Testing Gender and Ethnicity Invariance of the Comorbidity Structure of Common Mental Disorders

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Nicholas Robert Eaton, M.A.

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Robert Frank Krueger, Ph.D.

August 2012

Acknowledgements

This project reflects the professional support of numerous mentors and collaborators throughout the years. Primarily, I acknowledge the steadfast guidance provided by my graduate advisor, Robert F. Krueger, Ph.D., my post-baccalaureate advisor, Thomas F. Oltmanns, Ph.D., and my undergraduate advisor, Deanna M. Barch, Ph.D. Second, I acknowledge the collaboration of the following individuals: Steve Balsis, Ph.D., Bridget F. Grant, Ph.D., Ph.D., Deborah S. Hasin, Ph.D., Katherine M. Keyes, Ph.D., Kristian E. Markon, Ph.D., Arjen Noordhof, Ph.D., Andrew E. Skodol, M.D., and Susan C. South, Ph.D.

This project similarly reflects the significant personal support I have received over the years. I acknowledge in particular my devoted partner, Anthony B. Curtis, M.A., and my loving parents, Robert L. and Brenda E. Eaton.

Finally, I acknowledge funding support provided by the National Institutes of Health (U01AA018111, R01DA018652, K05AA014223, F31DA026689), the University of Minnesota, and Washington University in St. Louis.

Dedication

This dissertation is dedicated to all those who experience mental illness; ultimately, it is their well-being that motivates our efforts.

Abstract

Epidemiological studies of categorical mental disorders consistently report gender and ethnicity differences in many disorder prevalence rates. Further, these disorders are often comorbid. Can a dimensional multivariate liability model be developed to clarify how gender and ethnicity are associated with diverse, comorbid mental disorders? I pursued this possibility in the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC; $N = 43{,}093$). Gender and ethnicity differences in prevalence rates showed systematic patterns; for instance, women showed higher rates of all internalizing (mood and anxiety) disorders, and men showed higher rates of all externalizing (antisocial and substance use) disorders. I next investigated the latent associations underpinning disorder comorbidity and found that a dimensional internalizing-externalizing liability model fit the data well in all sub-populations. This model was gender and ethnicity invariant, indicating that observed gender and ethnicity differences in prevalence rates originated from the groups' different average standings on latent internalizing and externalizing liability dimensions. I discuss implications of these findings for understanding gender and ethnicity differences in psychopathology and for classification and intervention.

Table of Contents

List of Tables

List of Figures

Introduction

Mental disorders are most often defined as discrete, dichotomous entities, and the current nosological system, the *DSM-IV-TR* (American Psychiatric Association, 2000), loosely groups them under rationally derived headings such as "mood disorders" and "anxiety disorders" (Watson, 2005). However, high levels of comorbidity between these putatively distinct disorders—even across these broader groupings—highlight the interrelatedness of many manifestations of psychopathology (Krueger & Markon, 2006). For example, major depressive disorder, panic disorder, and generalized anxiety disorder co-occur more frequently than expected by chance alone (Krueger, 1999; Krueger, Caspi, Moffitt, & Silva, 1998; Mineka, Watson, & Clark, 1998). Such finding raise questions about what drives this observed comorbidity. Do some disorders have common etiological roots (e.g., environmental circumstances or genetic predispositions)? Can they be better conceptualized as manifestations of more general underlying liabilities? If so, what factors determine whether a liability will primarily manifest as one type of disorder (e.g., depression) instead of another (e.g., anxiety) at a specific point in an individual's life?

I review research attempts to answer these, and related questions, below. I begin by briefly discussing issues relevant to comorbidity and statistical modeling. The early structural literature regarding two dimensions that underlie common mental disorders—internalizing (INT) and externalizing (EXT)—is reviewed, followed by replications and expansions of these models. I discuss the link between the INT and EXT dimensions and review the literature on the longitudinal stability of INT-EXT. Finally, I discuss what little is known about INT and EXT as they occur in different sub-groups, such as

age cohorts across the lifespan and individuals from different countries around the globe.

Following the review of what is known about INT and EXT, I turn my attention to address a critical outstanding question in the structural psychopathology literature by empirical analysis: Are these latent liability dimensions invariant across gender and ethnicity? In other words, does the INT-EXT model underlie comorbidity patterns similarly in different sub-populations? This is a question of factorial invariance—whether *similar* or statistically the *same* dimensions are present in different groups. If invariance can be established, certain inferences regarding the origins of group differences in prevalence rates can be supported. Finally, I discuss the implications of my results for disorder conceptualization and classification, psychopathology research on etiology and gender/ethnicity, and intervention approaches.

Representations of Comorbidity

There are several ways to think about comorbidity, and it is worthwhile to discuss them briefly, because they form the basis of the conceptualization and associated statistical modeling that will follow. More in-depth discussions of comorbidity are presented elsewhere for the interested reader (e.g., Krueger & Markon, 2006; Lilienfeld, Waldman, & Israel, 1994). In general, for current purposes, these representations of comorbidity differ in terms of the type of variables analyzed and corresponding conceptual models. Three types of variables will be discussed in terms of comorbidity modeling: categorical, continuous, and categorical variables that represent dichotomizations of underlying continua. I then explore the difference between co-

occurrence and correlation of disorders in terms of comorbidity, which is illustrated with a real world example.

Comorbidity of putatively distinct categorical disorders. The first way to conceptualize comorbidity is in terms of co-occurring diagnoses that are putatively distinct. In the traditional medical model, disorders are typically conceived of as distinct categorical entities, with distinct etiopathophysiologies. For instance, a patient either has HIV or does not, and that same patient may also either have diabetes or not (Feinstein, 1970). Following this somatic medical model, mental disorder comorbidity is commonly conceptualized in this way, such that patients can meet diagnostic criteria for two or more categorical disorders at the same point in time.

Comorbidity of continuous disorders. Although the categorical disease model has prevailed for many years in the psychiatric community, more recent research has demonstrated the benefits of moving toward dimensional models of mental disorders (e.g., Helzer et al., 2008; Krueger & Piasecki, 2002; Widiger & Samuel, 2005). Under a dimensional system, comorbidity can be thought of differently than in the categorical disease model, and it is in this dimensional way that comorbidity is conceptualized in much of the research to be discussed henceforth. Instead of calculating the proportion of individuals who either have major depression, generalized anxiety disorder, or both, dimensional models of comorbidity typically utilize covariation between continuous symptom count variables for each disorder. For example, all individuals in a study might be given a structured clinical interview to determine whether or not each diagnostic criterion for major depression and generalized anxiety disorder is present or absent in their lives. The numbers of depression symptoms and of generalized anxiety

symptoms present are totaled separately for each individual, and covariance between the two symptom counts is calculated. The size of this covariance, or alternatively, the standardized covariance (i.e., correlation) between the two disorders' symptoms counts, can be thought of as the degree of symptomatic comorbidity these two disorders show (e.g., Krueger, Chentsova-Dutton, Markon, Goldberg, & Ormel, 2003).

Comorbidity of categorical variables modeled continuously. A third means of understanding and modeling comorbidity, germane to the empirical analyses I conduct below, combines the categorical diagnostic and dimensional approaches discussed above. Dichotomous categorical diagnoses of disorders can be treated statistically as continuous dimensions by utilizing tetrachoric correlations. Tetrachoric correlations are indices of association that assume a liability-threshold model (i.e., at a certain threshold point on the liability continuum of a disorder, the disorder switches from being "absent" to being "present;" see Kendler, 1993). Many studies examining the comorbidity between, and underlying structures of, common mental disorders utilize tetrachoric correlations (e.g., Krueger, 1999; Krueger, Caspi, Moffitt, & Silva, 1998), and therefore an understanding of the use of tetrachoric correlations is beneficial.

On the most basic level, the key concept underlying the analysis of tetrachoric correlations is that manifest dichotomous variables (e.g., a yes-no categorical diagnosis of major depressive disorder) can be modeled in such a way that they reflect latent dimensions. For example, a researcher might assume that an underlying distribution of Disease X symptomatology is continuously distributed. Individuals who do not receive a dichotomous diagnosis of Disease X would fall below a certain diagnostic threshold on this continuum, and individuals who fall at or above the dimension's diagnostic

threshold would receive a dichotomous diagnosis. If Disease X requires the presence of five symptoms to be present for an individual to receive the diagnosis, the location of the threshold on the latent Disease X symptomatology dimension would be at five symptoms.

Comorbidity as co-occurrence versus correlation. Comorbidity can be defined in a variety of ways. Two definitions of comorbidity with different implications for understanding disorder overlap are co-occurrence and correlation (see Krueger & Markon, 2006). The simultaneous presentation of two or more disorders in one individual does not necessarily indicate that the disorders are related. Instead, individuals experiencing one disorder may have the other disorder by chance. This scenario, co-occurrence, simply implies that, due to prevalence rates of each disorder, a particular number of individuals with one disorder will likely experience the other. The second scenario, correlation, is seen when two or more disorders relate more strongly than chance (reflecting their prevalence rates) would dictate. Thus, when comorbidity is thought of in terms of disorder-disorder correlation, mental disorders are present simultaneously due to some association between them. It is this correlational view of comorbidity that will be focused upon in this chapter, because the empirical evidence supports the existence of correlations among disorders; in other words, many common mental disorders do seem to co-occur more frequently than would be expected by chance alone.

The following example illustrates the difference between comorbidity as co-occurrence and correlation, which utilizes the basic dichotomous diagnosis conceptualization of comorbidity discussed above. Under a categorical nosological

system, comorbidity between dichotomous diagnoses of two disorders can be thought of in terms of a two by two contingency table, such those depicted in Table 1. The two tables presented represent the difference between co-occurrence and correlation as bases for comorbidity between major depressive disorder and generalized anxiety disorder. Prevalence rates of the two disorders and observed cases of comorbidity in these tables are based on data from 7,108 individuals in the national probability sample of the Midlife Development in the United States (MIDUS) study (see Brim, Ryff, & Kessler, 2004). These prevalence rates have been rescaled to a sample of 1,000 individuals for simplicity of illustration and rounded to the nearest whole number. They are presented as the marginal values in the table.

The upper table represents the overlap of major depression and generalized anxiety disorder if the disorders showed no interrelation, and thus represents comorbidity conceptualized as co-occurrence (and not correlation). The number of individuals in each cell is an expected value, however, based on the prevalence rates of major depression and generalized anxiety disorder in the sample. For example, we see, based on the marginal values, that 13.3% (i.e., [133/1000] * 100 = 13.3%) of individuals in the MIDUS sample experienced major depressive disorder, and 2.7% of individuals experienced generalized anxiety disorder. The expected frequencies of individuals in each cell were calculated using these prevalence rates. We would expect to see approximately four individuals out of 1,000 who experienced comorbid major depression and generalized anxiety based only on these prevalence rates (i.e., .133 * .027. * 1000 = 3.591, which rounds up to 4 individuals). Remember, this expected value

is calculated assuming that major depression and generalized anxiety disorder have no association and simply co-occur due to chance alone.

The lower table represents disorders that are comorbid not only because they co-occur but also because they correlate; they are associated with one another at greater than chance levels. The frequencies in the lower table, unlike those in the upper table, are observed values, and thus each cell represents the actual frequencies of individuals seen in the MIDUS study. The marginal prevalence rates remain the same in the lower table, but the cell values differ from those in the upper table. Of most importance for current purposes are the cells in bold: the number of individuals observed who experienced both major depressive disorder and generalized anxiety disorder. As mentioned above, we would expect four individuals to experience comorbidity of these disorders if they were unrelated. However, we see that 17 individuals experienced both disorders. This is more than 400% of the number of individuals with comorbidity we would expect if the disorders were, in fact, not associated. The marked elevation of observed values over expected values suggests that the disorders are correlated to some extent and thus are seen in tandem more frequently than chance levels would dictate, likely due to a relation between them (e.g., they may both be manifestations of the same latent construct, one disorder may "cause" the disorder, and so on; see Klein & Riso, 1993; Neale & Kendler, 1995).

Once we have established, as we have in the example above, that disorders are comorbid because of a correlation (and not simply a co-occurrence level due to prevalence rates), we may begin to ask why it is that this correlation exists. Numerous factors could account for this sort of comorbidity. It could be the case that one disorder

commonly causes another, and this etiological pattern could result in comorbidity. HIV infection is commonly seen in tandem with AIDS-related medical complications, for instance, because the HIV infection leads to suppressed immunity, which allows for the proliferation of the medical complications.

In terms of psychopathology, there is a relative lack of compelling data for most major mental disorders that indicates a clean, causal etiological pathway from one disorder another (Krueger & Markon, 2006). Thus, another hypothesis is needed to account for mental disorder comorbidity. A potentially compelling explanation for correlations between disorders is that the disorders are linked by a common latent construct. This hypothesis has begun to take hold as more and more disorders have been shown to interrelate to one another, and thus the presence of a psychologically meaningful underlying construct, or constructs, can be posited. I explore this line of thinking in the following section.

Comorbidity and Common Factors

As discussed above, many mental disorders show observed comorbidity levels that are higher than one would expect due to chance alone. A hypothesis to account for the observed comorbidities between many forms of psychopathology is that seemingly distinct mental disorders may be manifestations of common underlying constructs. That is, one or more unobserved latent liability dimensions/factors would account for the observed covariation between disorders. This is an application of the common factor model (Thurstone, 1947), which states that related observed variables are linear functions of one or more common factors and one unique factor per observed variable (these unique factors being typically understood as variable-specific variance, which

includes psychometric error). This common factor hypothesis, as applied to mental disorder comorbidity, can also be tested statistically, and, indeed, results of such analyses have generally been supportive.

The common factor model is depicted in Figure 1. In this hypothetical example, there are three manifest (i.e., observed) variables, depicted as rectangles by convention, and one latent common dimensional factor that links the three, depicted as ovular. A unique factor loading (denoted λ) links each manifest variable to the latent factor. Each observed variable also has a unique factor (denoted ε) that accounts for its specific variance, which is the variance in the observed variable not accounted for by the latent factor. Each individual has a score on the latent factor (commonly represented as η; not included in the figure). The observed scores for individuals on any given manifest variable is a linear composite the factor score (weighted by the factor loading) and the unique factor for that variable. For example, the level of major depression observed (which we might denote y_{MD}), is $y_{MD} = \lambda_{MD}\, \eta + \varepsilon_{MD}$. The common factor model and its related statistical techniques of factor analysis use observed variables, such as a symptom count variable for generalized anxiety disorder, to understand better the way the variables relate to one another and to ascertain the presence of any latent factors that might account for observed variable interrelations.

If a researcher had collected data on major depression, generalized anxiety, and panic disorder and noticed strong interrelations between the three disorders, he or she might be curious as to what was driving this observed covariation. The model shown in Figure 1 depicts an answer to this question. If the model shown in this figure represented the true state of nature, it would imply that all individuals have a standing

on a latent dimensional factor. This latent factor is related to each of the researcher's three variables in different ways (as represented by the three different factor loadings); it is these three variables' relations to the common factor that account for their observed comorbidity with one another. The researcher could then posit different models (e.g., a model with two latent factors) and, by comparing model fit indices, could determine which model best accounted for the observed covariances between disorders in a parsimonious way. This process, known as confirmatory factor analysis (CFA; as opposed to the more common exploratory factor analysis, EFA), is utilized in the majority of the research to be discussed subsequently, and the interested reader is referred to Brown (2006) for a solid introduction to the theory and application of these techniques.

The Structure of Common Mental Disorders: The INT-EXT Model

Researchers have long considered the structure of mental health problems, especially in light of the levels of comorbidity between certain disorders. Studies of the structure of psychopathology have an especially strong history in the area of child mental health research. Indeed, this child-oriented research posited the notion that two dimensional factors could account well for the comorbidity between many common psychopathological syndromes. Achenbach and Edelbrock (1978; 1984) reviewed this early thought about the structure of childhood psychopathology in detail.

Building upon the foundation of the child psychopathology literature, confirmatory factor analyses of the type described above have been increasingly applied to questions of psychiatric disorder comorbidity in adults in the past two decades. The results of these studies in adults, by and large, have also suggested that a model with

two broad dimensional factors, INT and EXT, accounts best for the observed covariances between many common major mental disorders (for adult research literature reviews, see Eaton, South, & Krueger, 2010, and Krueger & Markon, 2006; for child research literature reviews, see Achenbach & Edelbrock, 1978, 1984). INT is composed of disorders such as major depressive disorder, generalized anxiety disorder, agoraphobia, panic disorder, social phobia, specific phobia, dysthymic disorder, and obsessive-compulsive disorder. EXT is composed of disorders such as antisocial personality disorder, conduct disorder, and abuse/dependence of alcohol, nicotine, marijuana, and other drugs. It is worth noting that the studies discussed below focused on epidemiological data, which most often only assess relatively common mental disorders. As such, only common mental disorders are typically included in structural psychopathology studies. Disorders with low base-rates (e.g., psychotic disorders) are not easily amenable to such analyses, although the comparatively few studies that have included these disorders are discussed below.

The General Structure of INT-EXT

The *structures* of INT and EXT are the best understood aspects of these factors that characterize comorbidity between mental disorders. Krueger and colleagues (1998), drawing on previous research (e.g., Achenbach & Edelbrock, 1978), examined the relations between 10 common *DSM-III-R* psychiatric disorders (i.e., major depressive episodes, dysthymia, generalized anxiety disorder, agoraphobia, social phobia, simple phobia, obsessive-compulsive disorder, conduct disorder, and marijuana and alcohol dependence). Categorical diagnostic data were available for individuals followed longitudinally at age 18 (n = 930; we will refer to this as Time 1) and age 21 (n = 937;

Time 2). Due to the categorical nature of the data, the researchers adopted a liability-threshold model, and thus analyzed tetrachoric correlations between the disorder diagnoses. Because this study has been cited by subsequent research in this area, was a relatively early contribution to the adult literature on this topic, and serves as a model for the reader of most studies to be reviewed subsequently, it will be discussed at some length.

Krueger and colleagues (1998) used CFA to fit three different models to the data at Times 1 and 2. Figure 2 illustrates several types of the common structural models tested in this and subsequent studies. Because studies frequently differ in the particular disorders (and number of disorders) included in the models that are tested, eight of the disorders examined by those authors were selected to illustrate the models in Figure 2. Simple phobia and obsessive-compulsive disorder are not included for the purpose of simplifying these illustrated models. The first model fit in that study was defined as all 10 of the disorders included in their study loading on a single common factor (i.e., a "general psychopathology" factor). Model fit indices indicated that this model fit the data reasonably well but left noticeable room for improvement. The second model fit was a two-factor model in which major depressive episodes, dysthymia, generalized anxiety disorder, agoraphobia, social phobia, simple phobia, and obsessive-compulsive disorder were indicators of one latent factor (i.e., INT) and conduct disorder, marijuana dependence, and alcohol dependence were indicators of a second latent factor (i.e., EXT). This model fit the data very well according to fit indices, and it also provided a markedly better fit than did the one-factor model; Figure 2 includes a simplified

representation ("Two-factor INT-EXT model") to illustrate this model's two-factor structure.

The Time 1 data were fit to one final model by Krueger and colleagues (1998): a four-factor model (see Figure 2's simplified "Four-factor model"). As discussed above, the *DSM-IV* places disorders into rationally derived subgroups, and this model represented this approach. In this model, major depressive episodes and dysthymia made up one latent factor (i.e., affective disorders); generalized anxiety disorder, obsessive-compulsive disorder, social phobia, agoraphobia, and simple phobia made up another latent factor (i.e., anxiety disorders); alcohol and marijuana dependence made up a third latent factor (i.e., substance dependence); and conduct disorder served as an indicator for the fourth latent factor (i.e., antisocial behavior). Although this model fit the data well, it was clearly overparameterized, due to its large, unfavorable change in some fit statistics compared to the two-factor model. These results taken together highlighted the good fit of a two-factor INT-EXT model to account parsimoniously for the observed comorbidity between the disorders included.

The Time 2 data of individuals at age 21 years showed a similar pattern. The same two-factor model fit the data well. The one-factor model showed a worsening of fit compared to the two-factor model. The four-factor model again yielded a good model fit, but fit indices indicated it was overparameterized and thus increased fit at the expense of parsimony. Thus, the two-factor model was again preferred in this second wave of data. Of note, at both Time 1 and Time 2, the latent variables in the four-factor model tended to mimic the best-fitting, two-factor model. At Time 1, the correlation between the anxiety and affective factors was estimated at 1.0, and the correlation

between the antisocial behavior and substance dependence factors was estimated at .89. These correlations were .90 and .72 at Time 2, respectively. Thus, even when INT and EXT were split into two separate factors each (i.e., when INT was split into anxiety and affective factors, and EXT was split into antisocial behavior and substance dependence factors), those factors tended to correlate very highly together; this indicated the presence of two higher-order factors (i.e., INT and EXT) to account for these factor-factor correlations. This finding can be taken as further support for the hypothesis that the INT and EXT factors account for the observed comorbidity between numerous major mental disorders.

A study published the following year (Krueger, 1999) utilized diagnostic data from 8,098 individuals from the National Comorbidity Survey (NCS). Unlike the previous study, in which participants were in late adolescence and early adulthood, these individuals ranged in age from 15 to 45 years. Ten disorders were again modeled, and they were for the most part the same as those from the Krueger and colleagues (1998) study; however, panic disorder replaced obsessive-compulsive disorder, antisocial personality disorder replaced conduct disorder, and drug dependence replaced marijuana dependence. Four models were fit to these disorders, three of which were the same as those in the study above (i.e., a one-factor, a two-factor, and a four-factor model, which divided both the INT and EXT factors in two). Exploratory factor analyses of the data, however, had revealed the presence of two sub-factors for the INT factor: 1) an “anxious-misery” factor (referred to as “distress” in more recent literature), with indicators of major depressive episodes, dysthymia, and generalized anxiety disorder, and 2) a “fear” factor with social phobia, simple phobia, agoraphobia, and

panic disorder (see Figure 2 for a simplified representation: "Hierarchical INT-EXT model").

The results of model fitting analyses indicated that the three-factor model best balanced fit and parsimony in the total sample by a wide margin. The sample was then divided randomly in half, and the three-factor model fit best in both halves. The author of the study took these results (and other sub-sample results to be discussed below) collectively to indicate the superiority of the three-factor model in general, especially due to the fact that this model provided the best fit in five groups, including the large total sample. Overall, this study provided strong evidence for the three-factor model, wherein EXT is unitary and a higher-order INT is bifurcated into distress and fear sub-factors. We will refer to this model as the bifurcated or hierarchical INT-EXT model to remain consistent with much of the literature; technically, however, this is a higher-order, rather than hierarchical, model.

Replications of the INT-EXT Structure

The two studies above converged on a two-factor model of psychopathology: The observed covariances between disorders could be captured well, but also parsimoniously, by means of the INT and EXT latent factors. Although these studies had utilized different data drawn from different populations, further replication of this finding was warranted. Additional studies on this topic were also necessary to clarify whether INT was best conceptualized as a unitary factor or a bifurcated structure, with a higher-order factor subsuming distress and fear sub-factors.

Several research groups have replicated and extended the INT-EXT model in the ensuing years. These replications further support the hypothesis that the INT and EXT

factors link many common mental disorders and account for their comorbidity. In 2001, Vollebergh and colleagues modeled the structure of nine mental disorders in a community sample from the Netherlands. Diagnoses of *DSM-III-R* disorders (occurring the previous 12-months, and thus not lifetime diagnoses) were analyzed via tetrachoric correlations and fit to four models. The first was a one-factor, "general psychopathology" model, in which comorbidity between all nine disorders was accounted for by a single latent dimension. The second model reflected a two-factor structure of INT and EXT. A third model tested the bifurcated INT-EXT structure outlined by Krueger (1999) by including distress and fear sub-factors of INT. The fourth model placed conceptually similar disorders together. The authors of the study tested the fit of this fourth model by placing disorders into three *DSM*-based groups: mood disorders, anxiety disorders, and substance use disorders. Due to the longitudinal design of their assessment (discussed below), these models were fit to several subsets of the data.

The results of this study served as strong replication for previous INT-EXT research and also supported the bifurcated INT-EXT model. The one factor model tended to fit worst across all the analyses. Also providing a relatively poor fit was the three-factor, *DSM*-based model (i.e., mood disorder, anxiety disorder, and substance use factors). The two-factor INT-EXT model fit better. This INT-EXT model had a significantly better fit than the one-factor model, and did not show a worse fit than the more parameterized *DSM*-based three-factor model. However, the superior model was the bifurcated INT-EXT model. Across analyses, this model had the most favorable results on several fit statistics.

Other studies have also replicated the INT-EXT structure. Slade and Watson (2006) utilized 12-month diagnoses of *DSM-IV* and *ICD-10* (World Health Organization, 1992) collected in an Australian community sample. The four models fit were the same as those utilized by Vollebergh and colleagues (2001) discussed above, and the results were largely the same. For *DSM-IV* and *ICD-10* disorders, the one-factor model provided the worst fit according to a variety of fit indices. The *DSM*-based three-factor model fit somewhat better, but generally not as well as the two-factor INT-EXT model. Again, utilizing both diagnostic systems, the bifurcated INT-EXT model was superior with favorable fit index statistics. Not all studies have supported the bifurcated INT-EXT model over the simpler two-factor model, however. Seeley and colleagues (2011) attempted to clarify the optimal model for INT disorders: the two-factor model, the bifurcated INT-EXT model, or a *DSM-IV* organization-based model. The study yielded equivocal results, with all models fitting well and having similar predictive power for lifetime comorbidity.

The individual results of these studies were compelling, and a compilation at that time was warranted to unify the findings. As such, Krueger and Markon (2006) undertook a meta-analysis of these findings. The authors utilized tetrachoric correlation matrices reported in five major studies of psychopathological comorbidity (Kendler, Prescott, Myers, & Neale, 2003; Kessler, Chiu, Demler, & Walters, 2005; Krueger, 1999; Krueger et al., 1998; Vollebergh et al., 2001), which represented a total of 23,557 participants. Several models were fit to these data, and these analyses produced results that were largely congruent with those of previous studies. The one-factor "general psychopathology" model had a poor fit, a two-factor INT-EXT model fit better, and the

best fit overall was provided by the hierarchical INT-EXT model. This study, with its very large sample size and aggregation of different datasets, provided the strongest evidence yet for the latent INT-EXT structure. In addition, these results supported the bifurcation of the INT factor into two subfactors: distress and fear. After this meta-analysis was conducted, a subsequent manuscript provided a narrative review of the INT-EXT literature (Eaton, South, & Krueger, 2010).

While the majority of published structural psychopathology research has supported the existence of the INT-EXT structure, one research group to my knowledge has failed to replicate the findings. In one study, Wittchen and colleagues (2009) attempted to replicate the bifurcated INT-EXT model identified by Krueger (1999) in several age cohorts. Those authors were unable to replicate this, or any, structure across their age cohorts, concluding, "psychopathology cannot be reduced to any simple structure" (p. 189). In the same year, using the same sample, the research group published a second study attempting to replicate Krueger's (1999) model (Beesdo-Baum et al., 2009). In this second study, the authors reported finding evidence of a three-factor (distress-fear-EXT) model, but they concluded there was no evidence of a higher-order INT factor for the distress and fear factors. In addition, it is noteworthy that one study failed to demonstrate that neither the unitary INT model or the bifurcated INT model was superior to a model based on the *DSM-IV* organization in terms of model fit or predictive ability for lifetime comorbidity (Seeley et al., 2011).

Overall, the general structural research up to this point has produced relatively clear general results. The overall latent structure of INT-EXT has been replicated by a number of independent research teams. When compared with other models, such as a

one-factor "general psychopathology" model and *DSM*-based models, INT-EXT has been found to account best for the observed covariances between many common major mental disorders.

The Expansion of the INT-EXT Structure

The results of structural modeling studies depend to a large degree on the disorders included for analysis. The earlier work on the INT-EXT dimensions (e.g., Krueger et al., 1998; Krueger, 1999) found that a particular set of disorders could be accounted for by INT and EXT. The symptoms/disorders that related to INT were: major depressive episodes, dysthymia, generalized anxiety disorder, social phobia, simple phobia, agoraphobia, panic disorder, and obsessive-compulsive disorder. The disorders that related to EXT were: alcohol dependence, drug (e.g., marijuana) dependence, antisocial personality disorder, and conduct disorder.

Subsequent research has expanded the disorders (and updated the disorders' definitions to be congruent with *DSM-IV* and *ICD-10*) and syndromes associated with the INT factor. Krueger and colleagues (2003) parsed anxiety into "anxious worry" and "anxious arousal," both of which are associated with INT. They also found that neurasthenia, somatization, and hypochondriasis loaded on the INT factor. Slade and Watson (2006) demonstrated that posttraumatic stress disorder related to the INT distress sub-factor, and they replicated the inclusion of neurasthenia (also on the distress sub-factor). Recent work has also indicated that bulimia/binge-eating disorder loads on the INT factor (Kramer, Krueger, & Hicks, 2008). Bulimia/binge-eating disorder may not be what some readers would intuitively identify as an INT disorder. This study demonstrates that bulimia/binge-eating disorder has an INT factor loading similar in

magnitude to panic disorder, which is around half of the loading of more traditionally-conceived INT disorders such as generalized anxiety disorder and social phobia. However, the loading of bulimia/binge-eating disorder is notably larger than that of hypochondriasis and more than three times as large as the loading for obsessive-compulsive disorder in this study. Finally, in children and adolescents, it appears that separation anxiety disorder is related to INT (Lahey et al., 2008).

The disorders included in EXT have remained largely similar for the past decade, and fewer have been identified relative to the disorders of INT. This is due in part to a greater focus of diagnostic systems on more INT-related disorders, as well as inclusion of fewer EXT disorders in the datasets utilized for the statistical modeling. Adult antisocial behavior has been linked to EXT (Kramer, Krueger, & Hicks, 2008). In children and adolescents, inattention, hyperactivity-impulsivity, and oppositional defiant disorder seem related to EXT (Lahey et al., 2008). A study of categorical versus continuous liability models of EXT disorders elaborated on the previously identified EXT substance-related disorders. This study found an EXT-based interrelation between nicotine, alcohol, marijuana, cocaine, and other substance dependence disorders (Markon & Krueger, 2005). Finally, some EXT-specific behaviors have been linked to EXT, such as relational, destructive, and physical aggression; boredom proneness; low empathy; alcohol, marijuana, and drug use and problems; blame externalization; feelings of alienation; problematic impulsivity; low planful control; impatient urgency; theft; fraud; low honesty, irresponsibility; low dependability; rebelliousness; and excitement seeking (Krueger et al., 2007). These studies taken together indicate that

EXT, like INT, is a broad factor that underlies many temperamental and psychopathological constructs.

Recent structural psychopathology research has produced some relatively novel and informative findings with regard to the INT-EXT structure. Among these findings is the notion that not all disorders are manifestations of only one underlying factor; rather, a single disorder may be a manifestation of multiple latent liabilities. For instance, borderline personality disorder has recently been demonstrated empirically to cross-load on both the INT sub-factor of distress and EXT (Eaton et al., 2011; James & Taylor, 2008). Such results highlight the potential of single observed disorders to act as manifest confluences of more than one latent construct. Second, a study by Kendler and colleagues (2011; see also Røysamb et al., 2011) expanded the structural psychopathology literature by including all personality disorders (PDs) in addition to common Axis I disorders. All of the disorders modeled showed an INT-EXT latent structure, although that study—which investigated the genetic and environmental underpinnings of disorders rather than their phenotypic correlations—provided evidence that there may be separable INT and EXT factors for Axes I and II.

The expansion of the models of latent psychopathology structure does not solely mean fitting additional disorders into the INT-EXT structure. This model reflects the comorbidity patterns among many common forms of mental disorder, but it is not the case that this model is intended to capture the latent structure of all forms of psychopathology. Historically, particular disorders—most notably psychotic disorders—have been excluded from the model because (1) they are frequently not assessed (or are used as rule-outs for participation) in the large epidemiological samples

analyzed, and (2) they have comparatively low prevalence rates, which can complicate the statistical modeling. While there was at least one indication that some disorders might form a psychosis-related factor (Wolf et al., 1988), only recently have researchers begun to model such disorders regularly. These studies have demonstrated that factors other than INT and EXT—such as a factor commonly referred to as "thought disorder"—are necessary to capture the latent structure of the increasingly diverse sets of mental disorders being modeled (see Kotov et al., 2011a, 2011b; Markon, 2010). INT and EXT are two replicable latent comorbidity factors, but it is certain that other factors will be necessary to fully characterize, and create a comprehensive model of, the structure of all forms of psychopathology.

The Association between INT and EXT

Up to this point, we have only discussed the covariances between manifest variables (e.g., between measured major depression symptom counts and generalized anxiety disorder symptom counts) and have neglected possible covariances between INT and EXT themselves. It is not necessary that these factors are orthogonal (uncorrelated); they can be modeled obliquely (allowed to correlate). As mentioned above, the factors in the four-factor model by Krueger and colleagues (1998) correlated highly. What about the correlation between INT and EXT?

Many studies that have replicated the INT-EXT model have reported an association between INT and EXT. Krueger and colleagues (1998) estimated correlations between INT and EXT of .454 and .417 when their participants were 18 and 21 years old, respectively. Other studies have indicated similar degrees of relationship between INT and EXT. For example, Krueger (1999) found an INT-EXT

correlation of .51 in a different sample, which closely mirrored the correlation of .50 found via a large meta-analysis (Krueger & Markon, 2006). Correlations between INT and EXT estimated from two assessment waves in the Netherlands taken two years apart (.56 and .66) are consistent with the previously reported estimates as well (Vollebergh et al., 2001). Slade and Watson (2006) estimated an INT-EXT correlation of .65 when the factors were assessed via *DSM-IV* diagnostic criteria. The correlation between INT and EXT was almost identical (.61) when the *ICD-10* criteria were utilized to define the indicator disorders instead. Thus, even across classification systems, the INT-EXT correlation holds at relatively the same level of association.

To summarize, across studies, samples, diagnostic systems, and nations, correlations between INT and EXT have converged on a relatively small range of values. This moderate correlation has been reported to range from around .42 to around .66, with .5 seeming to be a reasonable compromise. This result indicates that approximately 25% (i.e., $.5^2 = .25$) of the variance of INT is accounted for by variance in EXT, and vice versa. It should be noted, however, that the common statistical parlance of "accounted for" should not be interpreted to mean "accounted for *causally*," but instead is a comment on the level of covariation between INT and EXT. This shared variance does not necessarily demonstrate that one factor causes the other.

The moderate INT-EXT relationship has several implications for research, theory, and practice regarding the INT and EXT latent factors of psychopathology. Most clearly, this correlation indicates that individuals with mental disorders do not always (nor necessarily even usually) fall into an INT or an EXT group. Indeed, the INT-EXT association suggests that individuals who have INT psychopathology are

likely to have some EXT psychopathology; alternatively, individuals who are externalizers are likely to experience some INT psychopathology. It is again important to note that this is not necessarily a causal relationship. For clarification of what this correlation might mean, consider the following example. Externalizing individuals might commit crimes and frequently use alcohol and drugs. This lifestyle, with its associated impulsivity, health problems, and mistreatment of others, could easily leave these individuals feeling alienated. These feelings of alienation and abandonment by family and friends could lead to feelings of depression. In addition, these individuals might have some anxiety arising out of fear of capture and incarceration. Thus, in this case, the INT symptoms (mood and anxiety) "arose from" EXT behaviors. Alternatively, individuals could experience significant INT psychopathology and self-medicate by using alcohol, marijuana, and harder drugs, thus later experiencing more EXT psychopathology. There are various possible interpretations, and future research may help clarify the etiological nature of the INT-EXT association.

Personality, temperament, and the INT-EXT link. One compelling hypothesis to account for the INT-EXT relationship incorporates a role for individuals' personality and temperament. If particular traits or predispositions are associated with both INT and EXT, these personality and temperamental constructs could serve as the bridge that links the two factors. The notion that personality/temperament and psychopathology are related has a long history in clinical thought and nosology. Previous editions of the *DSM*, as well as the most recent edition, have included PDs in one form or another. Although PDs are often considered to be distinct from other forms of psychopathology—hence their placement on a separate axis from most other forms of

psychopathology in *DSM-IV*—there is little compelling evidence to support this division. Indeed, reviews of the literature indicate that PDs are typically more similar to other mental disorders than they are different (Krueger, 2005). Normal (non-pathological) personality traits have also been linked to mental disorders (e.g., Trull & Durrett, 2005). Thus, the hypothesis that personality and temperament may show associations with the INT and EXT factors of psychopathology and, in fact, may link those factors, is not without a priori empirical support.

Several studies have tested the role of personality and temperament in INT and EXT disorders. This research has noted that both INT and EXT are associated with negative emotionality (or neuroticism), and EXT is further linked with disinhibition (e.g., low conscientiousness, low control, high novelty seeking, high disagreeableness; Clark, 2005; Krueger & Markon, 2006). These findings suggest that negative emotionality may account for a large part of the relation between INT and EXT and predispose one to INT and EXT general psychopathology. The presence or absence of disinhibition, on the other hand, may play a role in determining whether this underlying propensity for psychopathology is manifested as EXT (i.e., disinhibition present) or INT (i.e., disinhibition absent). Additionally, several personality traits account for some lower-level disorder-disorder comorbidity as well. In one study, neuroticism appears to account, to a strong degree, for the interrelations between the INT disorders; neuroticism and novelty seeking trait levels accounted for a good deal of the observed comorbidity between EXT disorders (Khan et al., 2005). Subsequent research has demonstrated that INT is very strongly correlated (r = .98) with measures of

neuroticism, and EXT shows a non-trivial association (r = .29) as well (Griffith et al., 2009).

The Genetic and Environmental Bases of INT-EXT

The previously studies were observational in nature, and they did not evaluate the risk factors that account for the observed comorbidity (Kendler, Prescott, Myers, & Neale, 2003). A determination of the origin of these disorders' comorbidity, and the etiological roots of individuals' standings on the latent liability factors factors could have implications for both conceptualization and treatment. Etiological questions of this sort can be investigated empirically by utilizing appropriate designs and samples with particularly informative characteristics (e.g., twins).

Behavior genetic methods can be applied to such etiological questions to determine, among other things, the role that genes and environments play in the development of comorbidity and liability. For readers unfamiliar with behavior genetic methodology, a brief description may prove helpful. Behavior genetic models, typically conducted with samples of identical (monozygotic) and fraternal (dizygotic) twins, parse variance in the observed variables into additive genetic, shared environmental, and non-shared environmental factors. Shared environment encompasses non-genetic factors that are shared by twins as they grow up, such as familial socioeconomic status, which serve to make the two twins more similar to one another. Non-shared environment is composed of non-genetic factors that differ between twins, such as one twin playing baseball while the other participates in a school orchestra, which serve to make the two twins less similar to one another. Error is also included in non-shared environment. Finally, it is important to note that the variance accounted for by genes,

shared environment, and non-shared environment is estimable in these studies because monozygotic and dizygotic twin pairs differ in the proportion of genes shared between the twins (100% and 50%, respectively); monozygotic and dizygotic twins who were reared together do not differ in the amount of shared and non-shared environment between twins in a twin pair (all twins who were reared together, regardless of zygosity, share 100% of the shared environment and 0% of the non-shared environment, by definition).

Kendler, Prescott, Myers, and Neale (2003) investigated the role of genetic, shared environmental, and non-shared environmental factors in the risk for developing many common disorders in a large sample (N = 5,600) of same-sex twin pairs. The authors modeled two different combinations of ten disorders (major depression, generalized anxiety disorder, phobia, animal phobia, situational phobia, panic disorder, alcohol dependence, drug abuse/dependence, adult antisocial behavior, and conduct disorder) to address the origins of disorders within both INT and EXT as well as between the distress and fear subfactors of INT. Their analyses utilized an independent pathway model, a full description of which is beyond the scope of this chapter, and interested readers are referred to the original manuscript. On a basic level, however, this model allowed for the estimation of the effects of genes, shared environment, and non-shared environment both: 1) as a higher-order, common factor that conferred risk on all related disorders, and 2) as unique effects that conferred risk to each disorder separately. This can be illustrated in Figure 1, if the common factor were, for instance, "genetic effect," and each of the unique variances of the observed variables were disorder-specific genetic effects.

The results of this study indicated that there is a strong genetic effect common to INT disorders and EXT disorders. In addition, genes also showed notable specific effects on alcohol dependence and drug abuse/dependence. Shared environment had effects for adult antisocial behavior and conduct disorder. Non-shared environment tended to show strong unique effects for each disorder. A subsequent study of many Axis I and Axis II disorders by Kendler and colleagues (2011) also found an underlying genetic structure reflecting INT and EXT.

These findings indicate that there is a high degree of genetic risk associated with the origins of the INT and EXT factors. However, there is less of a genetic impact on the development of individual disorders. Genes are thus conferring a common liability for comorbidity among the INT disorders as well as the EXT disorders. Thus, the pattern of lifetime disorder comorbidity that is commonly observed in many major mental disorders occurs primarily through genetic risk factors, especially at the level of the INT and EXT factors.

These studies also suggest that environmental influences play a role as well. There has been comparatively little work conducted to this point toward identifying particular environmental influences on latent factor levels, however. One recent study by Keyes and colleagues (2011) investigated the associations between retrospectively reported childhood abuse and neglect and adult psychopathology. That study found that childhood maltreatment was associated with subsequent psychopathology, and these relations were fully mediated through the latent INT and EXT factors. In other words, the impact of childhood maltreatment appeared to raise undifferentiated liability levels for INT and EXT; it was not associated with risk for any particular disorder. This

finding is congruent with research currently underway that indicates that discrimination and harassment experienced by sexual minority individuals serves to increase the level of INT and/or EXT individuals experience, thereby raising their risk for psychopathological manifestations via a pathway through the latent factors (Eaton, 2012).

The Stability of INT-EXT

Another necessary question to address in the structural psychopathology literature is how stable INT and EXT levels are over time. The determination of the stability of the INT-EXT factors could hold promise for our understanding of etiology, course, and remission of mental disorders. For example, if it were the case that an underlying liability to high levels of INT were relatively constant over time, but the particular manifestations of this INT liability (e.g., major depression, generalized anxiety disorder) changed over time, researchers and clinicians could develop a better conceptualization of the emergence of manifest mental disorders.

A few studies have examined the temporal stability of INT. The first study investigated the stability of INT and EXT from age 18 to age 21 (Krueger et al., 1998). The correlation between INT factor scores at ages 18 and 21 was around .69, and the correlation between EXT factor scores over the same timeframe was around .86. This study indicates that, over three years, both INT and EXT remain relatively stable. An individual's level of INT at age 18 accounted for approximately 48% ($.69^2$) of the variance in her level of INT at age 21. EXT showed an even higher level of stability: Age 18 EXT accounted for approximately three-quarters ($.86^2 = .74$) of the variance in EXT age 21. The second study to investigate INT-EXT stability over time followed a

large, nationally representative Dutch sample for one year (Vollebergh et al., 2001). This study, which replicated previous findings that INT has distress and fear sub-factors, examined the stability of the INT sub-factors individually. The stability of these sub-factors was high: distress over a year was stable at a correlation of .85, and fear showed a stability correlation of .89. Consistent with previous research (Krueger et al., 1999), EXT was significantly more stable than either of the INT subfactors, and it had a stability correlation of .96. Thus, over one year, an individual's level of latent EXT remained almost perfectly stable. These results for INT-EXT stability are somewhat higher than those reported by Krueger and colleagues (1998), which is to be expected. The Netherlands results investigated stability for one year, which was high; the Krueger and colleagues (1998) study investigated stability over three years, which resulted in slightly less stability over a longer period of time. A third study investigated the stability of INT (but not EXT) over the longest period yet investigated (Eaton, Krueger, & Oltmanns, 2011). In that study, INT at baseline, and INT assessed nine to 10 years later, were correlated at .74, a similar level to that reported previously. This suggests that the INT and EXT liabilities are quite stable over even relatively long periods of time.

The Presence of INT-EXT in Culturally Diverse Samples

As discussed above, the INT-EXT structure has emerged in a variety of different (sub-)samples. It has been identified in women and in men; it has been identified across the lifespan; it has been identified in samples from various Western countries. However, is INT-EXT a phenomenon of only Western cultures? If it were the case that other models more superiorly accounted for the comorbidity of mental disorders in non-

Western cultures, for instance, this could draw into question both the validity and the overall utility of the INT-EXT model.

Several studies have investigated the structure of common mental disorders in international samples, which allows us to address the concern of INT-EXT universality. These studies indicate that the INT-EXT structure does not appear to be a solely a Western phenomenon, and it has been found in psychiatric data from a variety of cultures. Below, I review those studies briefly below to illustrate the breadth of the cross-cultural data, which have converged on the notion that the INT-EXT model of comorbidity shows configural invariance.

As discussed above, while Krueger (1999) utilized a U. S. sample, which resulted in a hierarchical INT-EXT model, Krueger and colleagues (1998) analyzed data from individuals in New Zealand. A similar hierarchical structural result was obtained with an Australian sample (Slade & Watson, 2006). In addition, a study conducted in the Netherlands also found that the hierarchical INT-EXT model provided the best fit to the data (Vollebergh et al., 2001).

To address the latent factors of psychopathology to an even broader international degree, one study examined World Health Organization mental health data from 14 countries (Krueger et al., 2003). The countries represented a broad array of cultures and geographic locales: Brazil, Chile, China, France, Germany, Greece, India, Italy, Japan, the Netherlands, Nigeria, Turkey, the United Kingdom, and the United States. Four models, including a one-factor, a two-factor, and two three-factor models, were fit separately to symptom count data from individuals within each country, thus yielding 14 best fitting models. Even across these diverse cultures, the two-factor INT-

EXT model tended to fit the data best; the two-factor model was superior in 12 of the 14 countries. In two countries (the United States and Germany), a three-factor model was superior, but even it resembled INT-EXT, with an alcohol-related EXT-type factor and two other factors (depression-anxiety and somatization) that were more INT-related.

These results highlight the apparent universality of the INT and EXT factors of psychopathology. These studies also emphasize the promise that the INT-EXT model holds as a conceptualization that functions well for many people regardless of cultural, geographic, socioeconomic, and other individual differences. Indeed, the replication of the two-factor structure from Chile to China, Italy to India, Nigeria to the Netherlands, represents strong evidence in the search for the common underlying structure of numerous major mental disorders.

The Question of INT-EXT Invariance

INT-EXT structures have emerged in a wide variety of samples, including in data from women, men, and individuals of diverse ages and cultural/national origins. While these structures were putatively similar in terms of (1) the presence of INT and EXT factors and (2) on which factor(s) the indicator disorders loaded, there has been relatively little attention paid to testing this similarity in a formal way. The similar number of latent factors, and the similar pattern of where disorders load on those factors, observed in different sub-groups (e.g., women and men) suggests only a basic, descriptive level of structural similarity—not that the statistically *same* model is present across sub-groups. To test closer degrees of model similarity between sub-groups (e.g., equivalent factor loadings, intercepts, etc.), one must employ formal statistical

approaches, referred to as *factorial invariance* analyses, which specifically evaluate these possibilities.

The INT-EXT structure's degree of factorial invariance across different populations is a non-trivial issue, because meaningful comparisons of latent factor means, and of observed disorder prevalence rates, between groups require that these latent structures show particular levels of invariance (Gregorich, 2006; Horn & McArdle, 1992; Vandenberg & Lance, 2000). Further, an invariant comorbidity structure would indicate that observed group differences in prevalence rates are not due to unique features of each disorder but instead result from differences at the latent liability level of comorbidity structure. In such a scenario, the latent factors themselves would be the source of between-group prevalence rate differences as well as the source of comorbidity. Thus, research on group differences (e.g., gender differences) in disorder prevalence rates would benefit from considering the role of these factors, if they are indeed invariant; interventions targeting these influential latent liabilities would similarly seem a worthwhile avenue for future treatment research (see Barlow, Allen, & Choate, 2004; Barlow et al., 2011).

To this point, there has been very little investigation of the potential factorial invariance of the INT-EXT model across meaningful sub-groups. This dearth of research is striking given the potential promise that structural psychopathology research may hold. For example, it appears likely that INT and EXT research may serve to frame part of the "meta-structure"—the organizational scheme—of the upcoming *DSM-5* and *ICD-11* classification systems (see, e.g., Andrews et al., 2009; Regier et al., 2011). Further, recent research has indicated that INT and EXT liability factors have more than

descriptive utility: These dimensional factors have been shown to account for the continuity and development of comorbidity over time (Kessler et al., 2011) and, as noted above, they account for genetic liability to many Axis I and II disorders (Kendler et al., 2003, 2011; Lahey et al., 2011) and can serve as foci for treatment (e.g., Barlow, Allen, & Choate, 2004; Barlow et al., 2011). Unfortunately, it is currently unknown whether the apparent utility of INT and EXT for improving nosology and understanding etiology is impacted by membership in most sub-groups, such as those defined by gender and race/ethnicity.

Age invariance. As discussed above, the INT-EXT structure has been demonstrated in data from young children (e.g., Achenbach & Edelbrock,1978; 1984; Lahey et al., 2008) as well as samples ranging from ages 15 to over 98 (e.g., Eaton et al., 2011; Keyes et al., 2011; Krueger, 1999). Based on these findings, a logical research question was whether a single (factorially invariant) INT-EXT could function well across the lifespan. One study investigated the structural invariance of a hierarchical factor that resembled INT in several ways (Teachman, Siedleck, & Magee, 2007). The authors compared structural models of different age groups on state arousal, trait anxiety, general well-being, neuroticism, and state positive and negative affect. The best-fitting model, a hierarchical "tripartite" model, comprised of a higher-order negative affect factor with anxious-arousal and low positive affect subfactors, bore some resemblance to the three-factor INT model derived from previous research. Further analyses suggested that this model was invariant across different age groups.

Although these results are heartening in the search for an age unbiased conceptualization of psychopathology and are an important step in this direction, the

variables utilized can only serve as a proxy for psychopathology. Further, there are at least three different ways in which age invariance can be investigated. First, one can examine between-cohort group differences, as done by Teachman and colleagues (2007). Such an approach is informative and is permissible using easily collected/accessible cross-sectional data. Inferences about the invariance of INT-EXT as individuals age from such an approach are limited by the design, however, given the potential impact of confounding influences (e.g., cohort effects). The second means by which one could investigate invariance across the lifespan would be to follow a cohort longitudinally, and thus test the latent psychopathology structures' invariance at Time 1 and again at Time 2. This approach allows one to investigate whether the latent factors are invariant or experience within-individual change over time. A third approach, using a cross-sequential design, could combine the previous two methods and thus follow multiple cohorts of individuals longitudinally, permitting inferences about between-cohort group differences as well as within-individual change over time.

A recent study applied the cross-sequential design above to test INT invariance. In that study, three age cohorts (individuals 35 years and under, 36 to 50 years, and over 50 years) were followed longitudinally for approximately a decade in a probability sample of 7,108 individuals (Eaton, Krueger, & Oltmanns, 2011). Through formal tests of factorial invariance, the investigators demonstrated a notable degree of invariance between age cohorts; invariance within cohorts over time was also established. Taken together with the results from Teachman and colleagues (2007), these results suggest that the latent structures of psychopathology—and particularly INT—likely show notable levels of age invariance.

Gender invariance. Gender invariance would suggest that the INT-EXT structure captures a wide variety of psychopathology well in both men and women. If INT-EXT were not invariant, it is possible that men's mental disorders might best be captured by one model (e.g., a one-factor "general psychopathology" model) while women's might best be captured by another (e.g., a hierarchical INT-EXT model). Such an outcome could seriously limit the utility and universality of structural models of psychopathology.

Unlike the topic of age invariance, there has been very little work examining the potential gender invariance of INT-EXT. I am aware of only one study (Kramer, Krueger, & Hicks, 2008) that formally tested the potential factorial invariance of the INT and EXT latent factors, and that study had several significant methodological limitations. This study examined the gender invariance of the INT-EXT structure that resulted from analysis of 11 dimensional syndrome scales in a sample of middle-aged Minnesota Twin Registry participants (N = 2,992). A strong invariance model provided the best fit to the data, indicating substantial similarity of the INT-EXT structures that emerged in women and in men. While these results are encouraging theoretically, the study's generalizability is limited by the demographics of its sample, the moderate sample sizes when genders were modeled separately (ns = 1,880 women and 1,112 men), and the focus on past-month syndrome scales rather than on longer-term *DSM-IV* disorder diagnoses. While modeling continuous syndrome data can be very informative, modeling diagnoses as latent variable indicators addresses a different question and has clearer interpretations with regard to the role of latent comorbidity factors in observed disorder prevalence rates. Further, past-month assessment may capture current and

simultaneous presentation of disorders well, but it also restricts comorbidity patterns and decreases symptomatic variance. As such, most structural psychopathology studies use either 12-month or lifetime diagnoses.

The only other study of which I am aware that examined the potential gender invariance of latent psychopathology factors examined only EXT (Hicks et al., 2007). The results of this study, like those of Kramer, Krueger, and Hicks (2008), suggested a largely gender invariant EXT structure. While this study modeled diagnoses, and assessments of longer duration, its sample too was somewhat restricted: a sample of 626 twin pairs from Minnesota assessed at ages 17 and 24. Again, these results are compelling, but findings from a representative sample—and findings that included INT as well as additional indicators of EXT—would be informative in determining the level of gender invariance of the INT-EXT structure.

In summary, there is limited but promising evidence that the INT-EXT structure may show gender invariance. More definitive research is necessary to address this topic. In particular, research using a nationally representative sample, *DSM-IV* diagnoses, longer diagnostic assessment timeframes, and wider age ranges would be informative.

The question of INT and EXT's potential gender invariance is a non-trivial issue, given that gender differences in many common disorder prevalence rates have been well documented in epidemiological studies (for recent reviews, see Grant & Weissman, 2007; Shear, Halmi, Widiger, & Boyce, 2007; Widiger, 2007). For example, 12-month and lifetime prevalence rates from the NCS indicated that women showed markedly higher (and often approximately double) prevalence rates of major depression, dysthymia, generalized anxiety disorder, panic disorder, social phobia, and

specific phobia than did men, while men showed higher prevalence rates of antisocial personality disorder and alcohol and drug dependence (Kessler et al., 1993, 1994). Similar gender differences have been observed in the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC), the largest epidemiological study of psychopathology yet undertaken (Dawson et al., 2010; Keyes, Grant, & Hasin, 2008; Grant et al., 2004; Grant & Weissman, 2007; Trull et al., 2010; Vesga-López et al., 2008).

The origins of these gender differences in prevalence rates are not well understood, although various theories have been posited to explain how they arise. These explanations include response bias, differential service utilization rates, and various biological, social, and demographic influences (see Klose & Jacobi, 2004; Piccinelli & Wilkinson, 2000). Psychological explanations, such as increased rumination in women partially accounting for higher rates of unipolar depression, have also been posited (Nolen-Hoeksema, 1987; Nolen-Hoeksema, Wisco, & Lyubomirksy, 2008).

These theories of gender differences focus primarily on specific disorders, and rarely take into account comorbidity. When gender differences in prevalence rates and the INT-EXT structure of psychopathology are considered simultaneously, the possibility of a unifying model of gender and comorbidity emerges. Specifically, women show significantly higher prevalence rates of INT disorders, while men show significantly higher rates of EXT disorders (Grant & Weissman, 2007; Kessler et al., 1993, 1994). This observation suggests that gender differences in categorical prevalence rates might be due to gender differences in latent INT and EXT liability dimensions.

However, the establishment of factorial invariance is required for these inferences to be tenable. The utility of latent structural models for public health, epidemiology, psychopathology, and intervention research would be notably enhanced if they can encompass the role of gender in mental disorder prevalence by means of factorial invariance.

Ethnicity invariance. Like, studies of gender differences, epidemiological studies of psychopathology have demonstrated that prevalence rates of many common mental disorders differ across ethnic groups. Consider, for example, the results from the NCS and its subsequent replication (NCS-R), which indicated that lifetime and 12-month prevalence rates of major depressive disorder differed significantly by ethnicity (Blazer et al., 1994; Kessler et al., 1994, 2003). Similarly, findings from the NESARC highlighted ethnicity differences in the prevalence rates of mood, anxiety, alcohol use, drug use, and personality disorders (Grant et al., 2004; Huang et al., 2006; Smith et al., 2006). For instance, although 8.5% of that total sample reported experiencing an alcohol use disorder in the previous year, the prevalence rates differed markedly across ethnic groups: 4.5% of Asian/Pacific Islander, 6.9% of Black, 7.9% of Hispanic/Latino, 8.9% of White, and 12.1% of American Indian/Alaska Native individuals met diagnostic criteria for a past-year alcohol use disorder (Huang et al., 2006).

It is not only prevalence rates that differ across ethnic groups; bivariate comorbidity patterns also differ by ethnicity. Take for example the result that Asian/Pacific Islander individuals reported the lowest rates (4.5%) of 12-month alcohol use disorders in the study by Huang and colleagues (2006). In Asian/Pacific Islander individuals who *did* meet criteria for an alcohol use disorder, however, there was a

significantly higher risk of having a comorbid personality disorder (OR = 7.2) than there was in White (OR = 2.3), American Indian/Alaska Native (OR = 2.5), and Black (OR = 2.8) individuals with alcohol use disorders. This suggests that ethnicity is associated not only with risk to experience mental disorders but also with how disorders tend to present comorbidly.

The origins and meaning of ethnicity differences in prevalence rates and bivariate comorbidity patterns are important issues, because they may provide important clues about etiology and inform public health initiatives. A deeper understanding of these differences may emerge by modeling multivariate comorbidity explicitly. Ethnicity differences in latent factor means could account for observed ethnicity differences in prevalence rates and comorbidity, for instance, but the establishment of ethnicity invariance is required to support such an inference. However, I am unaware of any studies that have examined, in a factorial invariance framework, whether the underlying structure of common mental disorders differs by ethnicity. It is currently unknown whether the apparent utility of INT-EXT for improving nosology, understanding etiology, and guiding intervention is impacted by ethnicity. Given that prevalence rates and bivariate comorbidity patterns differ by ethnicity, multivariate invariance cannot simply be assumed, and formal evaluation is critical.

The Current Study

There remains an open question as to whether the INT-EXT model shows factorial invariance across gender and ethnicity. The current study was conducted to fill this gap in the literature by examining the structure of common mental disorders across (1) women and men, and (2) five ethnic groups—White, Hispanic/Latino, Black,

Asian/Pacific Islander, and American Indian/Alaska Native individuals—in a large nationally representative sample (N = 43,093). My specific aims were as follows. First, I would investigate common mental disorder prevalence rates, using lifetime and 12-month diagnoses, across gender and ethnic groups. Second, I would explore the latent comorbidity structure separately in the total sample, both genders, and each ethnic group. Third, I would conduct formal factorial invariance testing of the resulting latent comorbidity structure(s) of common mental disorders across the gender and ethnic groups. Finally, if invariance were established, I would compare the latent factor means in the gender and ethnic groups to determine the directionality and magnitude of liability differences. Such latent mean differences in an invariant structure would manifest as observed prevalence rate differences.

Method

Participants

I analyzed data from the 43,093 individuals who participated in the NESARC (2001-2002). The NESARC is a representative sample of the civilian, non-institutionalized United States population at least 18 years of age. Hispanic/Latino individuals, non-Hispanic Black individuals, and young adults were oversampled. Women composed 57% (n = 24,575) of the sample; the age range was 18-98 years. Race/ethnicity was selected by participants and defined via Census Bureau algorithms into: non-Hispanic/Latino White ("White" hereafter for brevity; 56.9%; n = 24,507), Hispanic or Latino ("Hispanic;" 19.3%; n = 8,308), non-Hispanic/Latino Black ("Black;" 19.1%; n = 8,245), non-Hispanic/Latino Asian/Native Hawaiian/Pacific Islander ("Asian/Pacific Islander;" 3.1%; n = 1,332), and non-Hispanic/Latino American Indian/Alaska Native ("American Indian/Alaska Native;" 1.6%; n = 701). The NESARC was weighted to be representative of the age, sex, and racial/ethnic distribution of the United States based on the 2000 census. For more information on the sampling and design of the NESARC, see Grant and Dawson (2006). The research protocol, including written informed consent, received full ethical review and approval from the U. S. Census Bureau and the U. S. Office of Management and Budget.

Assessment

Lifetime and 12-month *DSM-IV* diagnoses were made using the Alcohol Use Disorder and Associated Disabilities Interview Schedule—*DSM-IV* Version (AUDADIS-IV; Grant et al., 1995; 2003), a structured interview designed for administration by experienced lay interviewers. The present study examined major

depressive disorder, dysthymic disorder, generalized anxiety disorder, panic disorder, social phobia, specific phobia, alcohol dependence, nicotine dependence, marijuana dependence, other drug dependence, and antisocial personality disorder AUDADIS-IV diagnoses. I calculated the other drug dependence variable by collapsing relatively uncommon forms of drug dependence (i.e., stimulants, opioids, sedatives, tranquilizers, cocaine, solvents, hallucinogens, heroin, and any other drug not assessed) into one variable with sufficient variance for structural modeling. In keeping with the *DSM-IV* conceptualization of PDs as lifelong, antisocial PD was assessed only as a lifetime diagnosis, which was used in both lifetime and 12-month analyses.

The reliability levels of the AUDADIS diagnoses are generally good (e.g., kappas = .42 to .84) for mood and anxiety disorders (Canino et al., 1999; Grant et al., 2003; Ruan et al., 2008), substance use disorders (Grant et al., 1995; Chatterji et al., 1997; Hasin et al., 1997; Grant et al., 2003), and antisocial PD (Grant et al., 2003). The test-retest estimates for psychiatric disorders in the AUDADIS-IV are similar to other structured interviews (e.g., the DIS, the CIDI) used in large-scale psychiatric epidemiologic surveys (reviewed in Wittchen, 1994). Further, the AUDADIS-IV includes notable advantages over other structured interviews such as the Diagnostic Interview Schedule (DIS), including assessment of clinically significant distress and impairment after the syndrome is fully characterized (see Hasin et al., 2005).

Statistical Analyses

General. Analyses were conducted in Mplus version 6 (Muthén & Muthén, 2011) using the default delta parameterization and a weighted least squares (WLSMV) estimator, which allowed for treatment of dichotomous diagnoses as categorical

variables and to incorporate the NESARC's weighting, clustering, and stratification variables into all analyses. EFAs used the default oblique Geomin rotation as well as Promax. To determine the number of factors to extract, I relied upon scree plot analysis, fit indices, and substantive interpretability.

Fit indices. Fit indices considered in EFA and CFA were the comparative fit index (CFI), Tucker-Lewis index (TLI), and root mean squared error of approximation (RMSEA); in CFA, the number of free parameters in the model was also used. Chi-square goodness-of-fit and difference tests were not employed, because they are sensitive to large sample sizes; other fit indices, such as AIC and BIC, are not available with the WLSMV estimator. CFI/TLI values greater than .95 and RMSEA values less than .06 suggest reasonably good model fit (Hu & Bentler, 1999); the TLI can exceed 1.00 in cases of very good fit. Cheung and Rensvold (2002) conducted simulation studies of common fit indices and proposed a CFI difference critical value of .01 be used in factorial invariance research to determine whether the addition of constraints leads to notably worse model fit. Finally, the number of free parameters is the number of parameters that were freely estimated in the model. As the number of free parameters decreases, model parsimony increases. There were no missing values.

Factorial invariance. Meredith (1993) discussed a means by which increasingly stringent levels of factorial invariance could be tested. Each increasing level of invariance requires that the previous, less stringent levels also be met. The first, and least stringent, level of factorial invariance is *configural invariance*, established when indicator variables load on the same factors across groups. In *metric invariance*, the factor loadings that link each indicator to the latent factors are constrained to be equal

across groups. In *strong invariance*, the intercepts in equations connecting the latent factors with the indicator variables are constrained to equality across groups. Finally, in *strict invariance*, the residual variances for each indicator variable are constrained to equality across groups.

Although the Meredith (1993) set of factorial invariance levels is perhaps the most widely used in the literature, it requires that continuous manifest variables be modeled. Tests of factorial invariance for dichotomous indicators, such as AUDADIS-IV diagnoses, require methodology appropriate for modeling categorical variables (Millsap & Yun-Tein, 2004). In this categorical variable factorial invariance approach, factor loadings and thresholds are constrained to equality or freed, in tandem, across groups. Let us take gender as an example. In gender invariance analyses, my first model—the *unconstrained model*—was parameterized such that loadings and thresholds were free across gender, factor means were set to zero in both genders, and scaling factors were fixed to unity in both genders. I refer to this model as the unconstrained model, because factor loadings and thresholds would be free (unconstrained) to vary across genders. The second model—the *constrained model*—was parameterized such that loadings and thresholds were constrained to equality across gender, factor means were set to zero in men and were free in women, and scaling factors were fixed to unity in men and were free in women. I refer to this second model, which represents a gender invariant psychopathology structure, as the constrained model, because factor loadings and thresholds were constrained to equality across genders.

I conducted multi-group factorial invariance analyses of the latent comorbidity structure of common mental disorders. First, factorial invariance was tested across

gender by comparing latent structures in women and men simultaneously. Second, it was tested across ethnicity by comparing latent structures in the five ethnic groups simultaneously. In each set of analyses, I tested the structures that resulted from lifetime and from 12-month diagnoses separately; thus, I tested whether the lifetime disorder comorbidity structure was invariant across gender/ethnicity and also whether the 12-month disorder comorbidity structure was invariant across gender/ethnicity.

If the more parsimonious constrained model fit as well or better than the unconstrained model, factorial invariance would have been established. If invariance were present, I could conclude that the latent factors represented similar constructs across gender and across ethnicity, and comparison of the features of these latent factors would be empirically justified (Gregorich, 2006; Horn & McArdle, 1992; Vandenberg & Lance, 2000). In such a scenario, I proposed to compare the means of the latent factors across gender to determine whether women and men had different average standings on these latent factors; similarly, I proposed to compare latent factor means across ethnicity. I hypothesized that any differences between the groups in a factorially invariant latent comorbidity structure should manifest as differences in observed disorder prevalence rates. As such, my final analyses tested for significant differences in observed disorder prevalence rates and compared these differences to those expected based on the directionality of latent factor mean differences. Such a finding would further our understanding of multivariate comorbidity, and it would suggest that gender/ethnicity differences in observed disorder prevalence rates are manifestations of the gender/ethnicity mean differences at the level of latent common mental disorder liability factors.

Results

Gender Analyses

Prevalence rates. Table 2 presents the prevalence rates for the disorders included in the current study separately for women and men and for lifetime and 12-month diagnoses. All odds ratios, using men as the comparison group, were significant at $p < .001$ except other drug dependence, which was significant at $p = .005$. Across lifetime and 12-month prevalence rates, women showed significantly higher rates for all INT disorders, and men showed higher rates for all EXT disorders. In most cases, the magnitude of these differences was an approximately doubled prevalence rate in one gender versus the other. For instance, the 12-month prevalence of major depressive disorder in women was 10.1% and in men was 5.5%. The 12-month prevalence of alcohol dependence was 5.4% in men and 2.3% in women.

Model parameterization. As noted above, the question of whether the INT liability dimension should be parameterized as (1) a single dimensional factor, or as (2) two dimensional sub-factors (distress and fear) subsumed under a higher-order INT dimensional factor, remains somewhat unresolved in the literature. Previous research (Eaton et al., 2011; Keyes et al., 2011) with the NESARC sample has indicated good performance of the bifurcated INT structure across genders, and those reports utilized the bifurcated INT model when doing analyses by gender. To be congruent with this research, I opted to use the bifurcated INT structure for the gender analyses. This model was parameterized such that each diagnosis loaded on one of three factors: (1) distress: major depression, dysthymia, and generalized anxiety disorder; (2) fear: panic disorder, social phobia, and specific phobia; and (3) EXT: alcohol dependence, nicotine

dependence, marijuana dependence, other drug dependence, and antisocial PD. Distress and fear were parameterized to load on a higher-order INT factor, which was allowed to correlate with the EXT factor.

Model fit. The INT-EXT model provided a very good fit in the total sample for lifetime (CFI = .992, TLI = .989, RMSEA = .012) and 12-month diagnoses (CFI = .988, TLI = .984, RMSEA = .010). For fit information, see Table 3. Within each gender modeled separately, this INT-EXT model also fit very well. In women, the model provided good fit for lifetime (CFI = .993, TLI = .991, RMSEA = .009) and 12-month diagnoses (CFI = .990, TLI = .987, RMSEA = .008). In men, the model provided good fit for lifetime (CFI = .988, TLI = .984, RMSEA = .008) and 12-month diagnoses (CFI = .982, TLI = .976, RMSEA = .007). Thus, in the total sample, in women separately, and in men separately, the INT-EXT structure fit the data quite well, regardless of whether lifetime or 12-month diagnoses were modeled.

Invariance. Because an INT-EXT model fit well in women and men, our next question was how similar these models were across gender in terms of model parameters—that is, whether the parameters differed by gender or whether they showed invariance. I fit the unconstrained and constrained models in men and women simultaneously via a multi-group CFA, separately for lifetime and 12-month diagnoses (see Table 3). For lifetime diagnoses, the fits of the unconstrained (CFI = .991, TLI = .989, RMSEA = .012) and constrained (CFI = .991, TLI = .989, RMSEA = .012) models were identical, but the constrained model had fewer freely estimated parameters (k = 38) than the unconstrained model (k = 48). The constrained model for lifetime diagnoses is depicted in Figure 3.

For 12-month diagnoses, the constrained model (CFI = .988, TLI = .986, RMSEA = .009, k = 38) had a better fit with greater parsimony than did the unconstrained model (CFI = .987, TLI = .983, RMSEA = .010, k = 48). For lifetime and 12-month diagnoses, the CFI critical difference of .01 was not exceeded, further supporting the constrained model. These findings indicated that, in addition to the general configural structure, factor loadings and thresholds for all diagnoses could be conceptualized well as equivalent for women and men. Thus, the structure of these common mental disorders, including the connections between individual diagnoses and the underlying factors, was gender invariant.

Factor means. To investigate potential differences in factor means, the means of the latent INT and EXT factors were fixed to zero in men and freely estimated as .445 and -.378 in women, respectively, for lifetime diagnoses and as .428 and -.308 for 12-month diagnoses. All mean gender differences were significant at $p < .01$. These standardized means can be interpreted as z-scores (e.g., women were approximately .45 standard deviations *higher* on lifetime INT liability than men). Because complete factorial invariance had been established, these results demonstrated that the observed gender differences in the prevalence rates of these disorders could be accounted for by the genders' different average levels of latent INT and EXT.

Ethnicity Analyses

Prevalence rates. Table 4 presents the prevalence rates for the disorders included in the current study separately for each ethnic group and for lifetime and 12-month diagnoses. Chi-square analyses indicated all disorders showed significant ($p <$ .001) prevalence rate differences across the five ethnicities. This was true for lifetime

and 12-month diagnoses. As a general trend, Asian/Pacific Islander individuals tended to report the lowest rates of most disorders; Black and Hispanic individuals reported approximately similar rates to each other, which were slightly higher than those of Asian/Pacific Islander individuals; White individuals reported higher rates than Black, Hispanic, and Asian/Pacific Islander individuals; and American Indian/Alaska Native individuals tended to report the highest rates of most of these forms of psychopathology. Thus, the pattern of prevalence differences appeared to cut across INT and EXT groupings of disorders, with certain ethnic groups reporting relatively higher or lower levels of INT *and* EXT disorders in tandem. This pattern of ethnicity differences stood in contrast to the pattern of gender differences, wherein women reported higher levels of INT disorders, and men reported higher levels of EXT disorders.

Model parameterization. For ethnicity analyses, I chose to parameterize the INT factor as a single dimension, rather than a bifurcated model, for three reasons. First, some ethnic groups had relatively small sample sizes. As such, imposing a fine-grained distinction between the highly correlated distress and fear sub-dimensions seemed analytically questionable. As sample sizes increase, random error of interrelations among the diagnoses decreases; thus, small sample sizes have a higher likelihood of improperly indicating areas of misfit. The more complex bifurcated INT model had more potential for such misfit to occur. Second, using a unitary INT factor for the ethnicity analyses, and a bifurcated INT factor for the gender analyses, would allow for the testing of invariance of both the unitary and bifurcated INT-EXT structures.

The third rationale for choosing a unitary INT factor reflected the results of EFA. Unlike the strong support for a bifurcated INT-EXT structure in the NESARC in the total sample and across genders from previous research, there was no evidence, of which I was aware, that suggested the superiority of a unitary or bifurcated INT-EXT structure in different ethnic groups. That said, the structural psychopathology study that examined data from the most countries worldwide also employed a unitary INT-EXT model (Krueger et al., 2003). As such, EFA was conducted within each ethnicity separately for lifetime and 12-month diagnoses. In all cases, scree plot analyses indicated that a two- or three-factor solution was optimal (fit indices for one-factor solutions also did not reach acceptable thresholds), so I focused on fit indices and substantive interpretability. Tables 5 and 6 report Geomin rotated factor loadings and fit indices for two- and three-factor solutions, for each ethnicity, and for lifetime and 12-month diagnoses, respectively.

The two-factor models generally fit quite well across all ethnicities, for lifetime and 12-month diagnoses (Tables 5 and 6). These two-factor models were identified as INT-EXT models based on patterns of disorder loadings with no sizeable disorder cross-loadings for the lifetime diagnoses. The 12-month diagnoses showed a few minor departures from this simple structure (e.g., other drug dependence cross-loaded in Black, Asian/Pacific Islander, and American Indian/Alaska Native individuals, and social and specific phobia cross-loaded in American Indian/Alaska Native individuals). However, these results largely showed the same INT-EXT two-factor solution. I interpreted these results as being supportive of the generally robust presence of INT and EXT factors across ethnicities and across diagnostic duration.

The three-factor models also fit the data well (Tables 5 and 6). While some models replicated the bifurcated INT-EXT model (with distress and fear sub-factors of INT based on patterns of disorder loadings), others showed considerable cross-loadings and bore little resemblance to the bifurcated INT-EXT model. Several three-factor models lacked clear substantive interpretability and some produced Heywood cases (implausible values of factor loadings) using both Geomin and Promax rotations, suggesting three-factor solutions were over-extractions in some groups. Results indicated the highly correlated distress and fear INT sub-factors were thus not reliably distinguished across ethnicities.

All three rationales supported the modeling of INT as a single factor in the ethnicity analyses. As such, major depression, dysthymia, generalized anxiety disorder, panic disorder, social phobia, and specific phobia were parameterized to load on a single INT factor. Antisocial PD and alcohol, nicotine, marijuana, and other drug dependence loaded on an EXT factor.

As reported in Table 7, the two-factor structure fit well in each ethnicity modeled separately via CFA. In all five ethnicities, using both lifetime and 12-month diagnoses, every fit index ranged from good to excellent; no fit index fell below threshold for acceptable fit. In White individuals, the two-factor model fit well for lifetime (CFI = .983, TLI = .978, RMSEA = .014) and 12-month (CFI = .978, TLI = .971, RMSEA = .011) diagnoses. In Hispanic/Latino individuals, the two-factor model fit well for lifetime (CFI = .987, TLI = .983, RMSEA = .005) and 12-month (CFI = .973, TLI = .966, RMSEA = .005) diagnoses. In Black individuals, the two-factor model fit well for lifetime (CFI = .980, TLI = .974, RMSEA = .007) and 12-month (CFI

= .976, TLI = .969, RMSEA = .005) diagnoses. In Asian/Pacific Islander individuals, the two-factor model fit well for lifetime (CFI = .988, TLI = .985, RMSEA = .002) and 12-month (CFI = .989, TLI = .986, RMSEA = .001) diagnoses. In American Indian/Alaska Native individuals, the two-factor model fit well for lifetime (CFI = 1.000, TLI = 1.003, RMSEA = .000) and 12-month (CFI = .962, TLI = .952, RMSEA = .013) diagnoses. These results indicated that the INT-EXT model was a well fitting model of disorder comorbidity structure in all ethnicities using both lifetime and 12-month diagnoses.

Invariance. I next fit the unconstrained and constrained models in the five ethnic groups simultaneously via a multi-group CFA, separately for lifetime and 12-month diagnoses. As reported in Table 7, for lifetime diagnoses, the unconstrained multi-group model (CFI = .984, TLI = .979, RMSEA = .016, k = 115) fit slightly worse on all fit indices than did the more parsimonious constrained model (CFI = .987, TLI = .985, RMSEA = .013, k = 75). For 12-month diagnoses, the unconstrained model (CFI = .974, TLI = .967, RMSEA = .013, k = 115) fit slightly worse on all fit indices than did the more parsimonious constrained model (CFI = .978, TLI = .976, RMSEA = .011, k = 75). The .01 critical difference in CFI values proposed by Cheung and Rensvold (2002) was not reached when constraints were placed on the model. These results indicated that, for both lifetime and 12-month diagnoses, the INT-EXT model of common mental disorders was invariant across the ethnicities. The constrained lifetime model is depicted in Figure 4.

Factor means. The INT and EXT means among White individuals were both fixed to zero and served as a reference metric (due to White individuals composing the

largest ethnicity sub-group) for the means of the other ethnic groups. Table 7 details the standardized means and significant ($p < .05$) liability differences between ethnic groups. Given the ethnicity invariant structure, we can infer these ethnicity differences in latent INT and EXT factor means gave rise to the observed ethnicity differences in prevalence rates.

Discussion

General Findings

In this study, I examined the structure, and potential invariance, of common mental disorders across gender and ethnic groups. Analyses in the total sample, in both genders, and in each ethnic group suggested that comorbidity patterns could be accounted for by two broad latent factors: INT and EXT. This finding replicated past findings of configural invariance across datasets, such that the pattern of disorder relations to the latent factors was the same across sub-groups. Configural invariance is a low hurdle for determining the statistical similarity of two factor models, however, and more formal tests were necessary to provide a "risky test" of hypothesized gender and ethnicity invariance. Subsequent multi-group factorial invariance analyses of the INT-EXT structures indicated that a single structure—with identical factor loadings and indicator thresholds—functioned well in women and in men. Similarly, a single structure functioned well in all five ethnic groups.

Previous studies have indicated that apparently *similar* latent factors exist across diverse sub-groups of individuals but only rarely have studies tested this similarity statistically (see Eaton, Krueger, & Oltmanns, 2011; Hicks et al., 2007; Kramer et al., 2008). The present study extends this observation by finding that these latent structures are not only similar, but, statistically speaking, they are the *same*. In other words, an INT-EXT model, with a single set of parameters, could be fit well, and parsimoniously, to the data across genders and across ethnic groups. Thus, the INT and EXT factors modeled in this study were found to be both gender and ethnicity invariant.

The finding of gender invariance replicated previous research suggesting that the INT and EXT factors were gender invariant (see Hicks et al., 2007; Kramer et al., 2008), while also resolving many of the sampling limitations of those studies. In addition, by modeling disorders rather than syndromes scales, the current study could support stronger inferences about the origins of prevalence rate differences than could previously be made. These results indicate that the prevalence rate differences observed in the NESARC between women and men reflect significant gender differences in the average levels of INT and EXT liability.

The finding of ethnicity invariance was a novel contribution to the literature. While previous research had identified INT and EXT factors underlying comorbidity in samples from diverse nations, there had been no formal investigation of whether the emergent factors represented the same constructs in terms of statistical definition. Given frequently reported ethnicity differences in prevalence rates reported in epidemiological studies, this dearth of information about latent structure is striking. The results of the current study indicated that, across the ethnicity groups of the NESARC, INT and EXT factors underpinned the comorbidity of the disorders modeled. These INT and EXT factors could be defined identically in a statistical sense while still providing excellent fit to the data via a parsimonious model. The implication of this ethnicity invariance is that the prevalence rate differences observed in the NESARC between White, Hispanic, Black, Asian/Pacific Islander, and American Indian/Alaska Native individuals reflect significant ethnicity differences in the average levels of INT and EXT liability.

Generally, these results suggest that, as sub-groups differed on their average levels of the relatively undifferentiated *latent* INT and EXT liability factors, so too do

they differ on their levels (prevalence rates) of *manifest* diagnoses. In this way, the findings of factorial invariance suggest that disorder prevalence rate differences reflect differences in core, underlying psychiatric liabilities rather than meaningful disorder-specific prevalence rate differences. The significant factor mean differences that were estimated across gender and ethnic groups appear in these data to be the origination point for prevalence rate differences reported by previous epidemiological research and observed presently in the NESARC sample as well.

Implications for Classification

DSM-IV is currently under revision, and there has been a great deal of discussion about the general organization of *DSM-5* (Regier et al., 2011). Influenced by the replications of the INT-EXT structure in various samples in the literature, some researchers have proposed that an INT-EXT organizational "meta-structure" be used to frame many common mental disorders in *DSM-5* (Andrews et al., 2009). In such an organization, the unipolar mood and anxiety disorders, for instance, might be placed near each other in the document to highlight their strong degree of overlap; relatedly, both sorts of disorders might be subsumed under a general heading of "Internalizing Disorders."

The results of the current study support a proposed reorganization of the *DSM-5* (and *ICD-11*) to reflect an INT-EXT meta-structure for many common disorders. While there had been indications that these latent comorbidity factors were invariant across gender (Hicks et al., 2007; Kramer et al., 2008) and the lifespan (Eaton, Krueger, & Oltmanns, 2011), the current study contributed two critical missing pieces to the evaluation of this meta-structure. First, I demonstrated in a large, nationally

representative sample that the latent structure of common psychiatric diagnoses is invariant across gender, which expanded the limited previous work on the question. Second, I demonstrated, for the first time, that the latent structure of common psychiatric diagnoses is invariant across ethnicity. These findings, coupled with previous results, suggest that the use of INT and EXT disorder groupings would be an appropriate framing of patients' psychopathology regardless of their ages, genders, or ethnic identities.

Implications for Prevalence Rate Differences Conceptualization and Research

Gender. The conclusion that observed gender differences in prevalence rates systematically reflect differences in broad latent liability factors can unify the piecemeal gender differences literature, which typically focuses on one specific disorder at a time. For instance, one prominent theory to account for gender differences in major depressive disorder posits that women ruminate more frequently than men, focusing repetitively on their negative emotions and problems; this rumination stands in contrast to men's increased likelihood of engaging in more active problem solving behaviors (Nolen-Hoeksema, 1987; Nolen-Hoeksema, Wisco, & Lyubomirksy, 2008). The current results indicate that this theory need not stop with unipolar mood disturbance, however, and suggest that similar mechanisms may help account for gender differences in anxiety as well. Along these lines, the theory can be readily extended to anxiety (and other INT disorders) by noting that neuroticism, or negative affectivity, is strongly related to rumination, such that individuals who are more neurotic ruminate more frequently (Lam, Smith, Checkley, Rijsdijk, & Sham, 2003). Neuroticism is also very strongly related to the latent INT factor (r = .98; Griffith et al., 2010), a link accounted for

largely by genetic effects (Hettema et al., 2006). Finally, previous research has indicated that women tend to report higher levels of trait neuroticism (as well as conscientiousness and agreeableness) on average than do men (e.g., Donnellan & Lucas, 2008), which mirrors the current finding that women were estimated to have significantly higher mean levels of INT than men. It may be through neuroticism (and disinhibition-related traits in the case of EXT and men; e.g., Krueger et al., 2002; Miller & Lynam, 2001; Slutske et al., 2002) and its related psychological processes, such as rumination, that the latent liabilities to experience comorbid mental disorders are manifested and maintained.

Environmental influences also likely play a key role in gender (and ethnicity) differences in mental disorder prevalence rates. Given that women tend to report higher frequencies of some stressful life events than men prior to disorder onset (Harkness et al., 2010), understanding the relations between environmental stressors and latent comorbidity factors seems a particularly worthwhile focus for group differences research. Several possible association patterns are possible. For instance, in a diathesis stress model, one could conceptualize the latent factors as the diatheses and the environmental influences as stressors that activate the diatheses. There is little research on the topic, but it appears that negative environmental influences are prospectively predictive of increased latent factor levels (Eaton, 2012; Keyes et al., 2011). Alternatively, individuals with higher levels of comorbidity factors may tend to experience—or perceive—more negative life events.

Ethnicity. Because the INT-EXT structure was invariant across ethnicity, we can infer that observed prevalence rate ethnicity differences in categorical diagnoses of

many common disorders reflect ethnicity differences in latent dimensional factors. This finding supports latent comorbidity factors as major targets of inquiry for investigations of ethnicity differences in psychopathology. It appears that it is no longer sufficient to ask solely whether and why ethnic groups (or gender groups, or perhaps most groups of individuals) differ on particular categorical disorders. Instead, the research literature could benefit from investigations of whether and why the unifying dimensional latent liability factors differ across groups. Investigations of particular categorical disorders, while valuable, do not capture the broader, generalizable themes that latent factors represent.

Implications for Intervention

The results of the current study are congruent with findings that many different disorders respond to the same treatment modalities. In terms of psychopharmacological interventions, symptoms of anxiety and depression often improve with administration of tricyclics, SSRIs, and newer antidepressants (Goldberg et al., 2011). Although the mechanism of action for this effect is unclear, one compelling explanation, relating to the present findings, is that these psychopharmacological approaches may target the latent, shared core of various disorders—in the case of anxiety and mood disorders, the medications would impact latent INT levels—rather than particular disorders or symptomatological patterns (see Tang et al., 2009).

Many psychotherapeutic interventions similarly demonstrate effects on diverse manifestations of psychopathology. For example, cognitive-behavioral therapy is an empirically supported treatment for various mood and anxiety disorders (Nathan & Gorman, 2007). Even breaking up cognitive-behavioral therapy into its component parts

does not negate its diffuse effectiveness. Cognitive restructuring challenging maladaptive thoughts is effective for unipolar mood and anxiety disorders; behavioral activation to expose depressed individuals to positive emotions functions similarly to the exposure-based golden standard treatment of anxiety disorders. This effectiveness of single psychotherapeutic interventions to target multiple disorders has prompted some researchers to propose "transdiagnostic" treatment approaches. In such a modality, a set of generalized strategies is presented to the patient as a means to control a variety of subjectively diverse symptoms. For instance, Barlow and colleagues have developed a unified treatment protocol for unipolar mood and anxiety disorders that combines elements of cognitive restructuring, motivational interviewing, behavioral activation and exposure, and mindfulness (Barlow et al., 2011). The current findings support the applicability of such treatment protocols across genders and ethnicities, because their targets—the shared core of many common mental disorders—appear to be similar across groups.

Limitations

This study had several limitations. First, the NESARC's diagnostic interviews were administered by extensively trained lay interviewers rather than clinicians. This limitation is mitigated somewhat by the AUDADIS-IV's structured design and generally good psychometric properties. Second, lifetime prevalence rates based on retrospective reporting can be biased, and there is evidence that many prospectively assessed lifetime disorder prevalence rates are approximately double the rates obtained from retrospective lifetime reporting (Moffitt et al., 2010). While lifetime diagnoses may have a retrospection bias, this study supplemented lifetime diagnoses with 12-

month diagnoses. Results for 12-month diagnoses, which required comparatively much less retrospection, were fully consistent with those for lifetime diagnoses. Third, the current study investigated only common mental disorders and thus did not include other debilitating forms of psychopathology, such as schizophrenia. There are indications that some symptoms of psychotic disorders may relate to a separate latent factor (e.g., "thought disorder") while also showing associations with INT and neuroticism (e.g., Barrantes-Vidal, Ros-Morente, & Kwapil, 2009; Kotov et al., 2011a, 2011b; Markon, 2010; Wolf et al., 1988). Finally, the current study addressed diagnostic comorbidity with a focus on prevalence rates; thus, I modeled *DSM-IV* diagnoses. Future studies would benefit from also examining symptom-level data (e.g., Markon, 2010), which addresses a related but different set of questions.

Table 1

Illustration of disorders as independent (top panel) and correlated (bottom panel)

Generalized anxiety disorder and major depression as independent disorders:

	Generalized Anxiety		
Depression	Absent	Present	Marginals
Absent	844	23	867
Present	129	**4**	133
Marginals	973	27	$N = 1{,}000$

Generalized anxiety disorder and major depression as correlated disorders:

	Generalized Anxiety		
Depression	Absent	Present	Marginals
Absent	858	10	867
Present	115	**17**	133
Marginals	973	27	$N = 1{,}000$

Note: Independent presentation table represents *expected* values calculated from population prevalence rates of major depression and generalized anxiety disorder in 7,108 individuals from the Midlife Development in the United States (MIDUS) study, which utilized a national probability sample. Correlated presentation table represents *observed* values of MIDUS disorder comorbidity. Values were scaled to $N = 1{,}000$ for simplicity of presentation.

Table 2

Lifetime and 12-month DSM-IV *disorder prevalence rates (as percentages) for women and men*

	Lifetime Disorders			12-Month Disorders		
	Women	Men	Odds Ratio	Women	Men	Odds Ratio
Depression	22.9	13.1	1.46 (1.41-1.51)	10.1	5.5	1.38 (1.32-1.45)
Dysthymia	6.2	3.5	1.31 (1.25-1.38)	2.9	1.6	1.29 (1.20-1.39)
Generalized Anxiety	5.8	3.1	1.34 (1.27-1.42)	3.1	1.4	1.37 (1.28-1.47)
Panic Disorder	7.2	3.7	1.39 (1.32-1.47)	3.1	1.4	1.39 (1.29-1.49)
Social Phobia	5.8	4.3	1.16 (1.10-1.22)	3.4	2.1	1.22 (1.15-1.29)
Specific Phobia	12.4	6.2	1.47 (1.41-1.53)	9.6	4.6	1.46 (1.40-1.53)
Alcohol Dependence	8.0	17.4	0.63 (0.60-0.65)	2.3	5.4	0.68 (0.64-.0.72)
Nicotine Dependence	15.6	20.0	0.84 (0.81-0.87)	11.5	14.1	0.88 (0.85-0.91)
Marijuana Dependence	0.9	1.7	0.77 (0.71-0.83)	0.2	0.5	0.72 (0.63-0.83)
Other Drug Dependence	1.4	2.2	0.84 (0.79-0.90)	0.3	0.5	0.83* (0.72-0.95)
Antisocial Personality	1.9	5.5	0.68 (0.59-0.66)			

Note: All ORs significant at $p < .001$ except *$p = .005$. Men are OR comparison group. 95% confidence intervals are given in parentheses. Antisocial PD was only assessed as a lifetime disorder.

Table 3

Model fit statistics

	CFI	TLI	RMSEA	# Free
Total sample (*N* = 43,093)				
Lifetime diagnoses	.992	.989	.012	--
12-month diagnoses	.988	.984	.010	--
Women (*n* = 24,575)				
Lifetime diagnoses	.993	.991	.009	--
12-month diagnoses	.990	.987	.008	--
Men (*n* = 18,518)				
Lifetime diagnoses	.988	.984	.008	--
12-month diagnoses	.982	.976	.007	--
Multi-group (Women and Men)				
Lifetime diagnoses				
Unconstrained model	.991	.989	.012	48
Constrained model	.991	.989	.012	38
12-month diagnoses				
Unconstrained model	.987	.983	.010	48
Constrained model	.988	.986	.009	38

Note: Total sample analyses modeled women and men together. Multi-group analyses modeled women and men simultaneously as two separate groups. Unconstrained models allowed each gender to have unique model parameters; constrained (invariant) models constrained factor loadings and thresholds to equality across genders. CFI: comparative fit index. TLI: Tucker-Lewis index. RMSEA: root mean squared error of approximation. # Free: number of freely estimated parameters.

Table 4

Lifetime and 12-month DSM-IV *disorder prevalence rates (as percentages) for the five ethnic groups*

	Lifetime Disorders					12-Month Disorders				
	W	H	B	A	N	W	H	B	A	N
Depression	19.9	13.4	13.4	11.1	27.5	8.3	6.4	7.1	5.3	13.3
Dysthymia	5.3	3.4	4.2	3.0	8.2	2.3	1.8	2.3	1.6	4.0
Generalized Anxiety	5.0	3.0	3.4	2.3	6.5	2.4	1.8	2.1	1.4	3.0
Panic Disorder	6.1	4.0	3.8	2.3	1.0	2.4	1.8	1.7	0.9	4.8
Social Phobia	5.5	3.2	3.6	3.4	9.1	3.0	2.0	2.1	2.2	3.7
Specific Phobia	10.0	7.6	9.2	5.9	12.0	7.6	5.7	7.3	4.2	8.2
Alcohol Dependence	13.8	9.5	8.4	6.0	20.1	3.8	4.0	3.6	2.4	6.3
Nicotine Dependence	20.1	8.7	13.1	8.1	30.3	14.3	6.3	10.3	6.4	23.1
Marijuana Dependence	1.3	1.0	1.2	0.7	3.1	0.3	0.3	0.4	0.2	1.3
Other Drug Dependence	1.8	1.5	1.5	0.6	5.3	0.3	0.5	0.4	0.2	1.5
Antisocial Personality	3.6	3.3	3.7	1.8	9.7					

Note: Letters denote: White, Hispanic, Black, Asian/Pacific Islander, and American Indian/Alaska Native; see text for full description of ethnic groups. All prevalence rates differed significantly across groups ($p < .001$). Antisocial PD was assessed only as a lifetime disorder.

Table 5

Exploratory factor analysis results by ethnicity for lifetime diagnoses

	White					Hispanic					Black					Asian/ Pacific Islander					American Indian/ Alaska Native				
	1	2	1	2	3	1	2	1	2	3	1	2	1	2	3	1	2	1	2	3	1	2	1	2	3
MDD	**.87**	.01	**.88**	.01	.04	**.77**	.09	**.74**	.12	-.06	**.85**	.00	**.70**	.26	.00	**.70**	.19	**.86**	-.01	.08	**.81**	.15	**.82**	.14	-.01
Dysth	**.85**	-.01	**.88**	.03	-.06	**.85**	-.01	**.88**	-.01	-.33	**.81**	.02	**.91**	.02	.00	**.60**	.25	**.68**	.11	.12	**.83**	-.01	**.83**	.00	-.03
GAD	**.80**	-.04	**.56**	-.08	.37	**.86**	-.06	**.81**	.00	.01	**.84**	-.04	**.48**	**.47**	-.01	**.84**	.00	**.91**	-.16	-.02	**.88**	-.05	**.87**	-.07	.05
Panic	**.56**	.14	.28	.10	**.41**	**.63**	.08	**.62**	.11	.12	**.66**	.02	.31	**.44**	.05	**.69**	-.01	**.61**	.04	-.27	**.58**	.24	**.56**	.00	**.77**
Social	**.59**	.07	.17	.01	**.59**	**.67**	.03	**.64**	.08	.27	**.66**	.02	.01	**.72**	.09	**.64**	.20	**.41**	.38	**-.43**	**.67**	.02	**.67**	.02	-.03
Spec	**.47**	.15	-.01	.07	**.70**	**.54**	.00	**.61**	-.01	**.65**	**.61**	-.07	-.07	**.76**	-.02	**.67**	-.13	**.50**	-.01	**-.58**	**.49**	.01	**.46**	.01	.07
ASPD	.06	**.63**	.07	**.62**	.02	-.09	**.77**	-.10	**.78**	.05	.15	**.53**	.03	.22	**.52**	.00	**.88**	-.10	**.90**	.16	-.02	**.80**	.01	**.79**	.00
Alc	-.12	**.84**	-.06	**.82**	-.03	-.12	**.84**	-.12	**.84**	.01	-.10	**.86**	.00	.02	**.81**	.02	**.79**	.00	**.74**	.25	-.05	**.78**	-.01	**.79**	-.08
Nicot	.00	**.68**	.01	**.66**	.03	.01	**.70**	-.01	**.71**	-.05	.06	**.65**	.01	.16	**.62**	-.05	**.80**	-.33	**1.03**	-.02	.01	**.68**	.04	**.68**	-.06
Marij	.00	**.82**	.00	**.80**	.05	.03	**.82**	.02	**.83**	-.10	.00	**.81**	-.03	.14	**.78**	.12	**.96**	.09	**.89**	.29	.10	**.70**	.17	**.65**	.03
Drug	.01	**.86**	.50	**.84**	-.01	.05	**.80**	.01	**.82**	-.20	.16	**.68**	.30	-.03	**.64**	-.09	**.94**	.00	**.72**	**.74**	.01	**.97**	-.01	**.89**	.27
CFI	.982		.999			.983		1.000			.983		1.000			.989		1.000			1.000		1.000		
TLI	.971		.997			.973		1.000			.973		1.000			.982		1.003			1.003		1.005		
RMSEA	.017		.005			.006		.001			.006		.001			.002		.000			.000		.000		

Note: Geomin rotated factor loadings greater than or equal to .4 are bolded. CFI: comparative fit index. TLI: Tucker-Lewis index. RMSEA: root mean squared error of approximation. MDD: major depressive disorder. Dysth: dysthymic disorder. GAD: generalized anxiety disorder. Panic: panic disorder. Social: social phobia. Spec: specific phobia. ASPD: antisocial PD. Nic: nicotine dependence. Alc: alcohol dependence. Marij: marijuana dependence. Drug: other drug dependence. See text for full description of the ethnic groups.

Table 6

Exploratory factor analysis results by ethnicity for 12-month diagnoses

	White					Hispanic					Black					Asian/ Pacific Islander					American Indian/ Alaska Native				
	1	2	1	2	3	1	2	1	2	3	1	2	1	2	3	1	2	1	2	3	1	2	1	2	3
MDD	**.81**	.04	**.81**	.07	.02	**.72**	.13	**.76**	.11	.06	**.86**	.00	**.77**	.12	.04	**.75**	.00	-.11	**.79**	.16	**.97**	.00	**1.05**	-.17	-.10
Dysth	**.87**	-.07	**.97**	.00	-.16	**.78**	.11	**.87**	.08	-.02	**.91**	-.08	**.92**	.01	-.05	**.66**	.07	.04	**.67**	.00	**.80**	.02	**.91**	-.28	.02
GAD	**.82**	-.01	**.65**	.00	.28	**.85**	.00	**.77**	.01	.16	**.84**	.01	**.58**	**.40**	-.01	**.95**	-.05	-.01	**.83**	.27	**.90**	-.36	**.86**	.02	**-.48**
Panic	**.64**	.14	**.40**	.13	.38	**.75**	-.10	**.54**	-.06	.32	**.49**	.27	.26	**.40**	.23	**.56**	.18	**.53**	-.01	**.76**	**.65**	.07	**.73**	-.14	.02
Social	**.65**	.01	.30	-.02	**.54**	**.76**	-.05	**.42**	-.01	**.57**	**.51**	.18	.01	**.69**	.12	**.58**	.08	.31	.27	**.48**	.33	.23	.33	.24	-.03
Spec	**.48**	.11	.00	.09	**.73**	**.53**	.09	.18	.15	**.64**	**.45**	.07	-.12	**.75**	-.01	**.60**	-.06	.17	.22	**.49**	.30	.17	.27	.33	.06
ASPD	.08	**.60**	.02	**.59**	.14	-.02	**.71**	-.03	**.66**	**.42**	.10	**.53**	.05	.26	**.47**	-.03	**.82**	**.72**	.03	.02	.29	**.43**	.34	.09	**.40**
Alc	-.11	**.80**	-.09	**.77**	.04	-.24	**.84**	.01	**.75**	.07	.00	**.73**	.21	-.01	**.68**	.34	**.65**	**.56**	.36	-.03	-.05	**.99**	.00	**.53**	**.74**
Nicot	.06	**.63**	-.02	**.63**	.16	.08	**.59**	.20	**.53**	.17	.03	**.61**	.06	.17	**.54**	-.01	**.85**	**.97**	-.10	-.01	.26	**.45**	.31	.04	**.52**
Marij	-.01	**.81**	.03	**.81**	-.02	.01	**.96**	.00	**.88**	**.54**	-.11	**.96**	-.04	.24	**.86**	.19	**.93**	**1.14**	.01	.17	.17	**.80**	.01	**1.78**	.00
Drug	.09	**.78**	.15	**.77**	-.02	.02	**.88**	.35	**.80**	-.02	**.41**	**.61**	**.59**	-.01	**.56**	**.75**	**.46**	.00	**1.26**	**-.93**	**.45**	**.62**	**.56**	-.01	**.70**
CFI	.979		.999			.971		.993			.981		.999			.997		1.000			.982		1.000		
TLI	.965		.998			.953		.984			.97		.998			.995		.999			.971		1.007		
RMSEA	.012		.003			.006		.003			.005		.001			.001		.000			.002		.000		

Note: Geomin rotated factor loadings greater than or equal to .4 are bolded. CFI: comparative fit index. TLI: Tucker-Lewis index. RMSEA: root mean squared error of approximation. MDD: major depressive disorder. Dysth: dysthymic disorder. GAD: generalized anxiety disorder. Panic: panic disorder. Social: social phobia. Spec: specific phobia. ASPD: antisocial PD. Nic: nicotine dependence. Alc: alcohol dependence. Marij: marijuana dependence. Drug: other drug dependence. See text for full description of the ethnic groups.

Table 7

Model fit statistics and factor means

	CFI	TLI	RMSEA	# Free
White (*n* = 24,507)				
Lifetime diagnoses	.983	.978	.014	--
12-month diagnoses	.978	.971	.011	--
Hispanic (*n* = 8,308)				
Lifetime diagnoses	.987	.983	.005	--
12-month diagnoses	.973	.966	.005	--
Black (*n* = 8,245)				
Lifetime diagnoses	.980	.974	.007	--
12-month diagnoses	.976	.969	.005	--
Asian/Pacific Islander (*n* = 1,332)				
Lifetime diagnoses	.988	.985	.002	--
12-month diagnoses	.989	.986	.001	--
American Indian/Alaska Native (*n* = 701)				
Lifetime diagnoses	1.000	1.003	.000	--
12-month diagnoses	.962	.952	.003	--
Multi-group (All ethnicities)				
Lifetime diagnoses				
Unconstrained model	.984	.979	.016	115
Constrained model	.987	.985	.013	75
12-month diagnoses				
Unconstrained model	.974	.967	.013	115
Constrained model	.978	.976	.011	75

Spectrum Means

	Lifetime Diagnoses		12-month Diagnoses	
	INT	EXT	INT	EXT
Hispanic	$-.280^{A,N,W}$	$-.327^{N,W}$	$-.151^{N}$	$-.209^{N}$
Black	$-.241^{A,N,W}$	$-.346^{N,W}$	$-.065^{N}$	$-.126^{N}$
Asian/Pacific Islander	$-.514^{B,H,N,W}$	$-.351^{N,W}$	$-.386^{N,W}$	$-.081^{N}$
American Indian/Alaska Native	$.307^{A,B,H,W}$	$.504^{A,B,H,W}$	$.217^{A,B,H,W}$	$.595^{A,B,H,W}$

Note: Multigroup modeled ethnicities simultaneously. The unconstrained model allowed each ethnicity to have unique factor loadings and thresholds; the constrained (invariant) model constrained factor loadings and thresholds to equality across ethnicities. CFI: comparative fit index. TLI: Tucker-Lewis index. RMSEA: root mean squared error of approximation. # Free: number of parameters freely estimated in the model. INT: Internalizing. EXT: Externalizing. Spectrum means are given in standard deviation units. Superscript letters indicate a constrained model spectrum mean differs ($p < .05$) from that of ^{A}Asian/Pacific Islander, ^{B}Black, ^{H}Hispanic, ^{N}American Indian/Alaska Native, or ^{W}White individuals (White means were fixed to 0.00 to provide a reference group). See text for full description of the ethnic groups.

Figure 1

A theoretical example of the common factor model

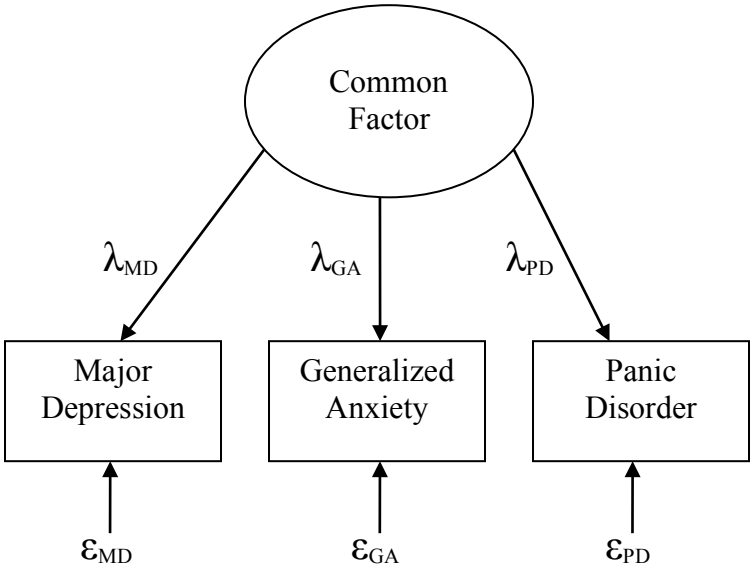
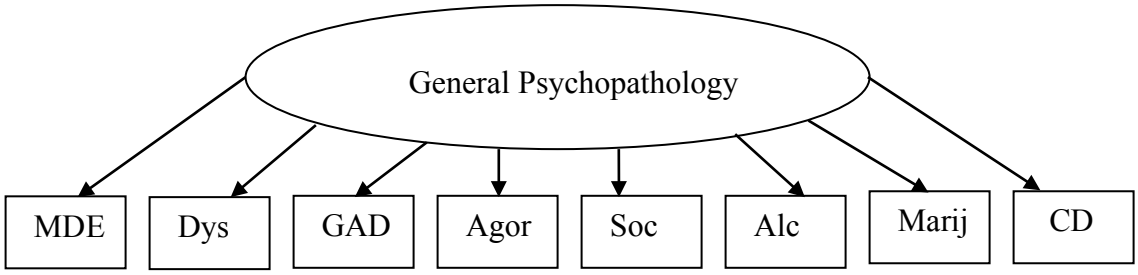


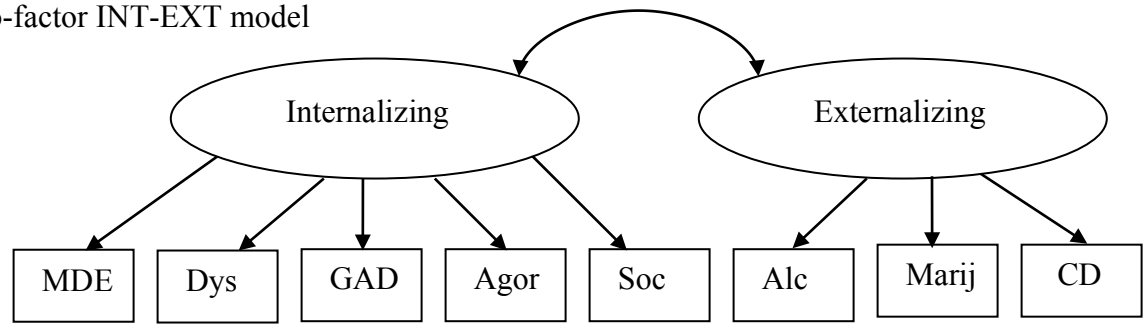
Figure 2

Simplified representations of commonly modeled structures

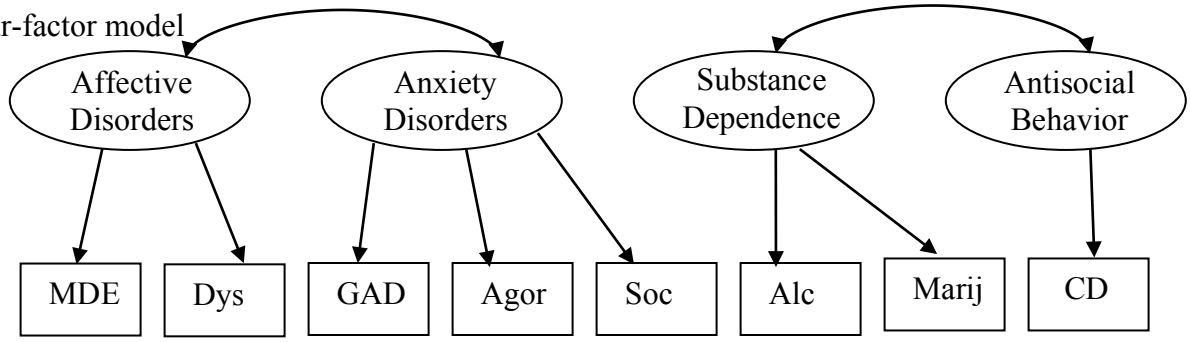
One-factor "general psychopathology" model



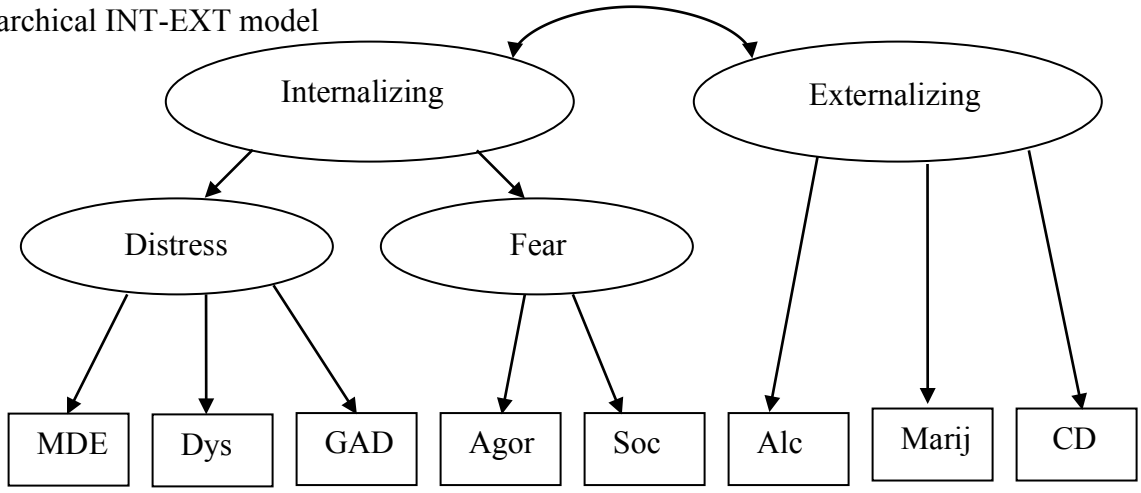
Two-factor INT-EXT model



Four-factor model



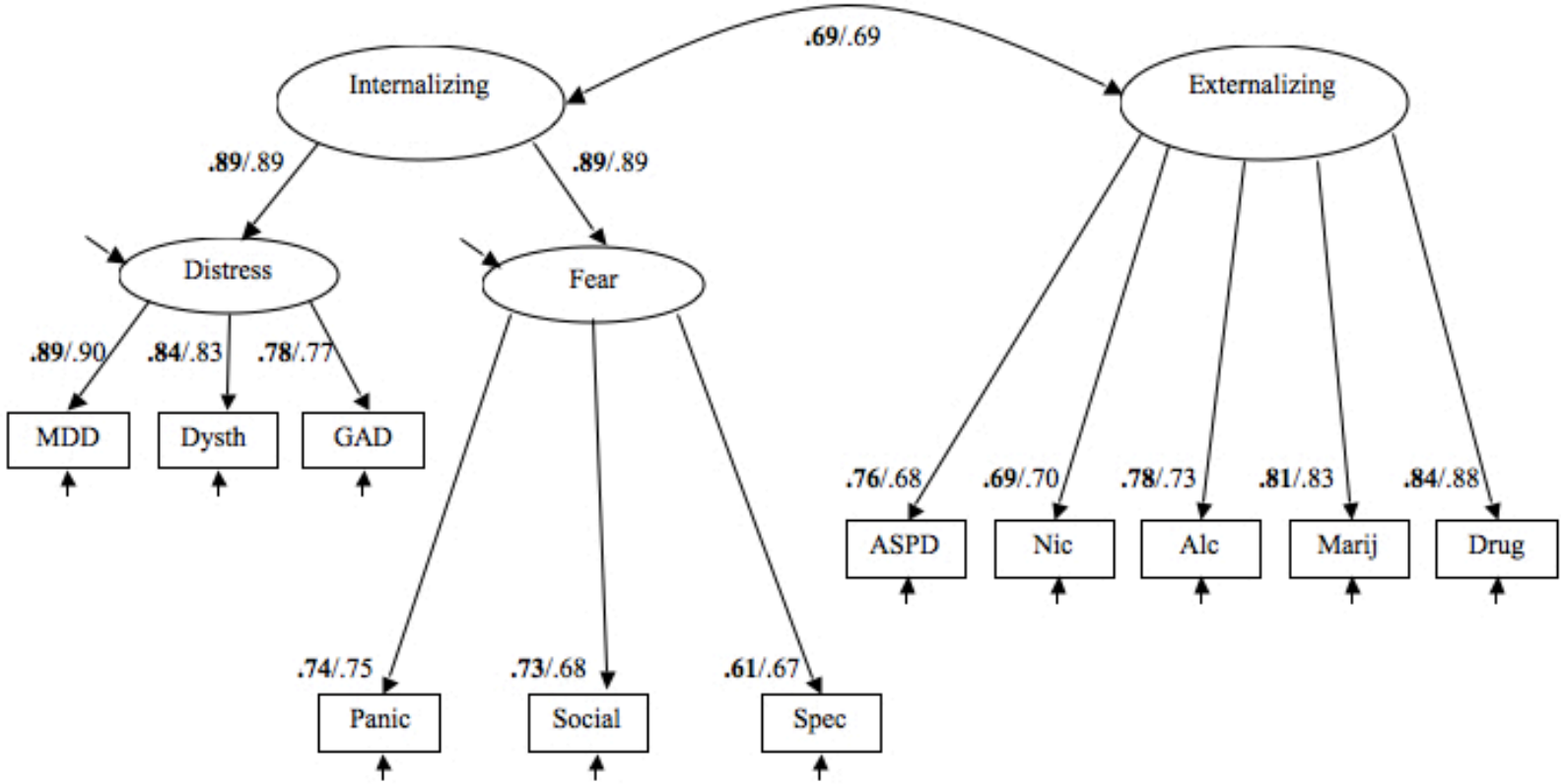
Hierarchical INT-EXT model



Circles represent latent variables, rectangles represent observed variables, straight arrows represent factor loadings, and curved arrows represent factor correlations. Specific variances have been omitted for simplicity. MDE: major depressive episode; Dys: dysthymia; GAD: generalized anxiety disorder; Agor: agoraphobia; Soc: social phobia; Alc: alcohol dependence; Marij: marijuana dependence; CD: conduct disorder.

Figure 3

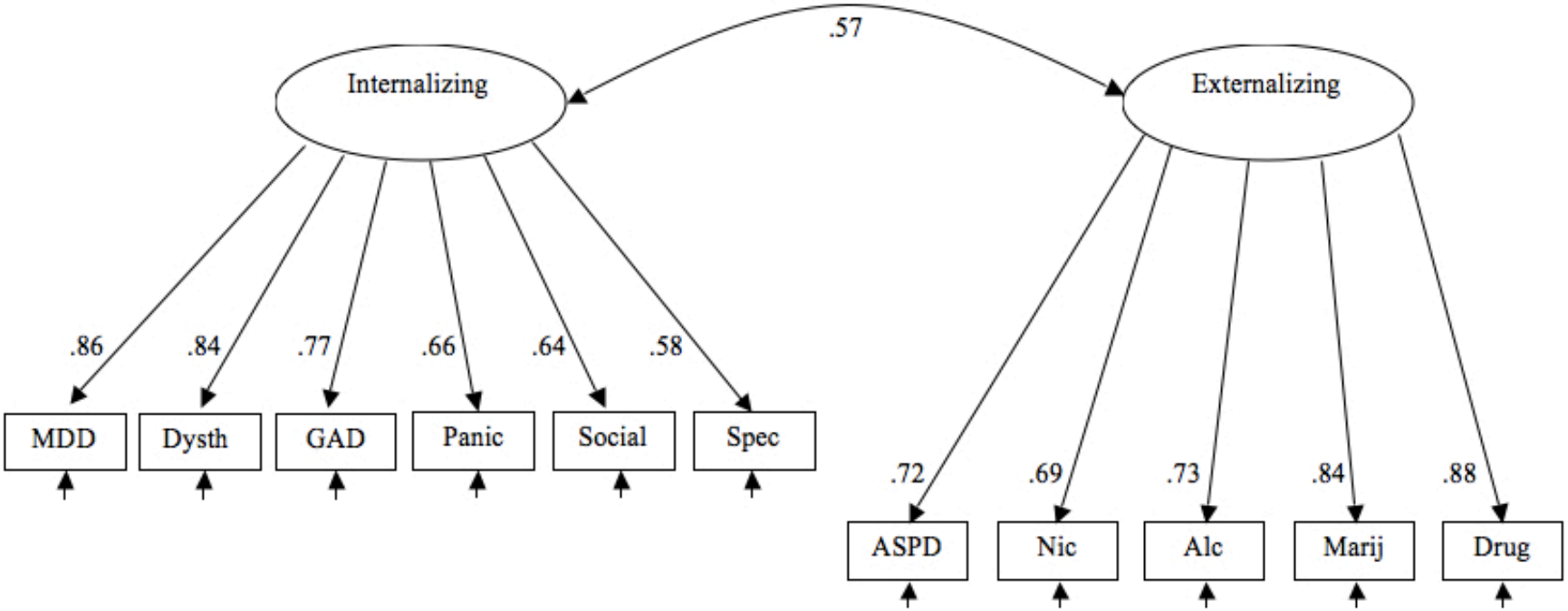
The constrained (gender invariant) model in women and men using lifetime diagnoses



Note. Values are standardized factor loadings (all significant $p < .001$). Values before slash and bolded are for women; values after slash are for men. Values differ slightly across gender due to standardization. MDD: major depressive disorder. Dysth: dysthymic disorder. GAD: generalized anxiety disorder. Panic: panic disorder. Social: social phobia. Spec: specific phobia. ASPD: antisocial PD. Nic: nicotine dependence. Alc: alcohol dependence. Marij: marijuana dependence. Drug: other drug dependence. Arrows without numbers indicate unique variances, including error.

Figure 4

The constrained (ethnicity invariant) model in White, Hispanic, Black, Asian/Pacific Islander, and American Indian/Alaska Native individuals using lifetime diagnoses



Note. Values are standardized and all significant ($p < .001$). For clarity, only estimates for White individuals' lifetime diagnoses are presented; values for Hispanic, Black, Asian/Pacific Islander, and American Indian/Alaska Native individuals differed only slightly due to standardization. MDD: major depressive disorder. Dysth: dysthymic disorder. GAD: generalized anxiety disorder. Panic: panic disorder. Social: social phobia. Spec: specific phobia. ASPD: antisocial PD. Nic: nicotine dependence. Alc: alcohol dependence. Marij: marijuana dependence. Drug: other drug dependence. Arrows without numbers indicate unique variances, including error. See text for full description of the ethnic groups.

References

Achenbach, T. M., & Edelbrock, C. S. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin, 85*(6), 1275-1301.

Achenbach, T. M., & Edelbrock, C. S. (1984). Psychopathology of childhood. *Annual Review of Psychology, 35*, 227-256.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

Andrews, G., Goldberg, D. P., Krueger, R. F., Carpenter, W. T., Jr., Hyman, S. E., Sachdev, P., & Pine, D. S. (2009). Exploring the feasibility of a meta-structure for DSM-V and ICD-11: Could it improve utility and validity? *Psychological Medicine, 39*, 1993-2000.

Barlow, D. H., Allen, L. B., & Choate, M. L. (2004). Toward a unified treatment for emotional disorders. *Behavior Therapy, 35*, 205-230.

Barlow, D. H., Farchione, T. J., Fairholme, C. P., Ellard, K. K., Boisseau, C. L., Allen, L. B., & Ehrenreich-May, J. (2011). *Unified protocol for transdiagnostic treatment of emotional disorders: Therapist guide.* New York: Oxford University Press.

Barrantes-Vidal, N., Ros-Morente, A., & Kwapil, T. R. (2009). An examination of neuroticism as a moderating factor in the association of positive and negative schizotypy with psychopathology in a nonclinical sample. *Schizophrenia Research, 115*, 303-309.

Beesdo-Baum, K., Höfler, M., Gloster, A. T., Klotsche, J., Lieb, R., Beauducel, A., Bühner, M., Kessler, R. C., & Wittchen, H.-U. (2009). The structure of common mental disorders: A replication study in a community sample of adolescents and young adults. *International Journal of Methods in Psychiatric Research, 18*(4), 204-220.

Blazer, D. G., Kessler, R. C., McGonagle, K. A., & Swartz, M. S. (1994). The prevalence and distribution of major depression in a national community sample: The National Comorbidity Survey. *American Journal of Psychiatry, 151*, 979-986.

Brim, O. G., Ryff, C. D., & Kessler, R. C. (Eds.). (2004). *How healthy are we? A national study of well-being at midlife*. Chicago: University of Chicago Press.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Canino, G., Bravo, M., Ramirez, R., Febo, V. E., Rubio-Stipec, M., Fernandez, R. L., & Hasin, D. (1999). The Spanish Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): Reliability and concordance with clinical diagnoses in a Hispanic population. *Journal of Studies on Alcohol, 60*, 790-799.

Chatterji, S., Saunders, J. B., Vrasti, R., Grant, B. F., Hasin, D., & Mager, D. (1997). Reliability of the alcohol and drug modules of the Alcohol Use Disorder and Associated Disabilities Interview Schedule—Alcohol/Drug-Revised (AUDADIS-ADR): An international comparison. *Drug and Alcohol Dependence, 47*, 171-185.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233-255.

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

Dawson, D. A., Goldstein, R. B., Moss, H. B., Li, T.-K., & Grant, B. F. (2010). Gender differences in the relationship of internalizing and externalizing psychopathology to alcohol dependence: Likelihood, expression and course. *Drug and Alcohol Dependence, 112*, 9-17.

Donnellan, M. B., & Lucas, R. E. (2009). Age differences in the Big Five across the life span: Evidence from two national samples. *Psychology and Aging, 23*, 558-566.

Eaton, N. R. (2012). Mental disorders in sexual minorities: The interplay between liability, discrimination, and harassment. Manuscript in preparation.

Eaton, N. R., Krueger, R. F., Keyes, K. M., Skodol, A. E., Markon, K. E., Grant, B. F., & Hasin, D. S. (2011). Borderline personality disorder co-morbidity: Relationship to the internalizing-externalizing structure of common mental disorders. *Psychological Medicine, 41*(5), 1041-1050.

Eaton, N. R., Krueger, R. F., & Oltmanns, T. F. (2011). Aging and the structure and long-term stability of the internalizing spectrum of personality and psychopathology. *Psychology and Aging, 26*(4), 987-993.

Eaton, N. R., South, S. C., & Krueger, R. F. (2010). The meaning of comorbidity among common mental disorders. In T. Millon, R., Krueger, and E. Simonsen (Eds.), *Contemporary directions in psychopathology: Scientific foundations of*

the DSM-V and ICD-11 (2nd ed., pp. 223-241). New York: Guilford Publications.

Feinstein, A. R. (1970). The pre-therapeutic classification of co-morbidity in chronic disease. *Journal of Chronic Disease, 23*, 455-468.

Goldberg, D., Simms, L. J., Gater, R., & Krueger, R. F. (2011). Integration of dimensional spectra for depression and anxiety into categorical diagnoses for general medical practice. In D. A. Regier, W. E. Narrow, E. A. Kuhl, & D. J. Kupfer (Eds.) *The conceptual evolution of DSM-5* (pp. 19-35). Arlington, VA: American Psychiatric Publishing, Inc.

Grant, B. F., & Dawson, D. A. (2006). Introduction to the National Epidemiologic Survey on Alcohol and Related Conditions. *Alcohol Research & Health, 29*(2), 74-78.

Grant, B. F., Dawson, D. A., Stinson, F. S., Chou, S. P., Dufour, M. C., & Pickering, R. P. (2004). The 12-month prevalence and trends in DSM-IV alcohol abuse and dependence: United States, 1991-1992 and 2001-2002. *Drug and Alcohol Dependence, 74*, 223-234.

Grant, B. F., Dawson, D. A., Stinson, F. S., Chou, S. P., Kay, W., & Pickering, R. (2003). The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): Reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug and Alcohol Dependence, 71*, 7-16.

Grant, B. F., Harford, T. C., Dawson, D. A., Chou, P. S., & Pickering, R. P. (1995). The Alcohol Use Disorder and Associated Disabilities Interview Schedule

(AUDADIS): Reliability of alcohol and drug modules in a general population sample. *Drug and Alcohol Dependence, 39*, 37-44.

Grant, B. F., & Weissman, M. M. (2007). Gender and the prevalence of psychiatric disorders. In W. E. Narrow, M. B. First, P. J. Sirovatka, & D. A. Regier (Eds.) *Age and gender considerations in psychiatric diagnosis: A research agenda for DSM-V* (pp. 31-46). Washington, DC: American Psychiatric Association.

Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care, 44*(11, Supplement 3): S78-S94.

Griffith, J. W., Zinbarg, R. E., Craske, M. G., Mineka, S., Rose, R. D., Waters, A. M., & Sutton, J. M. (2010). Neuroticism as a common dimension in the internalizing disorders. *Psychological Medicine, 40*, 1125-1136.

Harkness, K. L., Alavi, N., Monroe, S. M., Slavich, G. M., Gotlib, I. H., & Bagby, R. M. (2010). Gender differences in life events prior to onset of major depressive disorder: The moderating effect of age. *Journal of Abnormal Psychology, 119*(4), 791-803.

Hasin, D., Carpenter, K. M., McCloud, S., Smith, M., & Grant, B. F. (1997). The alcohol use disorder and associated disabilities interview schedule (AUDADIS): Reliability of alcohol and drug modules in a clinical sample. *Drug and Alcohol Dependence, 44*, 133-141.

Hasin, D. S., Goodwin, R. D., Stinson, F. S., & Grant, B. F. (2005). Epidemiology of major depressive disorder: Results from the National Epidemiologic Survey on

Alcoholism and Related Conditions. *Archives of General Psychiatry, 62*, 1097-1106.

Helzer, J. E., Kraemer, H. C., Krueger, R. F., Wittchen, H., Sirovatka, P. J., & Regier, D. A. (2008). *Dimensional approaches in diagnostic classification: Refining the research agenda for* DSM-V. Arlington, VA: American Psychiatric Publishing, Inc.

Hettema, J. M., Neale, M. C., Myers, J. M. ,Prescott, C. A., & Kendler, K. S. (2006). A population-based twin study of the relationship between neuroticism and internalizing disorders. *American Journal of Psychiatry, 163*, 857-864.

Hicks, B. M., Blonigen, D. M., Kramer, M. D., Krueger, R. F., Patrick, C. J., Iacono, W. G., & McGue, M. (2007). Gender differences and developmental change in externalizing disorders from late adolescence to early adulthood: A longitudinal twin study. *Journal of Abnormal Psychology, 116*, 433-447.

Horn, J. L., & McArdle, J. J. (1992). A pratical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*, 117-144.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Huang, B., Grant, B. F., Dawson, D. A., Stinson, F. S., Chou, S. P., Saha, T. D., et al. (2006). Race-ethnicity and the prevalence and co-occurrence of *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*, alcohol and drug use disorders and Axis I and II disorders: United States, 2001 to 2002. *Comprehensive Psychiatry, 47*, 252-257.

James, L. M., & Taylor, J. (2008). Revisiting the structure of mental disorders: Borderline personality disorder and the internalizing/externalizing spectra. *British Journal of Clinical Psychology, 47*, 361-380.

Kendler, K. S. (1993). Twin studies of psychiatric illness: Current status and future directions. *Archives of General Psychiatry, 50*, 905-915.

Kendler, K. S., Aggen, S. H., Knudsen, G. P., Røysamb, E., Neale, M. C., Reichborn-Kjennerud, T. (2011). The structure of genetic and environmental risk factors for syndromal and subsyndromal common DSM-IV axis I and all axis II disorders. *American Journal of Psychiatry, 168*, 29-39.

Kendler, K. S., Prescott, C. A., Myers, J., & Neale, M. C. (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry, 60*, 929-937.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month *DSM-IV* disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*, 617-627.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., et al. (2003). The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association, 289*, 3095-3105.

Kessler, R. C., McGonagle, K. A., Swartz, M., Blazer, D. G., & Nelson, C. B. (1993). Sex and depression in the National Comorbidity Survey I: Lifetime prevalence, chronicity and recurrence. *Journal of Affective Disorders, 29*, 85-96.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H.-U., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of *DSM-III-R* psychiatric disorders in the United States: Results from the National Comorbidity Study. *Archives of General Psychiatry, 51*, 8-19.

Kessler, R. C., Ormel, J., Petukhova, M., McLaughlin, K. A., Green, J. G., Russo, L. J., et al. (2011). Development of lifetime comorbidity in the World Health Organization World Mental Health Surveys. *Archives of General Psychiatry, 68*, 90-100.

Keyes, K. M., Eaton, N. R., Krueger, R. F., McLaughlin, K. A., Wall, M. M., Grant, B. F., & Hasin, D. S. (2012). Child maltreatment and the structure of common psychiatric disorders. *British Journal of Psychiatry, 200*(2), 107-115.

Keyes, K. M., Grant, B. F., & Hasin, D. S. (2008). Evidence for a closing gender gap in alcohol use, abuse, and dependence in the United States population. *Drug and Alcohol Dependence, 93*, 21-29.

Khan, A. A., Jacobson, K. C., Gardner, C. O., Prescott, C. A., & Kendler, K. S. (2005). Personality and comorbidity of common psychiatric disorders. *British Journal of Psychiatry, 186*, 190-196.

Klein, D. N. & Riso, L. P. (1993). Psychiatric disorders: Problems of boundaries and comorbidity. In C. G. Costello (Ed.), *Basic issues in psychopathology* (pp. 19-26). New York: Guilford.

Klose, M., & Jacobi, F. (2004). Can gender differences in the prevalence of mental disorders be explained by sociodemographic factors? *Archives of Women's Mental Health, 7*, 133-148.

Kovot, R., Ruggero, C. J., Krueger, R. F., Watson, D., Yuan, O., & Zimmerman, M. (2011a). New dimensions in the quantitative classification of mental illness. *Archives of General Psychiatry, 68*(10), 1003-1011.

Kotov, R., Chang, S.-W., Fochtmann, L. J., Mojtabai, R., Carlson, G. A., Sedler, M. J., & Bromet, E. J. (2011b). Schizophrenia in the internalizing-externalizing framework: A third dimension? *Schizophrenia Bulletin, 37*(6), 1168-1178.

Kramer, M. D., Krueger, R. F., & Hicks, B. M. (2008). Internalizing and externalizing liability factors account for gender differences in the prevalence of common psychopathological syndromes. *Psychological Medicine, 38*, 51-61.

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

Krueger, R. F. (2005). Continuity of Axes I and II: Toward a unified model of personality, personality disorders, and clinical disorders. *Journal of Personality Disorders, 19*, 233-261.

Krueger, R. F., Caspi, A., Moffitt, T. E., & Silva, P. A. (1998). The structure and stability of common mental disorders (*DSM-III-R*): A longitudinal-epidemiological study. *Journal of Abnormal Psychology, 107*(2), 216-227.

Krueger, R. F., Chentsova-Dutton, Y. E., Markon, K. E., Goldberg, D., & Ormel, J. (2003). A cross-cultural study of the structure of comorbidity among common psychopathological syndromes in the general health care setting. *Journal of Abnormal Psychology, 112*(3), 437-447.

Krueger, R. F., Hicks, B. M., Patrick, C. J., Carlson, S. R., Iacono, W. G., & McGue, M. (2002) Etiologic connections among substance dependence, antisocial

behavior, and personality: Modeling the externalizing spectrum. *Journal of Abnormal Psychology, 3*, 411-424.

Krueger, R. F., & Markon, K. E. (2006). Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology, 2*, 111-133.

Krueger, R. F., Markon, K. E., Patrick, C. J., Benning, S. D., & Kramer, M. D. (2007). Linking antisocial behavior, substance use, and personality: An integrative quantitative model of the adult externalizing spectrum. *Journal of Abnormal Psychology, 116*(4), 645-666.

Krueger, R. F., & Piasecki, T. M. (2002). Toward a dimensional and psychometrically informed approach to conceptualizing psychopathology. *Behaviour Research and Therapy, 40*, 485-500.

Lahey, B. B., Rathouz, P. J., Van Hulle, C., Urbano, R. C., Krueger, R. F., Applegate, B., Garriock, H. A., Chapman, D. A., & Waldman, I. D. (2008). Testing structural models of *DSM-IV* symptoms of common forms of child and adolescent psychopathology. *Journal of Abnormal Child Psychology, 36*, 187-206.

Lahey, B. B., Van Hulle, C. A., Singh, A. L., Waldman, I. D., & Rathouz, P. J. (2011). Higher-order genetic and environmental structure of prevalent forms of child and adolescent psychopathology. *Archives of General Psychiatry, 68*, 181-189.

Lam, D., Smith, N., Checkley, S., Rijsdijk, F., & Sham, P. (2003). Effect of neuroticism, response style and information processing on depression severity in a clinically depressed sample. *Psychological Medicine, 33*, 469-479.

Lilienfeld, S. O., Waldman, I. D., & Israel, A. C. (1994). A critical examination of the use of the term and concept of *comorbidity* in psychopathology research. *Clinical Psychology: Science and Practice, 1*, 71-83.

Markon, K. E. (2010). Modeling psychopathology structure: A symptom-level analysis of Axis I and II disorders. *Psychological Medicine, 40*, 273-288.

Markon, K. E., & Krueger, R. F. (2005). Categorical and continuous models of liability to externalizing disorders: A direct comparison in NESARC. *Archives of General Psychiatry, 62*, 1352-1359.

Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika, 58*(4), 525-543.

Miller, J. D., & Lynam, D. (2001). Structural models of personality and their relation to antisocial behavior: A meta-analytic review. *Criminology, 39*, 765-798.

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*(3), 479-515.

Mineka, S. Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49*, 377-412.

Moffitt, T. E., Caspi, A., Taylor, A., Kokaua, J., Milne, B. J., Polanczyk, G., & Poulton, R. (2010). How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective *versus* retrospective ascertainment. *Psychological Medicine, 40*, 899-909.

Muthén, L. K., & Muthén, B. O. (2011). *Mplus user's guide* (6th ed.). Los Angeles: Author.

Nathan, P. E., & Gorman, J. M. (2007). *A guide to treatments that work* (3rd ed.). New York: Oxford University Press, Inc.

Neale, M. C., & Kendler, K. S. (1995). Models of comorbidity for multifactorial disorders. *American Journal of Human Genetics, 57*, 935-953.

Nolen-Hoeksema, S. (1987). Sex differences in unipolar depression: Evidence and theory. *Psychological Bulletin, 2*, 259-282.

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science, 3*(5), 400-424.

Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression: Critical review. *British Journal of Psychiatry, 177*, 486-492.

Regier, D. A., Narrow, W. E., Kuhl, E. A., & Kupfer, D. J. (Eds.). (2011). *The conceptual evolution of DSM-5*. Arlington, VA: American Psychiatric Publishing, Inc.

Røysamb, E., Kendler, K. S., Tambs, K., Ørstavik, R. E., Neale, M. C., Aggen, S. H., Torgersen, S., & Reichborn-Kjennerud, T. (2011). The joint structure of *DSM-IV* axis I and axis II disorders. *Journal of Abnormal Psychology, 120*, 198-209.

Ruan, W. J., Goldstein, R. B., Chou, S. P., Smith, S. M., Saha, T. D., Pickering, R. P., Dawson, D. A., Huang, B., Stinson, F. S., & Grant, B. F. (2008). The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): Reliability of new psychiatric diagnostic modules and risk factors in a general population sample. *Drug and Alcohol Dependence, 92*, 27-36.

Seeley, J. R., Kosty, D. B., Farmer, R. F., & Lewinsohn, P. M. (2011). The modeling of internalizing disorders on the basis of patterns of lifetime comorbidity:

Associations with psychosocial functioning and psychiatric disorders among first-degree relatives. *Journal of Abnormal Psychology, 120*(2), 308-321.

Shear, K., Halmi, K.A., Widiger, T.A., & Boyce, C. (2007). Sociocultural factors and gender. In W. E. Narrow, M. B. First, P. J. Sirovatka, & D. A. Regier (Eds.) *Age and gender considerations in psychiatric diagnosis: A research agenda for DSM-V* (pp. 65-79). Washington, DC: American Psychiatric Association.

Slade, T., & Watson, D. (2006). The structure of common *DSM-IV* and *ICD-10* mental disorders in the Australian general population. *Psychological Medicine, 36*, 1593-1600.

Slutske, W. S., Heath, A. C., Madden, P. A., Bucholz, K. K., Statham, D. J., & Martin, N. G. (2002). Personality and the genetic risk for alcohol dependence. *Journal of Abnormal Psychology, 111*, 124-133.

Smith, S. M., Stinson, F. S., Dawson, D. A., Goldstein, R., Huang, B., & Grant, B. F. (2006). Race/ethnic differences in the prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Psychological Medicine, 36*, 987-998.

Tang, T. Z., DeRubeis, R. J., Hollon, S. D., Amsterdam, J., Shelton, R., & Schalet, B. (2009). Personality change during depression treatment: A placebo-controlled trial. *Archives of General Psychiatry, 66*, 1322-1330.

Teachman, B. A., Siedlecki, K. L., & Magee, J. C. (2007). Aging and symptoms of anxiety and depression: Structural invariance of the tripartite model. *Psychology and Aging, 22*(1), 160-170.

Thurstone, L. L. (1947). *Multiple-factor analysis.* Chicago: University of Chicago Press.

Trull, T. J., & Durrett, C. A. (2005). Categorical and dimensional models of personality disorder. *Annual Review of Clinical Psychology, 1*, 355-380.

Trull, T. J., Jahng, S., Tomko, R. L., Wood, P. K., & Sher, K. J. (2010). Revised NESARC personality disorder diagnoses: Gender, prevalence, and comorbidity with substance dependence disorders. *Journal of Personality Disorders, 24*(4), 412-426.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4-70.

Vesga-López, O., Schneier, F. R., Wang, S., Heimberg, R. G., Liu, S. M., Hasin, D. S., & Blaco, C. (2008). Gender differences in generalized anxiety disorder: Results from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *Journal of Clinical Psychiatry, 69*, 1606-1616.

Vollebergh, W. A. M., Iedema, J., Bijl, R. V., de Graaf, R., Smit, F., & Ormel, J. (2001). The structure and stability of common mental disorders: The NEMESIS study. *Archives of General Psychiatry, 58*, 597-603.

Watson, D. (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for *DSM-V*. *Journal of Abnormal Psychology, 114*(4), 522-536.

Widiger, T.A. (2007). DSM's approach to gender: history and controversies. In W. E. Narrow, M. B. First, P. J. Sirovatka, & D. A. Regier (Eds.) *Age and gender*

considerations in psychiatric diagnosis: A research agenda for DSM-V (pp. 31-46). Washington, DC: American Psychiatric Association.

Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the *Diagnostic and Statistical Manual of Mental Disorders—Fifth Edition*. *Journal of Abnormal Psychology, 114*(4), 494-504.

Wittchen, H.-U. (1994). Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research, 28*, 57-84.

Wittchen, H.-U., Beesdo-Baum, K., Gloster, A. T., Höfler, M., Klotsche, J., Lieb, R., Beauducel, A., Bühner, M., & Kessler, R. C. (2009). The structure of mental disorders re-examined: Is it developmentally stable and robust against additions? *International Journal of Methods in Psychiatric Research, 18*(4), 189-203.

Wolf, A. W., Schubert, D. S. P., Patterson, M. B., Grande, T. P., Brocco, K. J., & Pendleton, L. (1988). Associations among major psychiatric diagnoses. *Journal of Consulting and Clinical Psychology, 56*(2), 292-294.

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