Epidemiological studies of categorical mental disorders consistently report that gender differences exist in many disorder prevalence rates and that disorders are often comorbid. Can a dimensional multivariate liability model be developed to clarify how gender impacts diverse, comorbid mental disorders? We pursued this possibility in the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC; N = 43,093). Gender differences in prevalence were systematic such that women showed higher rates of mood and anxiety disorders, and men showed higher rates of antisocial personality and substance use disorders. We next investigated patterns of disorder comorbidity and found that a dimensional internalizing-externalizing liability model fit the data well, where internalizing is characterized by mood and anxiety disorders, and externalizing is characterized by antisocial personality and substance use disorders. This model was gender invariant, indicating that observed gender differences in prevalence rates originate from women and men’s different average standings on latent internalizing and externalizing liability dimensions. As hypothesized, women showed a higher mean level of internalizing, while men showed a higher mean level of externalizing. We discuss implications of these findings for understanding gender differences in psychopathology and for classification and intervention.

Keywords: comorbidity, gender differences, internalizing–externalizing, prevalence rates
largest epidemiological study of psychopathology yet undertaken (Dawson, Goldstein, Moss, Li, & Grant, 2010; Keyes, Grant, & Hasin, 2008; Grant et al., 2004; Grant & Weissman, 2007; Trull, Jahng, Tomko, Wood, & Sher, 2010; Véga-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, & Lyubomirsky, 2008).

These theories of gender differences focus primarily on specific disorders and rarely take comorbidity into account. A compelling account of mental disorder comorbidity focuses on unifying latent dimensional liabilities to experience multiple internalizing (mood and anxiety) or externalizing (antisocial and substance use) disorders (Eaton, South, & Krueger, 2010; Krueger, 1999; Slade & Watson, 2006; Vollebergh et al., 2001). Indeed, this internalizing–externalizing liability model is likely to frame key parts of the metastructure, or overall organization, of the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM–IV; e.g., Andrews et al., 2009; Regier, Narrow, Kuhl, & Kupfer, 2011). The internalizing dimension can be bifurcated into distress and fear subfactors; distress relates to disorders such as major depression, dysthymia, and generalized anxiety, and fear relates to disorders such as panic disorder, social phobia, and specific phobia. The externalizing dimension is associated with disorders such as antisocial personality disorder and alcohol, nicotine, and drug dependence. Further, these factors relate to normal personality: Internalizing correlates with neuroticism/negative affectivity (Griffith et al., 2010), and externalizing correlates with disinhibition (Krueger et al., 2002).

When gender differences in prevalence rates and the internalizing–externalizing liability 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 internalizing disorders, while men show significantly higher rates of externalizing 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 internalizing and externalizing liability dimensions. The utility of a dimensional liability model for public health, epidemiology, psychopathology, and intervention research would be notably enhanced if it could encompass the role of gender in mental disorder prevalence.

A few studies have evaluated the structure of psychopathology separately in women and men (e.g., Krueger, 1999; Kendler, Prescott, Myers, & Neale, 2003). However, we are aware of only two studies that have formally tested whether the latent structure of common mental disorders is gender invariant (Hicks et al., 2007; Kramer et al., 2008). While these studies were generally supportive of a gender invariant model, their findings’ generalizability was limited by nonrepresentative samples. Further, neither study focused on diagnoses from the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM–IV; American Psychiatric Association, 1994). Addressing these limitations is critical, given the potential use of a dimensional liability model to frame key aspects of DSM–5. If the model is to be applied to DSM–5 mental disorders—and to both women and men—factorial invariance (a lack of bias) across gender should be demonstrated in a maximally representative sample.

The current study examined a nationally representative sample of 43,093 individuals assessed in the first wave of the NESARC. Our initial goal was to examine and present any gender differences in prevalence rates of common mental disorders in this sample. We were interested not only in the significance of these gender differences but also their directionality: We hypothesized that women would show significantly higher rates of internalizing disorders than men, and men would show significantly higher rates of externalizing disorders. Second, we sought to determine the latent comorbidity structure of these disorders in women and men, separately. Third, we aimed to formally test whether or not the emergent structures in women and men could be considered gender invariant (i.e., women and men showing equivalent structures of psychopathology). Finally, if invariance were found, we hypothesized that women would have a higher mean standing on the internalizing liability dimension than men, and men would have a higher mean standing on the externalizing liability dimension. The presence of a gender invariant structure of common mental disorders would indicate that these gender differences in latent internalizing and externalizing liabilities account for the observed gender differences in prevalence rates. That is, gender differences in the prevalence of different manifest categorical disorders would be a function of mean-level gender differences in underlying liability dimensions. As such, these underlying dimensions, as opposed to their manifestations as specific observed categories of psychopathology, would be highlighted as important organizing constructs for official nosologies and for research on the role of gender in psychopathology.

Method

Participants

This study utilized data from 43,093 individuals who participated in the first wave of the NESARC, conducted in 2001–2002. The NESARC study’s design has been detailed elsewhere (Grant & Dawson, 2006). The first wave of NESARC was a representative sample of the civilian, noninstitutionalized United States population, aged 18 and older. Young adults, African Americans, and Hispanics were oversampled. Women composed 57% (n = 24,575). Race/ethnicity was selected by participants, using censusedefined categories: White (56.9%), Hispanic or Latino (19.3%), African American (19.1%), Asian/Native Hawaiian/Pacific Islander (3.1%), and American Indian/Alaska Native (1.6%). Participants provided written informed consent after reviewing a complete description of the study.

Assessment

Lifetime and past-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), a structured interview designed for experienced lay interviewers. Major depressive disorder, dysthymic disorder, general-
ized anxiety disorder, panic disorder, social phobia, specific phobia, alcohol dependence, nicotine dependence, marijuana dependence, other drug dependence, and antisocial personality disorder AUDADIS–IV diagnoses were examined. The other drug dependence variable was created to collapse 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 whose variance would be sufficient for covariance structure modeling; the internal consistency of this variable was good ($\alpha = .77$). In keeping with DSM–IV notions of personality disorder stability, antisocial personality disorder was assessed on a lifetime basis only; this lifetime diagnosis was used in both lifetime and 12-month analyses. The reliability of the AUDADIS–IV diagnoses examined have been reported elsewhere and are generally good to excellent (e.g., kappas = .42 to .84; see Hasin et al., 2005). Test–retest estimates for AUDADIS–IV disorders are similar to other structured interviews (e.g., the Diagnostic Interview Schedule [DIS], the Composite International Diagnostic Interview [CIDI]) used in large psychiatric epidemiologic surveys (reviewed in Wittchen, 1994). Further, the AUDADIS–IV has advantages over structured interviews such as the DIS, including assessment of clinically significant distress and impairment after the syndrome is fully characterized (Hasin et al., 2005).

### Statistical Analyses

All analyses were conducted in Mplus version 6 (Muthén & Muthén, 2010) using the Mplus defaults of delta parameterization and WLSMV estimator. WLSMV allowed us to treat diagnostic variables as categorical and use the NESARC’s weighting, clustering, and stratification variables. Odds ratios used men as the reference comparison group. To evaluate model fit in confirmatory factor analyses (CFAs), we considered the comparative fit index (CFI), Tucker-Lewis index (TLI), RMSEA, and the number of freely estimated parameters in the model. CFI/TLI values >.95 and RMSEA values <.06 suggest good model fit (Hu & Bentler, 1999). Based on simulations, Cheung and Rensvold (2002) proposed that a CFI difference critical value of .01 be used to test whether the addition of constraints leads to notably worse model fit in factorial invariance studies. More parsimonious models use fewer freely estimated parameters. In model comparisons, therefore, we defined the optimal model by means of the best fit (CFI, TLI, and RMSEA), model parsimony (number of free parameters), and the CFI critical difference of .01. Distress and fear loadings on higher-order internalizing were constrained to equality to ensure model identification.

### Results

#### Prevalence Rates

Table 1 presents the prevalence rates for the disorders included in the current study, separately for men and women and for lifetime and 12-month disorders. All odds ratios were significant (all $p < .001$ except other drug dependence at $p = .005$). Across lifetime and 12-month prevalence rates, women showed higher rates for all internalizing disorders, and men showed higher rates for all externalizing disorders.

#### Comorbidity Structure

In CFAs (see Table 2), guided by previous studies and exploratory factor analyses (not reported here for brevity), we parameterized each diagnosis to load 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) externalizing: alcohol dependence, nicotine dependence, marijuana dependence, other drug dependence, other drug dependence, and antisocial personality disorder. Distress and fear were parameterized to load on a higher-order internalizing factor, which was allowed to correlate with the externalizing factor. This internalizing–externalizing model provided a very good fit in the total sample for both lifetime and 12-month diagnoses. Within each gender modeled separately, this internalizing–externalizing model continued to fit very well for lifetime and 12-month diagnoses.

#### Invariance

Because an internalizing–externalizing model fit well for both women and men, our next question was how similar these models

---

### Table 1

**Disorder Prevalence Rates and Odds Ratios by Gender**

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Lifetime disorders</th>
<th>12-month disorders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Women (%)</td>
<td>Men (%)</td>
</tr>
<tr>
<td>Depression</td>
<td>22.9</td>
<td>13.1</td>
</tr>
<tr>
<td>Dysthymia</td>
<td>6.2</td>
<td>3.5</td>
</tr>
<tr>
<td>Generalized Anxiety</td>
<td>5.8</td>
<td>3.1</td>
</tr>
<tr>
<td>Panic Disorder</td>
<td>7.2</td>
<td>3.7</td>
</tr>
<tr>
<td>Social Phobia</td>
<td>5.8</td>
<td>4.3</td>
</tr>
<tr>
<td>Specific Phobia</td>
<td>12.4</td>
<td>6.2</td>
</tr>
<tr>
<td>Alcohol Dependence</td>
<td>8.0</td>
<td>17.4</td>
</tr>
<tr>
<td>Nicotine Dependence</td>
<td>15.6</td>
<td>20.0</td>
</tr>
<tr>
<td>Marijuana Dependence</td>
<td>0.9</td>
<td>1.7</td>
</tr>
<tr>
<td>Other Drug Dependence</td>
<td>1.4</td>
<td>2.2</td>
</tr>
<tr>
<td>Antisocial Personality</td>
<td>1.9</td>
<td>5.5</td>
</tr>
</tbody>
</table>

*Note.* All odds ratios (OR) significant at $p < .001$, except *$p = .005$. Men are OR comparison group. 95% confidence intervals are given in parentheses. Antisocial personality disorder was only assessed as a lifetime disorder.
were across gender in terms of model parameters—that is, whether the magnitude of parameters differed by gender or whether they showed invariance. Tests of invariance for indicators such as diagnoses require methodology appropriate for modeling categorical variables (Millsap & Yun-Tein, 2004). In this approach, factor loadings and thresholds are constrained to equality or freed, in tandem, across genders. In our first model (the “unconstrained model”), loadings and thresholds were free across genders, factor means were set to 0 in both genders, and scaling factors were fixed to 1 in both genders. In the second model (the “constrained model”), loadings and thresholds were constrained to equality across genders, factor means were set to 0 in men and were free in women, and scaling factors were fixed to 1 in men and were free in women. This model represented a gender invariant psychopathology structure.

We fit the unconstrained and constrained models in men and women simultaneously via a multiple group CFA, separately for lifetime and 12-month diagnoses (see Table 2). For lifetime diagnoses, the fits of the two models were identical, but the constrained model had fewer freely estimated parameters than the unconstrained model. The constrained model for lifetime diagnoses is depicted in Figure 1. For 12-month diagnoses, the constrained model had a better fit, with greater parsimony, than did the unconstrained model. 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 structure, factor loadings and thresholds for all diagnoses were equivalent for women and men. Thus, the structure of these common mental disorders, including the connections between individual diagnoses and the underlying factors, could be considered gender invariant.

In terms of factor means, means of these latent internalizing and externalizing factors were fixed to 0 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 were across gender in terms of model parameters—that is, whether the magnitude of parameters differed by gender or whether they showed invariance. Tests of invariance for indicators such as diagnoses require methodology appropriate for modeling categorical variables (Millsap & Yun-Tein, 2004). In this approach, factor loadings and thresholds are constrained to equality or freed, in tandem, across genders. In our first model (the “unconstrained model”), loadings and thresholds were free across genders, factor means were set to 0 in both genders, and scaling factors were fixed to 1 in both genders. In the second model (the “constrained model”), loadings and thresholds were constrained to equality across genders, factor means were set to 0 in men and were free in women, and scaling factors were fixed to 1 in men and were free in women. This model represented a gender invariant psychopathology structure.

We fit the unconstrained and constrained models in men and women simultaneously via a multiple group CFA, separately for lifetime and 12-month diagnoses (see Table 2). For lifetime diagnoses, the fits of the two models were identical, but the constrained model had fewer freely estimated parameters than the unconstrained model. The constrained model for lifetime diagnoses is depicted in Figure 1. For 12-month diagnoses, the constrained model had a better fit, with greater parsimony, than did the unconstrained model. 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 structure, factor loadings and thresholds for all diagnoses were equivalent for women and men. Thus, the structure of these common mental disorders, including the connections between individual diagnoses and the underlying factors, could be considered gender invariant.

In terms of factor means, means of these latent internalizing and externalizing factors were fixed to 0 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 internalizing liability than men). Because complete factorial invariance had been established, these results demonstrated that the observed differences in the prevalence rates of the specific disorders modeled between women and men could be accounted for by the genders’ different average levels of latent internalizing and externalizing.

Discussion

The current study sought to synthesize two lines of research—patterns of gender differences in prevalence rates and a potentially gender invariant latent structure of psychopathology—via factorial invariance analyses of liability dimensions underlying DSM–IV disorder comorbidity. We found that the underlying structure of common mental disorders is gender invariant, with significant gender differences in mean liability levels. This provides compelling evidence that observed gender differences in prevalence rates of many common mental disorders originate at the level of latent internalizing and externalizing liabilities.

Limitations

This study is not without its limitations. First, we examined lifetime diagnoses, which can be subject to memory biases. However, our results from lifetime diagnoses were highly congruent with results from 12-month diagnoses, which require much less retrospection. Second, our diagnostic information was collected by extensively trained lay interviewers rather than clinicians. This being said, it is noteworthy that the instrument used to assess symptomatology was fully structured, which resulted in generally good diagnostic reliability levels. Finally, 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 psychiatric disorders may relate to a separate liability factor (e.g., thought disorder liability) while also showing associations with internalizing liability/neuroticism (e.g., Barrantes-Vidal, Ros-Morente, & Kwapił, 2009; Markon, 2010).

Implications

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). Based largely on replications of the internalizing–externalizing model, an organizational meta-structure for many common disorders reflecting this structure has been advocated (Andrews et al., 2009). Our findings support this proposal in two ways. First, we replicated the internalizing–externalizing structure in the NESARC, the largest epidemiologic study of psychopathology yet undertaken. Second, our results indicated that this structure is gender invariant. The current study represents the first time that gender invariance has been tested and successfully incorporated into the internalizing–externalizing liability model of comorbidity among categorical DSM–IV mental disorders in a representative sample. Taken together, these findings support an internalizing–externalizing metastructure for many disorders in DSM–5, especially because the model is applicable in both women and men.

Gender differences research. Our conclusion—that the observed gender differences in prevalence rates systematically reflect gender differences in broad latent liability factors—ties together distinct lines of research and theory on gender differences in prevalence rates for specific disorders. For instance, one major theory to account for gender differences in depression involves the notion that women ruminate more frequently than men, focusing repetitively on their negative emotions and problems rather than engaging in more active problem solving (Nolen-Hoeksema, 1987; Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008). This theory can be readily extended to anxiety (and other internalizing 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 strongly related to \( r = .98 \), and nearly isomorphic with, the latent internalizing dimension reflecting the multivariate comorbidity among DSM–IV mood and anxiety disorders (Griffith et al., 2010). This link between internalizing and trait neuroticism is itself 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 our finding that women had significantly higher mean levels of internalizing than did men. It may be through neuroticism (and disinhibition-related traits in the case of externalizing and men; e.g., Krueger et al., 2002; Miller & Lynam, 2001; Slutske et al., 2002) that psychological processes impact latent propensities to experience comorbid mental disorders. Given that women tend to report higher frequencies of some stressful life events than men do prior to disorder onset (Harkness et al., 2010), the interaction between these liabilities and environmental stressors seems a particularly worthwhile focus for gender differences research.

Intervention and prevention. Our results support recent efforts to develop interventions that target latent disorder liabilities. For instance, both anxious and depressive symptoms often respond well to cognitive–behavioral therapy, and there have been efforts to develop psychotherapeutic interventions that address the shared internalizing liability rather than solely focusing on its manifestations (Barlow et al., 2011). Along these lines, prevention efforts that focus on gender-linked core psychological processes are likely to be effective in impacting multiple disorders. In women, these preventative measures might focus, for instance, on coping and cognitive restructuring skills to reduce the likelihood of rumination and cognitive distortions developing into clinically significant depression or anxiety. In men, prevention might focus on rewarding patterned behaviors and shaping disinhibitory tendencies into outlets that are not destructive to the self or others.

References


Received December 22, 2010
Revision received April 27, 2011
Accepted May 3, 2011