seen as important and defining characteristics may not actually be so. For example, the presence or absence of sadness and / or anhedonia – a key defining feature of depression according to current conceptualizations – seems to add little to our ability to predict and hence to understand other relevant variables, either as main effects or as interactions. That is, the relation between (a) the continuous syndrome measure of the symptoms of depression and (b) a number of dependent variables theoretically as well as empirically linked to depression did not differ as a function of the presence or absence of a key defining feature of depression, sadness and / or anhedonia. What this means is that the syndrome of depressive symptoms with sadness and / or anhedonia, and a syndrome of depressive symptoms without sadness or anhedonia do not differ from each other in their relations to a number of theoretically linked constructs. This in turn raises questions regarding the validity of depression as a unique diagnostic or categorical entity, at least as currently defined.

It is interesting parenthetically to note that, although not directly relevant to the primary focus of the present study, in many instances the level of depressive symptoms did not show significant unique effects when controlling for level of functioning.

There are two levels of implication for this study, the first pertaining to the need for replication, and the second to our findings if replicated. In regards to the first, it is obvious that the present results need to be replicated and extended, across a number of different dimensions. First, because we had access to a rich dataset, we were able to assess the relations between the different components of the depression diagnosis and a relatively wide range of dependent variables (life events, cognitions, social support) theoretically as well as empirically linked to depression. Nonetheless, it will be important for future work to determine if these results replicate with other dependent variables relevant to depression (e.g., cognitions that are specifically linked to depression but not other forms of psychopathology). Second, the present study utilized a treatment sample for two reasons: (a) because clinical trial studies often use diagnoses as selection and outcome criteria, they hence represent an important research domain in which to assess the utility and validity of diagnoses; and (b) because it allowed for use of a true dichotomous dependent variable. However, it is possible that the treatment itself somehow differentially affected continuous syndrome versus diagnostic depression, thus biasing our results vis-à-vis the true utility and validity of diagnoses. However, why and how this would have happened is not obvious, and if it did, one might expect that it would bias results more in the direction of increasing the utility of diagnoses, since the treatment focused per se on depression as a diagnostic entity.

Third, another area where it would be important to extend the present findings would be to compare an interview vs. self-report version of the same measure, in order to determine whether resource intensive interviews provide utility over and above self-report. Fourth it will be important to see if similar findings emerge with respect to other diagnostic categories. It is possible that for other diagnostic categories (e.g., schizophrenia) that assess behavior and affect that differs more qualitatively from normative behavior, diagnoses may be more likely to show utility over and above continuous syndrome measures. And finally, this study involved a sample of depressed adults and therefore the applicability of the results to children and adolescents is unclear.

In regards to the second level of implications of our results, if these findings were replicated, it would raise the question as to whether obtaining structured diagnoses is worth the expenditure of resources that they require, or whether a simple syndrome measure would suffice. More significantly, the results if replicated also raise fundamental questions about diagnoses and their validity as reflecting categorically distinct entities.

Conclusion

Despite the need for replication across this and other domains of psychopathology and cohorts, this study provides valuable information about categorical versus continuous measures of depression and the larger question of the utility of categorical versus continuous measures of psychopathology. Two broad conceptualizations of psychopathology are widely recognized – psychopathology as distinct diagnostic classes and psychopathology as continuous quantitative syndromes. The relative utility of each

of these conceptualizations has been debated for many years and continues to be a hot topic in the contemporary psychopathology and treatment literatures. The present study focused on one common form of psychopathology – Major Depressive Disorder – and the relative utility of the two conceptualizations of this disorder. Our findings suggest that categorical diagnostic depression variables have relatively little unique predictive utility over and above continuous symptom scores of depression. Additionally, our findings suggest that key symptoms and duration criteria, components of diagnoses thought to be important and unique in distinguishing psychopathology from normality, also have relatively little predictive utility, although level of functioning did have substantial predictive utility in most cases. In fact, our findings suggest that functioning may be a better predictor of related dependent variables in some cases than level of depressive symptoms. Future efforts should be made to determine whether our findings replicate within the same diagnostic category and across other diagnostic categories and should study more fully the implications these findings have for treating psychopathology.

REFERENCES

Achenbach, T.M. & Edelbrock, C.S. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, VT: University Associates in Psychiatry.

Agresti, A. (1996). *An introduction to categorical data analysis*. New York: Wiley.

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

Amir, N., Beard, C., & Przeworski, A. (2005). Resolving ambiguity: The effect of experience on interpretation of ambiguous events in generalized social phobia. *Journal of abnormal psychology*, 114(3), 402-408.

Beauchaine, T.P. (2003). Taxometrics and developmental psychopathology. *Development and psychopathology*, 15, 501-527.

Bockting, C.L.H., Schene, A.H., Spinhoven, P., Koeter, M.W.J., Wouters, L.F., Huyser, J., Kamphuis, J.H., & The DELTA Study Group. (2005). Preventing relapse/recurrence in recurrent depression with cognitive therapy: A randomized controlled trial. *Journal of consulting and clinical psychology*, 73(4), 647-657.

Cantwell, D.P. (1996). Classification of child and adolescent psychopathology. *Journal of child psychology and psychiatry*, 37(1), 3-12.

Chambers W.J., Puig-Antich J., Hirsch M., Ambrosini, M.A., & Tabrizi, M.D. (1985). The assessment of affective disorders in children and adolescents by semi-structured interview: Test-retest reliability of the Schedule for Affective Disorders and Schizophrenia for School-Age Children, Present Episode Version. *Archives of general* psychiatry, 42(7), 696-702.

Chard, K.M. (2005). An evaluation of cognitive processing therapy for the treatment of posttraumatic stress disorder related to childhood sexual abuse. *Journal of consulting and clinical psychology*, 73(5), 965-971.

Clark, L.A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual review of psychology*, 46, 121-153.

Cohen, J. (1983). The cost of dichotomization. *Applied psychological measurement*, 7(3), 249-253.

Elkin, I., Parloff, M.B., Hadley, S.W., & Autry, J.H. (1985). NIMH treatment of depression collaborative research program: background and research plan. *Archives of general psychiatry*, 42, 305-316.

Endicott, J. & Spitzer, R.L. (1978). A diagnostic interview: the schedule for affective disorders and schizophrenia. *Archives of general psychiatry*, 35(7), 837-844.

Ferdinand, R.F., Heijmens Visser, J., Hoogerheide, K.N., van der Ende, J., Kasius, M.C., Koot, H.M., & Verhulst, F.C. (2004). Improving estimation of the prognosis of childhood psychopathology; combination of DSM-III-R/DISC diagnoses and CBCL scores. *Journal of child psychology and psychiatry*, 45(3), 599-608.

Fergusson, D.M. & Horwood, L.J. (1995). Predictive validity of categorically and dimensionally scored measures of disruptive childhood behaviors. *Journal of the American academy for child and adolescent psychiatry*, 34(4), 477-485.

Fergusson, D.M., Horwood, L.J., Ridder, E.M., & Beautrais, A.L. (2005). Sub-threshold depression in adolescence and mental health outcomes in adulthood. *Archives of general psychiatry*, 62, 66-72.

First, M.B., Spitzer, R.L., Gibbon, M., and Williams, J.B.W. (2002). *Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition.* (SCID-I/P). New York: Biometrics Research, New York State Psychiatric Institute.

Hinshaw, S.P., Lahey, B.B., & Hart, E.L. (1993). Issues of taxonomy and comorbidity in the development of conduct disorder. *Development and psychopathology*, 5, 31-49.

Jensen, P.S. (1995). Discussion: scales versus categories? Never play against a stacked deck. *Journal of the American academy of child and adolescent psychiatry*, 34(4), 485-487.

Jensen, P.S., Koretz, D., Locke, B.Z., Schneider, S., Radke-Yarrow, M., Richters, J.E., & Rumsey, J.M. (1993). Child and adolescent psychopathology research: problems and prospects for the 1990s. *Journal of abnormal child psychology*, 21(5), 551-580.

Jensen, P.S., Watanabe, H.K., Richters, J.E., Roper, M., Hibbs, E.D., Salzberg, A.D., & Liu, S. (1996). Scales, diagnoses, and child psychopathology: II. Comparing the CBCL and the DISC against external validators.

Kraemer, H.C., Noda, A., & O'Hara, R. (2004). Categorical versus dimensional approaches to diagnosis: methodological challenges. *Journal of psychiatric research*, 38, 17-25.

Krueger, R.F., Watson, D., & Barlow, D.H. (Eds.). (2005). Toward a dimensionally based taxonomy of psychopathology [Special section]. *Journal of abnormal psychology*, 114(4), 491-569.

MacCallum, R.C., Zhang, S., Preacher, K.J., & Rucker, D.D. (2002). On the practice of dichotomization of quantitative variables. *Psychological methods*, 7(1), 19-40.

Maxwell, S.E. & Delaney, H.D. (1993). Bivariate median-splits and spurious statistical significance. *Psychological bulletin*, 113, 181-190.

McDermott, P.A. & Weiss, R.V. (1995). A normative typology of healthy, sub-clinical, and clinical behavior styles among American children and adolescents. *Psychological assessment*, 7(2), 162-170.

Morey, L.C. (1991). *Personality assessment inventory – professional manual*. Florida, USA: Psychological assessment resources, Inc.

Rohde, P., Lewinsohn, P.M., & Klein, D.N. (2005). Association of parental depression with psychiatric course from adolescence to young adulthood among formerly depressed individuals. *Journal of abnormal psychology*, 114(3), 409-420.

Spitzer, R.L. & Endicott, J. (1978) Schedule for Affective Disorders and Schizophrenia—Change Version, 3rd ed. New York, New York State Psychiatric Institute, Biometrics Research.

Trull, T.J., Widiger, T.A., & Guthrie, P. (1990). Categorical versus dimensional status of borderline personality disorder. *Journal of abnormal psychology*, 99(1), 40-48.

Weertman, A., Arntz, A., Schouten, E., & Dreessen, L. (2005). Influences of beliefs and personality disorders on treatment outcome in anxiety patients. *Journal of consulting and clinical psychology*, 73(5), 936-944.

Weissman, A.N. (1979). The Dysfunctional Attitude Scale: A validation study. *Dissertation abstracts international*, 40, 1389B-1390B, (University Microfilms No. 79-19, 533).

Weissman, A.N. & Beck, A.T. (1978, November). Development and validation of the Dysfunctional Attitudes Scale: A preliminary investigation. Paper presented at the meeting of the Association for the Advancement of Behavior Therapy, Chicago.

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(4), 494-504.

World Health Organization. (1992). *International statistical classification of diseases and related health problems* (10th ed.). Geneva: Author.