Published in final edited form as:

J Abnorm Psychol. 2011 May ; 120(2): . doi:10.1037/a0022621.

The Modeling of Internalizing Disorders Based on Patterns of Lifetime Comorbidity: Associations with Psychosocial Functioning and Psychiatric Disorders among First-Degree Relatives

John R. Seeley, Derek B. Kosty, Richard F. Farmer, and Peter M. Lewinsohn
Oregon Research Institute

Abstract

Two broad-band latent factors, “internalizing” and “externalizing,” have frequently been identified in studies of the hierarchical structure of psychopathology. In the present research, three competing measurement models of putative internalizing disorders (i.e., a parsimonious single-factor model, the *DSM-IV* model, and an alternative model proposed by Krueger and Watson) were evaluated in terms of their ability to account for lifetime patterns of diagnostic comorbidity. Four diagnostic assessments were performed on an age-based cohort of 816 persons over a 15-year interval. Each of the three measurement models demonstrated adequate or good fit to the data and similar approximating abilities. Additional analyses, however, suggested that non-specific aspects of lifetime mood/anxiety or distress/fear disorders (i.e., general negative affect) largely accounted for indicators of psychosocial functioning at age 30 as well as densities of specific psychiatric disorders among the first-degree relatives of probands. The relevance of these findings for theoretical and descriptive models of internalizing disorders is discussed.

Keywords

Internalizing disorders; mood disorders; anxiety disorders; measurement model; hierarchical organization; psychosocial functioning

Perspectives on the optimal modeling or categorization of psychiatric disorders have been frequently framed within the lumpers/splitter dialectic. Lumpers are more likely to emphasize the significance of similarities among category exemplars over their differences, and propose fewer overall diagnostic categories to account for the most clinically important variations within a population. Splitters, in contrast, emphasize the differences among exemplars, and seek to identify meaningful distinguishing features or characteristics. Proportionately more diagnostic categories are usually proposed to account for these distinctive qualities. Whereas lumpers prefer simplicity, breadth, and coherence, splitters prefer complexity, conciseness, and precision.

Please address editorial correspondence to: John R. Seeley, Ph.D., Oregon Research Institute, 1715 Franklin Blvd., Eugene, Oregon 97403, Tel: 541-484-2123, Fax: 541-484-1108, johns@ori.org.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/pubs/journals/abn

Advances in statistical modeling, such as structural equation modeling, have permitted the evaluation of hierarchical models of psychiatric disorders that can incorporate the perspectives of both lumpers and splitters. In such models, members who display characteristic sets of narrowly-defined features can be simultaneously represented as members of increasingly broader groups at higher levels within the hierarchy. Investigations of comorbidity patterns among psychiatric disorders, for example, frequently suggest that subgroups of disorders demonstrate non-random patterns of co-occurrence. Such comorbidity across hierarchical lines of inclusiveness suggests the possibility of “liability spectra,” whereby the emergence of more narrowly defined disorders within the hierarchy is regarded as expressions of an underlying latent liability factor. In addition to suggesting higher-order factors, comorbidity among subsets of psychiatric disorders might explain, in part or whole, the increased risk for spectrum-related disorders over a lifetime, and potentially have relevance for etiologic theories and therapy effectiveness (Krueger, 1999; Watson, 2005).

When defined by phenotypic features, considerable comorbidity between mood and anxiety disorders or covariation among their associated symptoms is regularly observed (e.g., Watson, 2009). Similarly, genetic influences for depression and anxiety phenotypes substantially overlap, and it is often difficult to distinguish disorders from these two domains at the genetic level of analysis (Gregory & Eley, 2007; Middeldorp, Cath, Van Dyck, & Boomsma, 2005). The substantial comorbidity and covariation among mood and anxiety disorders or their defining features is often explained in terms of a general internalizing factor or continuous dimension (Krueger, Caspi, Moffitt, & Silva, 1998; Krueger & Finger, 2001; McGlinchey & Zimmerman, 2007; South & Krueger, 2008; Watson, 2005; Watson, O'Hara, & Stuart, 2008).

Researchers have investigated whether this general dimension or domain can be meaningfully differentiated into distinct correlated dimensions or subfactors. Some studies in this area suggest the superiority of a parsimonious single-factor representation of the internalizing domain. In Krueger, Chentsova-Dutton, Markon, Goldberg, and Ormel (2003), for example, the structure of seven disorders defined by the World Health Organizations was examined based on samples from 14 countries. For participants from 12 of the countries represented, a two-factor model consisting of a superordinate internalizing factor (defined by depression, somatization, hypochondriasis, neurasthenia, anxious worry, and anxious arousal) and an externalizing factor defined by hazardous use of alcohol accounted for the data better than one- and three-factor alternatives, the latter of which included a two-factor modeling of the internalizing domain. Similarly, when comorbidity patterns among *DSM*-defined disorders were evaluated, Krueger et al. (1998) reported the superiority of a two-factor model over other alternatives, with this model defined by a general internalizing domain (that included major depressive disorder [MDD], dysthymia [DYS], generalized anxiety disorder [GAD], agoraphobia, social phobia [SOC], simple/specific phobia [PHOB], and obsessive-compulsive disorder [OCD]) and a general externalizing domain (that included conduct disorder, marijuana dependence, and alcohol dependence). Other research also support the viability of the single-factor modeling of internalizing disorders (e.g., South & Krueger, 2008) or a single dimension of internalizing pathology where symptom concentrations associated with both anxiety and depressive disorders are located on the upper end of this continuum (Krueger & Finger, 2001; McGlinchey & Zimmerman, 2007).

In a longitudinal investigation with an age-based cohort, Fergusson, Horwood, and Boden (2006) examined the underlying structure of MDD, GAD, PHOB, and panic disorder [PAN] based on separate assessments conducted at three ages (18, 21, and 25). About one-half to three-quarters of the variance of the within-time comorbidity of symptoms among these disorders was accounted for by a general internalizing factor, with the remaining variance

uniquely associated with individual disorders. Similarly, across measurement occasions, a general internalizing factor also accounted for most of the covariation in symptom features associated with these disorders. Such findings support the explanatory power of the single factor model in accounting for both within-time and across-time patterns of internalizing symptom covariation. Consistent with these observations based on phenotypic features, some genetic modeling studies suggest that in addition to common genetic risk factors associated with the broad internalizing/neuroticism domain, there are also genetic risk factors that are unique to MDD, GAD, and PAN not shared with the general domain (Hettema, Neale, Myers, Prescott, & Kendler, 2006).

Other perspectives suggest a more complex structural organization of internalizing psychopathology. Recent editions of the *Diagnostic and Statistical Manual of Mental Disorders (DSM)*; American Psychiatric Association, 1980, 1987, 1994), for example, distinguish “mood disorders” from “anxiety disorders.” As reviewed by Watson (2005, 2009) and exemplified in research by Slade and Watson (2006) and Vollebergh et al. (2001), however, the organization of specific disorders into these two broad categories has frequently not been supported relative to other theoretical models. As a consequence, researchers have evaluated alternative conceptual frameworks for organizing disorders currently ascribed to these two categories.

In an early application of a measurement model framework to the evaluation of the hierarchical structure of mental disorders, Krueger (1999) evaluated the fit of four competing models with confirmatory factor analytic (CFA) methods based on National Comorbidity Survey data. A three-factor model, replicated across random halves of the sample and among men and women, demonstrated the best fit to the data. This three-factor model included two correlated internalizing factors (i.e., labeled “anxious–misery” and “fear”) and one externalizing factor. MDD, DYS, and GAD defined the anxious–misery subfactor, and SOC, PHOB, agoraphobia, and PAN defined the fear subfactor. The externalizing factor was defined by alcohol dependence, drug dependence, and antisocial personality disorder. Subsequent measurement model studies have generally supported a broad internalizing domain further bifurcated into two correlated latent factors (Cox, Clara, Enns, 2002; Slade & Watson, 2006; Vollebergh et al., 2001; see also a meta-analysis by Krueger & Markon, 2006). Consistent with the two-factor organization of the internalizing domain, genetic modeling studies suggest that MDD and GAD share a common genetic diathesis (Kendler, 1996), and that PAN and PHOB only have modest to moderate genetic associations with MDD and GAD (Kendler et al., 1995).

Watson (2005) subsequently proposed a refinement of Krueger’s hierarchical internalizing model based on these and other findings, and suggested this as an alternative to the *DSM-IV* classification framework. In this revised model, mood and anxiety disorders are organized within the same general class (i.e., emotional disorders). This superordinate class is further divided into three subclasses defined by distress disorders (MDD, DYS, GAD, posttraumatic stress disorder [PTSD]), fear disorders (PAN, agoraphobia, SOC, PHOB), and bipolar disorders (i.e., bipolar I and II, cyclothymia). Other putative disorders within this general class, such as OCD, are proposed for further research and evaluation.

Although there is extensive support for the Krueger–Watson hierarchical model of internalizing disorders, there are also suggestions that alternative models might be equally viable. Beesdo-Baum et al. (2009), for example, evaluated model fit across three competing structural models of internalizing and externalizing psychopathology: (a) the Krueger three-factor model that included one externalizing factor, a higher-order internalizing factor, and two internalizing subfactors (i.e., anxious–misery and fear); (b) a variant of the Krueger three-factor model that included one externalizing factor and two correlated internalizing

subfactors, but no higher-order internalizing factor; and (c) an alternative two-factor model based on single internalizing and externalizing factors. With respect to Krueger's original three-factor model, the anxious-misery first-order factor and the second-order internalizing factor were statistically identical in this research, a finding similar to that reported in Krueger (1999), Slade and Watson (2006), and Vollebergh et al. (2001), each of whom found these factors to be almost perfectly correlated (all $r_s \geq .92$). These observations suggest some conceptual problems with this basic model, most notably the redundancy of the higher-order internalizing factor in relation to other model elements. The remaining two models fit the data well, with the Krueger three-factor variant fitting the data slightly better on some indicators than the simpler two-factor model (Beesdo-Baum et al., 2009).

There are additional indications that the basic Krueger model, or variants thereof, may not be robust under varying conditions. Wittchen et al. (2009) found that a variation of the basic Krueger model that excluded the higher-order internalizing factor was not reliably supported across different age ranges between adolescence and early adulthood. They also reported that this model was not supported when additional putative internalizing disorders were added. Although such findings potentially place some constraints on the generalizability of the Krueger–Watson model, they are nonetheless consistent with emerging research that suggests genetic effects on symptoms of anxiety and depression are developmentally dynamic from the young adolescence through the young adult years (Kendler, Gardner, & Lichtenstein, 2008). In the aggregate, such findings suggest that evaluations of comorbidity patterns within relatively brief periods of time (e.g., 12 months) may result in unreliable outcomes, especially when limited numbers of specific internalizing disorders are evaluated.

This Study

There are three primary goals associated with the present research: (a) to comparatively evaluate three competing and theoretically plausible measurement models of putative internalizing disorders based on the patterns of lifetime comorbidity, (b) to evaluate the extent to which viable measurement models are associated with the density of specific psychiatric disorders among first-degree relatives of probands, and (c) to investigate the extent to which viable models of lifetime disorder comorbidity account for psychosocial functioning among probands at age 30. In the context of these evaluations, we simultaneously evaluate both the "lumper" and "splitter" perspectives by examining the relative contributions of the intermediate and upper levels of an internalizing disorder hierarchy in accounting for the patterns of lifetime comorbidity, family aggregation of mental disorders, and psychosocial impairments in adulthood. In so doing, we use a prospective approach to evaluate lifetime comorbidity patterns, and we include in the models evaluated additional putative internalizing disorders that have been frequently omitted in similar past research.

In relation to the first of these goals, an age-based cohort was followed longitudinally, and assessed for internalizing psychiatric disorders on four separate occasions from adolescence through 30. Nine putative internalizing disorders were assessed: MDD, DYS, GAD, PTSD, PHOB, SOC, PAN, OCD, and separation anxiety disorder (SAD). Patterns of lifetime comorbidity among these disorders were, in turn, evaluated with reference to three measurement models, each of which is described immediately below.

Single-factor internalizing model (INT1)

This model, the most parsimonious of the models under study, posits a single latent factor for the nine putative internalizing disorders. Representative research consistent with this model includes Fergusson et al. (2006), Krueger et al. (1998), Krueger et al. (2003), and South and Krueger (2008).

DSM-framework model (INT2a)

This two-factor model is consistent with the organizational scheme implied in *DSM-IV*. In this model, *mood disorders* (MDD, DYS) are distinguished from *anxiety disorders* (GAD, PTSD, PHOB, SOC, PAN, OCD, SAD). According to *DSM-IV*, SAD is an anxiety disorder in name, although it is listed in the "infancy, childhood, or adolescence" section of *DSM-IV* and not within the group designated as "anxiety disorders." SAD's inclusion within the childhood section is likely related to inclusionary criteria that require that the onset of this disorder occur prior to age 18, thus precluding the possibility that this condition can first emerge during adulthood. There is, however, some controversy as to whether this age-based criterion is warranted (Klein, 2009). In addition to the descriptor "anxiety disorder" occurring within the diagnostic label, SAD has also been strongly associated with several other *DSM*-defined anxiety disorders or symptom sets associated with these disorders (Klein, 1995; Lahey et al., 2004; Lahey et al., 2008; Lipsitz et al., 1994).

Prior research that compared the *DSM* organizational framework with alternative models (Slade & Watson, 2006; Vollebergh et al., 2001) within a truncated period of time (i.e., last 12 months) found the *DSM*-based model to be inferior the other models examined.

Krueger–Watson two-factor model (INT2b)

This model distinguishes two latent factors in a manner consistent with a model originally proposed by Krueger (1999) and further refined by Watson (2005): *distress disorders* (MDD, DYS, GAD, PTSD) and *fear disorders* (PHOB, SOC, PAN). This model is informed by research that has generally supported the superiority of this two-factor account of diagnostic comorbidity (Cox et al., 2002; Krueger, 1999; Krueger & Markon, 2006; Slade & Watson, 2006; Vollebergh et al., 2001). The main distinction between this model and the *DSM-IV* model is that GAD and PTSD are modeled within the same factor as MDD and DYS. Research reviewed by Watson (2005) and other subsequent reports (Krueger & Markon, 2006; Miller, Fogler, Wolf, Kaloupek, & Keane, 2008; Moffitt et al., 2007) generally support the inclusion of GAD and PTSD within a factor also defined by mood disorders.

This two-factor model, however, slightly departs from the revised model proposed by Watson (2005). First, it does not include agoraphobia, conceptualized by Watson as a fear disorder, due to a low rate of occurrence in our sample. Second, because of low rates of occurrence in our sample, our evaluation of this model does not include a third latent factor defined by the bipolar disorders (bipolar I, bipolar II, cyclothymia). Third, because of the almost perfect correlations between the anxious–misery/distress subfactor and the higher-order internalizing factor reported in several studies (Beesdo-Baum et al., 2009; Krueger, 1999; Slade & Watson, 2006; Vollebergh et al., 2001), we omitted the higher-order internalizing factor from this model.

In addition to testing the fit of the Krueger–Watson model based on seven disorders specifically identified by Watson (2005), we also evaluated an extension of this model that included both OCD and SAD. Watson's version of this model does not specifically include OCD among the internalizing or emotional disorders. There are, however, findings that support OCD as a robust marker of a superordinate internalizing factor (Krueger et al., 1998). Other studies locate OCD at an intermediate level in the hierarchy, specifically within a latent factor defined by fear disorders that is distinct from a second intermediate-level internalizing latent factor associated with distress disorders (Miller et al., 2008; Slade & Watson, 2006). Lahey et al. (2008) also reported that OCD and fear disorder-related symptoms were best modeled as markers of the same broad latent factor at higher levels in the hierarchy. At lower levels, however, this general factor was bifurcated into two narrower factors, whereby OCD symptoms were modeled as related to but distinct from other fear-

related symptoms (e.g., those associated with PHOB, SAD, and agoraphobia). Although evidence such as this suggests that OCD is a marker of fear disorders at intermediate levels of the hierarchy, other findings reviewed in Watson (2009) suggest that OCD is not a particularly good marker of either fear or distress disorders. As a consequence, in this study we initially evaluate whether OCD is more strongly associated with distress or fear disorders before including it within the Krueger–Watson model evaluated in this research.

We also evaluated whether SAD is more strongly associated with distress or fear disorders with respect to the Krueger–Watson's model. Extant research associates SAD symptom features with those of PHOB, SOC, PAN, OCD, and agoraphobia (Lahey et al., 2004; Lahey et al., 2008). Similarly, childhood SAD is good a predictor of PAN, OCD, SOC, and agoraphobia in adulthood (Klein, 1995; Lipsitz et al., 1994). Overall, these findings suggest that SAD might be more strongly aligned with the fear disorders factor in the Krueger–Watson model, a proposition that we also test.

A second major goal of this research was to evaluate whether viable measurement models are associated with the density of specific psychiatric disorders among first-degree relatives of probands. The intent of such analyses is to evaluate whether there is a non-specific or general internalizing propensity within families, or if there are tendencies for disorders to demonstrate some degree of specificity with respect to family liability.

The third major goal of this research was to investigate the extent to which viable models of lifetime disorder comorbidity account for psychosocial functioning at age 30. McGlinchey and Zimmerman (2007) and Krueger and Finger (2001), based on patient samples, found that a general internalizing/negative affect factor was strongly related to poor social functioning, occupational impairment, and frequency of inpatient psychiatric hospitalization. Other research, however, suggests that prototypical levels of distress and impairment are often greater among persons with *DSM*-defined depressive disorders compared to those with anxiety disorders (Rapoport, Clary, Fayyad, & Endicott, 2005). An exception in this latter research was PTSD, which was associated with comparable levels of impairment as MDD and DYS. Similarly, research by Kessler, DuPont, Berglund, and Wittchen (1999) suggests that GAD and MDD, two disorders that contribute to the distress latent subfactor, are associated with similar levels of functional impairment. Such findings raise the possibility that functional impairment might distinguish latent factors associated with the models evaluated in the present research, and perhaps be more strongly associated with mood and distress latent subfactors among the two-factor models. Consistent with McGlinchey and Zimmerman (2007) and Krueger and Finger (2001), however, Kessler et al. (1999) further reported that the presence of both a mood (i.e., MDD) and an anxiety (i.e., GAD) disorder was consistently associated with more impaired functioning than either type of disorder alone. In the aggregate, these findings suggest that non-specific factors associated with these two classes of disorders largely account for overall functional impairment.

METHOD

Proband Sample

Probands for this research were recruited for the Oregon Adolescent Depression Project, and each participant completed four diagnostic assessments conducted between ages 14 and 33. Lewinsohn, Hops, Roberts, Seeley, and Andrews (1993), Lewinsohn, Rohde, Seeley, Klein, and Gotlib (2003) and Rohde et al. (2007) provide detailed descriptions of sample recruitment, composition, and retention. Briefly, participants were initially recruited from nine randomly selected high schools within western Oregon. The first assessment wave (T_1) included 1,709 adolescents between the ages of 14 and 19 ($M = 16.6$, $SD = 1.2$). About a year after the first assessment, a second assessment wave (T_2) was conducted ($T_1 - T_2$

interval: $M = 13.8$ months, $SD = 2.3$), in which 1,507 members of the original sample (88%) participated. As participants approached their 24th birthday, a third assessment wave (T_3) was conducted. Persons eligible to participate in this assessment included all persons with a positive psychiatric history by T_2 ($n = 644$), as well as a randomly selected subgroup of persons who were negative for any history of mental disorder by T_2 ($n = 457$ of 863 persons). Additionally, to enhance the ethnic diversity of the sample, all non-white T_2 participants were recruited for T_3 . Of the 1,101 persons who were eligible to participate given these selection criteria, 941 persons (or 85%) completed T_3 diagnostic assessments. Finally, all T_3 participants were recruited for a fourth diagnostic assessment as they reached their 30th birthday (T_4). Of the 941 eligible persons, 816 (87%) completed this fourth assessment, with this group constituting the reference sample for this research. These 816 participants were, on average, 30.6 years of age ($SD = 0.6$) at T_4 , predominantly white (89%), female (59%), and married (53%), with fewer than half (41%) having earned a bachelor's degree or higher.

Attrition over the four assessment waves was relatively modest, and differences between persons who continued and discontinued participation were generally small (Rohde et al., 2007). Discontinuation was associated with male gender, cigarette and substance use, a history of a disruptive behavior disorder, lower socioeconomic status, and fewer persons living at home during one's adolescent years.

Diagnostic Assessments of Probands

Participants were interviewed with a version of the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS) at T_1 , T_2 , and T_3 . The version of the K-SADS used combined features of the Present Episode and Epidemiologic versions (Chambers et al., 1985; Orvaschel, Puig-Antich, Chambers, Tabrizi & Johnson, 1982). In conjunction with the K-SADS, the Longitudinal Interval Follow-Up Evaluation (LIFE; Keller et al., 1987) was used to evaluate the presence and course of disorders since participation in the previous diagnostic interview. T_4 diagnostic assessments were based on the Structured Clinical Interview for Axis I *DSM-IV* Disorders–Non-Patient Edition (SCID-NP; First, Spitzer, Gibbon, & Williams, 1994), with the LIFE also used to evaluate disorder presence and course since T_3 . Symptom endorsements during these interviews were subsequently evaluated in conjunction with *DSM-III-R* criteria (American Psychiatric Association, 1987) at T_1 and T_2 and *DSM-IV* criteria (American Psychiatric Association, 1994) at T_3 and T_4 .

Consistent with the methodology employed in similar studies (Krueger, 1999; Krueger et al., 1998; Vollebergh et al., 2001), disorders with a low lifetime rate ($< 1\%$) were not considered in the present research. A total of 9 putative internalizing disorders were examined with respect to lifetime occurrence. Unweighted and weighted lifetime prevalence rates, respectively, through T_4 were: MDD (58.8%; 51.6%), DYS (5.5%; 4.1%), GAD (2.2%; 2.0%), PTSD (8.6%; 7.0%), PHOB (4.7%; 3.6%), SOC (4.4%; 3.6%), PAN (7.2%; 6.1%), OCD (1.5%; 1.1%), and SAD (5.4%; 3.9%). Weighted rates adjust for the probability sampling procedures introduced at T_3 described earlier in this section.

A subset of interviews from each wave was rated from audio or videotapes by a second interviewer for reliability purposes: $T_1 = 263$, $T_2 = 162$, $T_3 = 190$, and $T_4 = 124$ interviews. Diagnostic agreement was calculated for disorders that were diagnosed 10 or more times by both raters combined. Across the four assessment waves, inter-rater diagnostic reliability was good to excellent for disorders that occurred with sufficient frequency. The reliability of one disorder, MDD, could be determined for each of the four waves ($M\kappa = .84$; range: .81 – .86). Kappa coefficients for DYS (.56), PTSD (.73), PHOB (.66), PAN (.81), and SAD (.83) could only be determined for one of the four assessment waves. The remaining

disorders (GAD, SOC, OCD) were not diagnosed with sufficient frequency during any assessment wave so as to allow an evaluation of diagnostic reliability.

Diagnostic Assessments of Family Members

A total of 2,414 first-degree relatives of 732 probands were interviewed with the SCID-NP (First et al., 1994) or, if under age 18, the K-SADS (Chambers et al., 1985). Retrospective assessments of past psychiatric disorders were assessed with the revised Family Informant Schedule and Criteria, modified for *DSM-IV* (FISC; Mannuzza & Fyer, 1990), or the Epidemiologic version of the K-SADS (Orvaschel et al., 1982) if the participant was less than 18-years old. Of the first-degree relatives interviewed, 30.2% ($n = 730$) were the mothers of probands, 29.8% ($n = 719$) were fathers, 19.7% ($n = 476$) were female siblings, and 20.3% ($n = 489$) were male siblings. Diagnostic categories assessed with both interviews and the corresponding mean proportion (density) of family members who met lifetime criteria for specific diagnostic categories were as follows: MDD = .28, DYS = .04, PHOB = .03, SOC = .03, PAN = .03, and substance use disorder (SUD = .36), which includes the abuse of or dependence on alcohol, cannabis, or other substances. Although OCD was assessed with both interviews, their prevalence among family members (< 1%) was too low to include in subsequent analyses.

As detailed in Klein, Lewinsohn, Seeley, and Rohde (2001), the best estimate method (Leckman, Sholomskas, Thompson, Balanger, & Weissman, 1982) for determining lifetime psychiatric diagnoses among family members was used. Inter-diagnostician agreement, based on all available interview data and indexed by kappa, was excellent for the diagnostic categories evaluated: MDD = .91, DYS = .89, PHOB = .91, SOC = .94, PAN = .90, and SUD = .97.

Psychosocial Functioning at T₄

Measurement models of putative internalizing disorders were also separately evaluated in terms of their ability to account for variance in several indicators of functional impairment assessed at T₄ (age 30). Areas assessed included unemployment during the past year (6-point scale of number of weeks unemployed), annual household income (6-point scale ranging from no income to \$50,000 or more), history of divorce or separation, relationship quality with family and friends ($\alpha = .90$; 20 items; Prociano & Heller, 1983), level of social adjustment during the past two weeks ($\alpha = .70$; 54 items; Weissman & Bothwell, 1976; higher scores indicate poorer adjustment), and *DSM-IV* Global Assessment of Functioning scale.

Statistical Analyses and Model Selection

Model evaluation and selection were guided by statistical indicators of model fit and parsimony (Myung, 2000). Models were first contrasted in terms of their goodness-of-fit to the data with confirmatory factor analytic (CFA) methods. This was followed with model comparisons based on information-theoretic criteria, which balances model complexity and model fit, with overall preference given to well-fitting and simpler models. Several indicators were used in the evaluation of model fit and interpreted in conjunction with critical cut points recommended by Yu (2002). Fit indicators (with associated cut-off values for denoting “good fit”) were: chi-square test of model fit (χ^2 ; $p > .05$), comparative fit index (CFI; .96), Tucker–Lewis fit index (TLI; .95), root mean square error of approximation (RMSEA; .05), and weighted root mean square residual (WRMR; 1.0). CFA analyses were performed with Mplus statistical software (version 5.1; Muthén & Muthén, 1998–2007), and model parameters were estimated using the weighted least squares estimator procedure with means and variances adjusted (WLSMV). This estimator is

preferable in the present research as it is comparatively robust against violations of distribution normality (Curran, West, & Finch, 1996; Flora & Curran, 2004).

Model selection was guided by two information criterion statistics based on full-information maximum likelihood estimation¹: Akaike information criterion (AIC) and the sample-size adjusted Bayesian information criterion (Adj. BIC). These statistics are similar in that they permit comparisons between processes denoted in specified theoretical models and characteristics present in the data, and they each provide an index of the optimal balance of model fit with parsimony. These statistics differ, however, in relation to how model complexity is penalized. BIC more so than AIC favors less complex models, particularly when sample sizes are small to moderate in size, by imposing a harsher penalties on additional model parameters. The sample-size Adj. BIC used here is less biased by sample size, does not impose as harsh a penalty in fitting additional parameters as BIC, and is a more accurate statistic than BIC under some circumstances (Henson, Reise, & Kim, 2007; Yang, 2006).

Lower Adj. BIC and AIC values indicate greater parsimony and fit. Because AIC values are relative, they were contrasted among competing models based on procedures outlined by Burnham and Anderson (2004). Difference scores (Δ AIC) were derived by subtracting the AIC value associated with the best approximating model from AIC values of other models within the set. AIC values ≤ 2.0 are interpreted as similar competing models in terms of their approximating abilities, those within the range of 4.0 $\leq$ AIC $\leq$ 7.0 indicate models that have considerably less support, and models with AIC values > 10.0 have essentially no support relative to the best approximating model. There are presently no comparable guidelines for interpreting Adj. BIC. Model selection was therefore guided by (a) fit statistics from CFA analyses, (b) AIC and Adj. BIC values, and (c) interpretative guidelines provided by Burnham and Anderson based on AIC values.

In the event that either two-factor measurement model (INT2a or INT2b) fit the data well in the CFA analyses, we also performed parallel sets of analyses that separately evaluated shared and unique contributions of latent factors to the prediction of family density of psychopathology and psychosocial outcomes. The intent of such analyses was to clarify if any observed associations among these variables were the result of variance uniquely associated with a given latent factor, variance common to both factors, or both unique and common variance. The outcomes of such analyses, in conjunction with findings from the CFAs, would be relevant to the "lumper" versus "splitter" dialectic, and suggest levels of categorical description that are potentially most useful for a given purpose.

Finally, because of the implementation of stratified sampling procedures at T₃ that were continued through T₄, Caucasian participants without a history of psychiatric diagnosis by T₂ were under-sampled in subsequent assessment waves. To adjust for this, a weighting procedure was used whereby Caucasian participants without a lifetime diagnosis by T₂ were assigned a weight of 2.05. This value reflects the probability of this subgroup being sampled at T₃. All participants with a psychiatric diagnosis by T₂ and all non-Caucasian participants were assigned a weighting of 1.0. Findings presented in the following section are based on weighted data.

¹Maximum likelihood estimation with robust standard errors produces unbiased estimates for models with categorical outcomes (Brown, 2006). However, model fit indices (i.e., CFI, TLI, RMSEA, WRMR) are not available when maximum likelihood methods are used with categorical outcomes. As such, model fit and factor loadings were derived from weighted least squares estimation and model selection was based on AIC and BIC values using maximum likelihood estimation.

RESULTS

Associated Risks among Diagnostic Categories

Tetrachoric correlations and odds ratios based on diagnostic categories are presented in Table 1. The tetrachoric correlation matrix was derived with Mplus and odds ratios were calculated with SPSS statistical software by exponentiating the beta coefficients in logistic regression models. Correlations among the internalizing disorders were mostly moderate in magnitude and positive, and generally comparable albeit somewhat attenuated to those reported by others (see Watson, 2009, for an integrative summary). Odds ratios also suggest that the occurrence of any disorder often substantially increases the risk for any other disorder from the internalizing domain. Some exceptions are evident, however, and these are indicated by the 95% confidence intervals (CIs) that include 1.0.

Preliminary Analyses: The Placement of OCD and SAD within the Krueger–Watson Framework

Both OCD and SAD have been previously associated with the family of internalizing disorders (e.g., Lahey et al., 2004; 2008), and both are regarded as examples of anxiety disorders within the *DSM-IV* classification framework. Although not explicitly represented within Krueger–Watson hierarchical internalizing model, research reviewed earlier generally supports the inclusion of OCD and SAD within the “fear disorders” domain. In tests of these hypotheses, OCD and SAD were separately regressed on the distress and fear latent factors, with these factors defined by disorder categories explicitly identified by Watson (2005) in his framework (distress disorders = MDD, DYS, GAD, and PTSD; fear disorders = PHOB, SOC, and PAN). As expected, OCD demonstrated a stronger association with the fear disorders latent factor than the distress disorders latent factor (standardized coefficients = 0.78 and –0.05, respectively), as did SAD (standardized coefficients = 0.42 and 0.20, respectively). Consequently, analyses presented below involving the extended Krueger–Watson model include OCD and SAD among the fear disorders.

Comparative Evaluation of Internalizing Measurement Models

The single-factor (INT1), *DSM-IV* (INT2a), and Krueger–Watson (INT2b) models were evaluated for model fit. In one set of analyses, models were evaluated with reference to the seven diagnostic categories that are explicitly common to each of the three models (i.e., MDD, DYS, GAD, PTSD, PHOB, SOC, and PAN). In the second set of analyses, OCD and SAD were added to the single-factor model, the anxiety disorders latent factor of the *DSM-IV* model, and the fear disorders latent factor of the Krueger–Watson model. The resultant model fit indicators and selection criteria for each set of analyses for the three competing models are presented in Table 2.

When measurement models are based on seven diagnostic categories, fit indicators suggest adequate to good fit for each model, with the adjusted BIC, AIC, and AIC selection criteria not clearly elevating one model above another. Similarly, fit indicators and selection criteria for each of the extended nine-category models that include OCD and SAD indicate good model fit in each instance and similar approximating abilities. Measurement models with parameter estimates for the nine-variable single factor (INT1), *DSM-IV* (INT2a), and Krueger–Watson (INT2b) models are presented in Figure 1.

Associations with Family History Data Assessed at T₃

Family history data were evaluated in relation to the latent factors associated with the three measurement models presented in Figure 1. In these regression-based analyses performed with Mplus (Version 5.1; Muthén & Muthén, 1998–2007), factor scores associated with each measurement model were outputted from the CFA analyses and subsequently modeled

as predictors of psychiatric disorder density among first-degree relatives of probands. In addition, to obtain different perspectives on shared versus unique contributions of mood/anxiety or distress/fear disorder latent factors in relation to family density data, two data analytic approaches were performed for the two-factor models (INT2a and INT2b). In the *unadjusted approach*, factor scores associated with each latent factor were evaluated with separate regression models. By modeling the association between a given predictor and outcome in this manner, possible shared variance with the latent factor not included in the model was preserved. In the *adjusted approach*, factor scores associated with the two latent factors associated with a given measurement model were entered simultaneously as predictors in the regression models. By modeling latent factors in this way, shared or common variance among latent factors was removed from the model, with the resultant regression coefficients denoting the unique (i.e., non-shared) contributions of a given latent factor to the prediction of family history data.

Except for SOC, all other psychiatric disorder categories among first-degree relatives had significant associations with the one-factor model, the mood and anxiety latent factors of the *DSM-IV* model (INT2a), and the distress and fear latent factors associated with the Krueger–Watson model (INT2b) when the unadjusted approach was used (Table 3). When the adjusted approach was used, however, all coefficients were no longer significant except for PAN when associated with the anxiety disorder latent factor of the *DSM-IV* model (adjusted coefficient = .46, $p = .014$, semi-partial $r = .149$). The overall pattern of findings from the unadjusted and adjusted analyses suggests that the associations with the psychiatric disorders among relatives of probands is almost completely accounted for by shared variance among mood/anxiety and distress/fear latent factors, and that unique variance associated with latent factors that define both two-factor models adds little to their prediction.

Associations with Psychosocial Functioning at T₄

The above approaches to evaluating model associations with family density data were similarly used in the evaluation of T₄ psychosocial functioning measures. When latent factors were evaluated separately with respect to their ability to predict T₄ psychosocial functioning (i.e., the unadjusted approach; Table 4), each of the latent factors demonstrated robust and strikingly similar associations with impairment-related variables. The single internalizing factor (INT1), mood and anxiety latent factors of the *DSM-IV* model (INT2a) and the distress and fear latent factors of the Krueger–Watson model (INT2b) were each significantly associated with all six variables. These latent factors were positively associated with unemployment during the last 12 months, history of divorce or separation, and poor social adjustment. Significant negative associations for each of the latent factors were obtained for the following outcome variables: income over the last 12 months, quality of relationships, and global functioning.

When two-factor models (INT2a and INT2b) were reanalyzed with the adjusted method, many of the significant coefficients from the unadjusted analysis were no longer significant. This pattern of findings suggests that significant coefficients presented in Table 4 are largely the product of shared variance among latent factors. However, some unique effects associated with the anxiety latent factor of the *DSM-IV* model (INT2a) were obtained. These were evident in significant and positive associations with history of divorce or separation (adjusted coefficient = .30, $p = .023$, semi-partial $r = .093$) and poor social adjustment (adjusted coefficient = .26, $p = .014$, semi-partial $r = .079$), and a significant negative association with global functioning (adjusted coefficient = $-.34$, $p = .007$, semi-partial $r = -.121$). With respect to the Krueger–Watson model (INT2b), the only significant effect emerged for the fear disorders latent factor in relation to poor social adjustment (adjusted coefficient = .46, $p = .014$, semi-partial $r = .074$). Overall, the difference in the

unadjusted and adjusted findings suggest that the robust associations that latent factors have with psychosocial functioning in adulthood is largely related to non-specific or shared variance associated with these factors, and that comparatively little additional predictive power is achieved when the superordinate internalizing factor (INT1) is further bifurcated into lower-order factors (either INT2a or INT2b).

DISCUSSION

Three competing measurement models (single-factor, *DSM-IV*, and Krueger–Watson) were evaluated in terms of their ability to account for patterns of lifetime comorbidity among nine putative internalizing disorders. Each model emerged as viable and basically indistinguishable when referenced to model fit and selection indicators. These findings are consistent with the range of results from similar studies reported in the literature that support both the single-factor internalizing model as well as the Krueger–Watson model. The main departure between our findings and those reported by others (Slade & Watson, 2006; Vollebergh et al., 2001) is that the *DSM-IV* distinction of “anxiety” and “mood” disorders performed as well as other competing models evaluated in the present research. Overall, both lumpers and splitters perspectives were supported with respect to the measurement model findings. These findings also highlight the substantial overlap among internalizing domain disorders, and are broadly supportive of the reorganization of mood and anxiety disorders into an inclusive super-class of emotional or internalizing disorders (Fergusson, et al., 2006; Watson, 2005).

Findings from this study also supported PTSD, OCD, and SAD as robust indicators of general internalizing pathology. GAD, in contrast, demonstrated comparatively modest associations with latent factors at the upper and intermediate levels of the hierarchy. This finding raises the possibility that GAD might generally function as a comparatively weak liability marker, or might be more optimally modeled according to an alternative theoretical framework not examined here. Whereas research reviewed by Watson (2005) and subsequent research by Moffitt et al. (2007) support the inclusion of GAD within the same latent factor as MDD, other research suggests that GAD is more appropriately modeled as an anxiety disorder rather than a mood disorder and that GAD and MDD are unlikely to be markers of the same latent factor given several distinguishing features (Kessler et al., 2008; Verduin & Kendall, 2003).

Tetrachoric correlations and odds ratios for the disorders examined in the present research occasionally depart from those reported by others. Watson (2009), for example, presents weighted mean tetrachoric correlations based on four national epidemiological studies. The correlations obtained in the present research are generally similar, albeit somewhat attenuated, to those reported by Watson. There are also some significant departures, however, most notably in the correlations for DYS and GAD, DYS and SOC, and GAD and PAN. Several factors likely accounted for these divergent observations. First, correlations presented by Watson (2009) are based on period-prevalence data (i.e., diagnostic comorbidity during the last 12 months) from four large epidemiological samples of adults, whereas data reported here are based on patterns of lifetime comorbidity as assessed at four time periods within a 15 year interval from adolescence to adulthood. Structural relations among diagnostic comorbidity have been observed to vary as a function of age (Wittchen et al., 2009), and there are similar indications that genetic effects on symptoms of anxiety and depression are developmentally dynamic through the adolescent and early adult years (Kendler et al., 2008). In support of this explanation, Griffith et al. (2010) reported correlations between lifetime rates of disorders in an adolescent sample that were also considerably lower than those reported by Watson (2009). Second, diagnostic interviewers in the present study followed *DSM* exclusionary criteria when considering diagnoses at each

assessment wave. Data utilized by Watson (2009) were from studies that did not impose hierarchical exclusion rules. Adherence to these diagnostic decision rules in this study likely reduced the base-rate of both GAD and DYS, and attenuated relations among other mood and anxiety disorders (Watson, 2009). Third, criteria for some diagnostic concepts, such as GAD, were substantially modified from *DSM-III-R* to *DSM-IV*, perhaps resulting in inconsistent associations between GAD and other disorders across assessment waves. As noted by Watson, changes in GAD criteria were primarily undertaken to better differentiate GAD from PAN. In our sample, most of the diagnosed cases of PAN occurred at T₃ and T₄ when *DSM-IV* criteria were in effect, likely resulting in attenuated patterns of comorbidity compared to what might be observed with *DSM-III-R* diagnostic decision rules. In addition, from *DSM-III-R* to *DSM-IV*, several hierarchical exclusion rules were eliminated. Such changes in diagnostic criteria and diagnostic decision rules likely resulted in different patterns of association among pairs of disorders across different editions of *DSM*. Fourth, in the present research, GAD and SOC were two of three diagnostic categories for which diagnostic reliability could not be determined, thus raising the possibility that these inconsistent findings could be related to the degree to which these categories were reliably assessed.

The three viable measurement models that emerged from this research were further evaluated in terms of their associations with family densities of specific psychiatric disorders and psychosocial impairments at age 30. Findings were remarkably similar for both variable sets, with most of the variance accounted for in these analyses related to a general internalizing factor. This general factor emerged as a potent predictor of the family density of five of the six psychiatric disorders evaluated among first-degree relatives (i.e., MDD, DYS, PHOB, PAN, and SUD) and each of the six psychosocial functioning indicators evaluated among probands. The finding that variance associated with a general internalizing factor largely accounted for the prediction of the family density of most internalizing psychiatric disorders is consistent with accumulating data that implicate neuroticism as a heritable personality trait (Flint, 2004), with as much as 40% of variance in this trait attributable to genetic variation (Lake, Eaves, Maes, Heath, & Martin, 2000). The present research also replicated and extended findings previously reported by McGlinchey and Zimmerman (2007) and Krueger and Finger (2001), who found that a general internalizing/negative affect factor was related to several aspects of psychosocial functioning.

Although the more complex two-factor models of *DSM-IV* and Watson generally accounted for similar amounts of overall variance among family density and psychosocial functioning as the single-factor model, the intermediate-level latent factors associated with each of these models were not well-distinguished in relation to these variables. That is, once common or non-specific variance shared by both intermediate factors was removed, the unique variance associated with individual factors of these two-factor models generally accounted for no additional significant variance in the predicted variables.

Overall, findings from the present research are consistent with an emerging developmental view that the lifetime comorbidity of anxiety and depressive disorders is not random, but rather reflects the influence or impact of related processes. There are, for example, indications that genes related to the emergence of anxiety disorders in middle childhood are the same genes associated with the expression of depressive disorders during adolescence (Silberg, Rutter, & Eaves, 2001). Similarly, episodes of anxiety disorders usually precede depressive disorders (e.g., Kovacs, Gatsonis, Paulauskas, & Richards, 1989), and childhood anxiety disorders frequently predict future episodes of depression (Caspi, Moffitt, Newman, & Silva, 1996; Costello, Mustillo, Erkanli, Keeler, & Angold, 2003). Previously we reported data from this sample that also suggested that the onset of anxiety disorders usually predates

the onset of depressive disorders (Orvaschel, Lewinsohn, & Seeley, 1995; Rohde, Lewinsohn, & Seeley, 1991), and Kessler, Avenevoli, and Merikangas (2001) have similarly reported that about three-quarters of youth with a depressive disorder have an earlier history of at least one anxiety disorder. Overall, anxiety disorders and depressive disorders are strongly related, likely share common etiological determinants, and have distinct temporal associations. As this study also demonstrates, a lifetime history of clinically significant levels of general negative affect also predicted histories of specific psychiatric disorders among first degree relatives and impaired psychosocial functioning in early adulthood.

Study Strengths and Limitations

There are several strengths associated with the present research that, in the aggregate, might have resulted in greater support for multiple measurement models in comparison to other studies that generally provide support for one model over other alternatives. Participants in this study were selected from an age-based cohort that was followed longitudinally and assessed within an age range during which the first emergence of several psychiatric disorders is likely (e.g., Kessler et al., 2005). Similar studies have frequently utilized samples that are heterogeneous in age, which can introduce a potential bias if disorders are more prevalent among persons within certain age ranges and if these age ranges are under-represented in a sample.

The estimation of lifetime incidence and comorbidity in the present study was based on repeated diagnostic assessments over two consecutive decades. Our emphasis on lifetime diagnostic comorbidity through age 30 covers a time period that is substantially longer than most studies, a majority of which usually assess for a time period covering the last 12 months. Other studies that have evaluated lifetime comorbidity patterns are usually based on recollections of past psychiatric history, a process that is known to produce biased reports and substantially lower prevalence rates than those found in prospective studies (Moffitt et al., 2010; Wells & Horwood, 2004). Such recall biases were likely minimized in the present study by conducting multiple diagnostic assessments across the study period. Because of our emphasis on lifetime psychiatric history rather than current psychiatric conditions, prevalence rates were higher for several disorders, thus allowing us to incorporate more putative internalizing disorders into our measurement models than is typically found in research of this type.

Although this study included some desirable design features, there are also some limitations that should be considered. Several of the mood and anxiety disorders in *DSM* are defined by symptom sets that are themselves multidimensional (Watson, 2005). Given the focus of this research on psychiatric disorders, with diagnostic thresholds based on notions concerning disorder severity and indications for treatment (Fergusson, Horwood, & Lynskey, 1994), this research was primarily concerned with patterns of covariation found at the extremes of variables that have an underlying continuous distribution. When analyses are focused at the taxonomic level, the underlying dimensionality associated with diagnostic criteria for specific disorders is ignored. Consequently, associations among symptoms that define *DSM* mood and anxiety disorders might show different patterns of associations than those obtained here based on diagnostic categories.

Second, different methods were used to assess diagnostic categories across waves. Whereas the K-SADS was used from T₁ to T₃, the SCID was used at T₄. Similarly, *DSM-III-R* criteria were used at T₁ and T₂, and *DSM-IV* criteria at T₃ and T₄. Changes in diagnostic criteria and assessment approaches could have introduced method bias into the study that, in turn, influenced comorbidity patterns among disorders.

Third, inter-rater reliability was not established for GAD, SOC and OCD due to the low number of positive diagnoses among interviews selected for reliability evaluation. Reliability data on other disorders were also not consistently available for each of the four assessment waves. On occasions when reliability data were available, however, the level of agreement among raters was good to excellent. The extent to which reliability issues might account for the comparatively low coefficients for GAD and SOC in the measurement models presented in Figure 1, however, remains uncertain.

Fourth, the composition of the sample warrants further discussion. Participants for this study were ethnically and geographically homogeneous; consequently, the extent to which study findings are generalizable to other populations remains unclear. Similarly, because of the small number of participants with lifetime diagnoses of some disorders, we did not separately evaluate models based on gender. Prior research, however, has found that similar measurement models are frequently gender invariant (Kendler, Prescott, Myers, & Neale, 2003; Krueger, 1999; Lahey et al., 2008). The lifetime prevalence of MDD in the present sample was also substantially higher than other reports. When compared to the Dunedin study that employed a prospective design covering a similar age span (Moffitt et al., 2010), however, the weighted lifetime rate of MDD in the current research was similar (41% vs. 51%, respectively). In contrast, the lifetime prevalence rates of several anxiety disorders in the present sample were lower than those obtained in other studies (e.g., Kessler et al., 2005; Moffitt et al., 2010). Hence, the diagnostic homogeneity among sample participants may have constrained the degree to which the various models could be differentiated. Lastly, the size of the sample utilized here was smaller than most of the previous studies that have examined the structure of internalizing disorders resulting in lower statistical power compared to previous investigations.

Conclusions and Future Directions

As findings from this study illustrate, three competing measurement models accounted well for lifetime comorbidity patterns among internalizing-related diagnostic categories. Therefore, it may be premature to substantially modify the existing *DSM* nosologic framework given the absence of a clearly superior alternative model. Additional tests of theoretically plausible measurement models are needed, as is research on the validity of different model configurations based on other theoretically-relevant variables (e.g., genetics, treatment response, disorder course). Other areas for future investigation include the covariation of specific sets of internalizing symptoms or behaviors, both within discrete moments of time as well as over the lifetime, and a corresponding evaluation at the symptom level of the amount of a general distress/negative affect component contributing to its occurrence versus the level of specificity associated with the symptom (Watson, 2009). Such analyses would likely yield substantial benefit, particularly given the heterogeneity of symptom dimensions within *DSM* diagnostic categories (Watson, 2005). Previous symptom-based analyses, for example, illustrate that some symptoms of depression and anxiety have specificity, whereas others are common or shared across these two constructs (Watson, 2009; Zinbarg & Barlow, 1996). Symptoms that have specific or relatively unique associations with depressed mood include sadness, loss of interest or pleasure, hypersomnia, suicidal ideation, and helplessness. Symptoms relatively specific to anxiety include apprehension and those linked to autonomic activation tied to the fight/flight/freeze response, such as symptoms common to episodes of panic (e.g., heart palpitations, rapid and shallow breathing, trembling, muscle tension) (Barlow, Chorpita, & Turovsky, 1996; Brown, Chorpita, & Barlow, 1998; Clark & Watson, 1991; Watson, 2009; Zinbarg & Barlow, 1996).

Additional research might also seek to establish if examples of diagnostic comorbidity observed here and elsewhere indicate problems with discriminant validity of *DSM* disorder

criterion sets, whereby two or more disorder labels are invoked to account for the same underlying process or phenomenology (Watson, 2005). Future refinements of diagnostic categories could be guided by such findings and, in turn, produce more reliable and valid constructs that hold greater promise for furthering the science of psychopathology.

Acknowledgments

National Institute of Mental Health Grants MH40501 and MH50522, and National Institute on Drug Abuse Grant DA12951 supported this research.

REFERENCES

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 3rd ed.. Washington, D.C: Author; 1980.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 3rd ed.. Washington, D.C: Author; 1987.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed.. Washington, DC: Author; 1994.
- Barlow, DH.; Chorpita, BF.; Turovsky, J. Fear, panic, anxiety, and disorders of emotion. In: Hope, DA., editor. Perspectives on anxiety, panic, and fear (The 43rd Annual Nebraska Symposium on Motivation; Lincoln, NE. Nebraska University Press; 1996. p. 251-328.
- Beesdo-Baum K, Höfler M, Gloster AT, Klotsche J, Lieb R, Beauducel A, et al. 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*. 2009; 18:204–220. [PubMed: 20024895]
- Brown, TA. Confirmatory factor analysis for applied research. New York: Guilford Press; 2006.
- Brown TA, Chorpita BF, Barlow DH. Structural relationships among dimensions of the *DSM-IV* anxiety and mood disorders and dimensions of negative affect, positive affect, and autonomic arousal. *Journal of Abnormal Psychology*. 1998; 107:179–192. [PubMed: 9604548]
- Burnham KP, Anderson DR. Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*. 2004; 33:261–304.
- Caspi A, Moffitt TE, Newman DL, Silva PA. Behavioral observations at age 3 years predict adult psychiatric disorders: longitudinal evidence from a birth cohort. *Archives of General Psychiatry*. 1996; 53:1033–1039. [PubMed: 8911226]
- Chambers WJ, Puig-Antich J, Hirsch M, Paez P, Ambrosini PJ, Tabrizi MA, et al. The assessment of affective disorders in children and adolescents by semistructured interview: Test-retest reliability of the Schedule for Affective Disorders and Schizophrenia for School-Age Children, Present Episode version. *Archives of General Psychiatry*. 1985; 42:696–702. [PubMed: 4015311]
- Clark LA, Watson D. Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*. 1991; 100:316–336. [PubMed: 1918611]
- Costello E, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry*. 2003; 60:837–844. [PubMed: 12912767]
- Cox BJ, Clara IP, Enns MW. Posttraumatic stress disorder and the structure of common mental disorders. *Depression and Anxiety*. 2002; 15:168–171. [PubMed: 12112721]
- Curran PJ, West SG, Finch JF. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*. 1996; 1:16–29.
- Fergusson DM, Horwood LJ, Boden JM. Structure of internalising symptoms in early adulthood. *British Journal of Psychiatry*. 2006; 189:540–546. [PubMed: 17139039]
- Fergusson DM, Horwood LJ, Lynskey MT. The comorbidities of adolescent problem behaviors: A latent class model. *Journal of Abnormal Child Psychology*. 1994; 22:339–354. [PubMed: 8064037]

- First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. Structured Clinical Interview for Axis I DSM-IV Disorders–Non-Patient Edition. New York: Biometrics Research Department; 1994.
- Flint J. The genetic basis for neuroticism. *Neuroscience and Biobehavioral Reviews*. 2004; 28:307–316. [PubMed: 15225973]
- Fleiss, JL. Statistical methods for rates and proportions. 2nd ed.. New York: Wiley; 1981.
- Flora DB, Curran PJ. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*. 2004; 9:466–491. [PubMed: 15598100]
- Griffith JW, Zinbarg RE, Craske MG, Mineka S, Rose RD, Waters AM, et al. Neuroticism as a common dimension in the internalizing disorders. *Psychological Medicine*. 2010; 40:1125–1136. [PubMed: 19903363]
- Gregory AM, Eley TC. Genetic influences on anxiety in children: What we’ve learned and where we’re heading. *Clinical Child and Family Psychology Review*. 2007; 10:199–212. [PubMed: 17503180]
- Henson JM, Reise SP, Kim KH. Detecting mixtures from structural differences using latent variable mixture modeling: A comparison of relative model fit statistics. *Structural Equation Modeling*. 2007; 14:202–226.
- Hettema JM, Neale MC, Myers JM, Prescott CA, Kendler KS. A population-based twin study of the relationship between neuroticism and internalizing disorders. *American Journal of Psychiatry*. 2006; 163:857–864. [PubMed: 16648327]
- Keller MB, Lavori PW, Friedman B, Nielsen E, Endicott J, McDonald-Scott PA. The Longitudinal Interval Follow-up Evaluation: a comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry*. 1987; 44:540–548. [PubMed: 3579500]
- Kendler KS. Major depression and generalised anxiety disorder: Same genes, (partly) different environments-revisited. *British Journal of Psychiatry*. 1996; 168:68–75. [PubMed: 8770431]
- Kendler KS, Gardner CO, Lichtenstein P. A developmental twin study of symptoms of anxiety and depression: Evidence for genetic innovation and attenuation. *Psychological Medicine*. 2008; 38:1567–1575. [PubMed: 18578897]
- Kendler KS, Prescott CA, Myers J, Neale MC. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry*. 2003; 60:929–937. [PubMed: 12963675]
- Kendler KS, Walters EE, Neale MC, Kessler RC, Heath AC, Eaves LJ. The structure of the genetic and environmental risk factors for six major psychiatric disorders in women: Phobia, generalized anxiety disorder, panic disorder, bulimia, major depression, and alcoholism. *Archives of General Psychiatry*. 1995; 52:374–383. [PubMed: 7726718]
- Kessler RC, Avenevoli S, Merikangas KR. Mood disorders in children and adolescents: An epidemiologic perspective. *Biological Psychiatry*. 2001; 49:1002–1014. [PubMed: 11430842]
- Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*. 2005; 62:593–602. [PubMed: 15939837]
- Kessler RC, DuPont RL, Berglund P, Wittchen H-U. Impairment in pure and comorbid generalized anxiety disorder and major depression at 12 months in two national surveys. *American Journal of Psychiatry*. 1999; 156:1915–1923. [PubMed: 10588405]
- Kessler RC, Gruber M, Hettema JM, Hwang I, Sampson N, Yonkers KA. Comorbid major depression and generalized anxiety disorders in the National Comorbidity Survey follow-up. *Psychological Medicine*. 2008; 38:365–374. [PubMed: 18047766]
- Klein DN, Lewinsohn PM, Seeley JR, Rohde P. A family study of major depressive disorder in a community sample of adolescents. *Archives of General Psychiatry*. 2001; 58:13–20. [PubMed: 11146753]
- Klein R. Is panic disorder associated with childhood separation anxiety disorder? *Clinical Neuropharmacology*. 1995; 18:S7–S14.
- Klein RG. Anxiety disorders. *Journal of Child Psychology and Psychiatry*. 2009; 50:153–162. [PubMed: 19220598]

- Kovacs M, Gatsonis S, Paulauskas SL, Richards C. Depressive disorders in childhood. IV: A longitudinal study of comorbidity with and risk for anxiety disorders. *Archives of General Psychiatry*. 1989; 46:776–783. [PubMed: 2774847]
- Krueger RF. The structure of common mental disorders. *Archives of General Psychiatry*. 1999; 56:921–926. [PubMed: 10530634]
- Krueger RF, Caspi A, Moffitt TE, Silva PA. The structure and stability of common mental disorders (DSM-III-R): A longitudinal-epidemiological study. *Journal of Abnormal Psychology*. 1998; 107:216–227. [PubMed: 9604551]
- Krueger RF, Chentsova-Dutton YE, Markon KE, Goldberg D, Ormel J. A cross-cultural study of the structure of comorbidity among common psychopathological syndromes in the general health care setting. *Journal of Abnormal Psychology*. 2003; 112:437–447. [PubMed: 12943022]
- Krueger RF, Finger MS. Using item response theory to understand comorbidity among anxiety and unipolar mood disorders. *Psychological Assessment*. 2001; 13:140–151. [PubMed: 11281035]
- Krueger RF, Markon KE. Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology*. 2006; 2:111–133.
- Lahey BB, Applegate B, Waldman ID, Loft JD, Hankin BL, Rick J. The structure of child and adolescent psychopathology: Generating new hypotheses. *Journal of Abnormal Psychology*. 2004; 113:358–385. [PubMed: 15311983]
- Lahey BB, Rathouz PJ, Van Hulle C, Urbano RC, Krueger RF, Applegate B, et al. Testing structural models of DSM-IV symptoms of common forms of child and adolescent psychopathology. *Journal of Abnormal Child Psychology*. 2008; 36:187–206. [PubMed: 17912624]
- Lake RIE, Eaves LJ, Maes HHM, Heath AC, Martin NG. Further evidence against the environmental transmission of individual differences in neuroticism from a collaborative study of 45,850 twins and relatives on two continents. *Behavior Genetics*. 2000; 30:223–233. [PubMed: 11105396]
- Leckman JF, Sholomskas D, Thompson WD, Balanger A, Weissman MM. Best-estimate of lifetime psychiatric diagnoses. *Archives of General Psychiatry*. 1982; 39:879–883. [PubMed: 7103676]
- Lewinsohn PM, Hops H, Roberts RE, Seeley JR, Andrews JA. Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology*. 1993; 102:133–144. [PubMed: 8436689]
- Lewinsohn PM, Rohde P, Seeley JR, Klein DN, Gotlib IH. Psychosocial functioning of young adults who have experienced recovery from major depressive disorder during adolescence. *Journal of Abnormal Psychology*. 2003; 112:353–363. [PubMed: 12943014]
- Lipsitz JD, Martin LY, Mannuzza S, Chapman TF, Liebowitz MR, Klein DF, et al. Childhood separation anxiety disorders in patients with adult anxiety disorders. *American Journal of Psychiatry*. 1994; 151:927–929. [PubMed: 8185008]
- Mannuzza, S.; & Fyer, AJ. *Family Informant Schedule and Criteria (FISC)*. New York, NY: New York State Psychiatric Institute; 1990.
- McGlinchey JP, Zimmerman M. Examining dimensional representation of depression and anxiety disorders' comorbidity in psychiatric outpatients with item response modeling. *Journal of Abnormal Psychology*. 2007; 116:464–474. [PubMed: 17696702]
- Middeldorp CM, Cath DC, Van Dyck R, Boomsma DI. The co-morbidity of anxiety and depression in the perspective of genetic epidemiology: A review of twin and family studies. *Psychological Medicine*. 2005; 35:611–624. [PubMed: 15918338]
- Miller MW, Fogler JM, Wolf EJ, Kaloupek DG, Keane TM. The internalizing and externalizing structure of psychiatric comorbidity in combat veterans. *Journal of Traumatic Stress*. 2008; 21:58–65. [PubMed: 18302181]
- Moffitt TE, Caspi A, Taylor A, Kokaua J, Milne BJ, Polanczyk G, et al. How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. *Psychological Medicine*. 2010; 40:899–909. [PubMed: 19719899]
- Moffitt TE, Harrington H, Caspi A, Kim-Cohen J, Goldenberg D, Gregory AM, Poulton R. Depression and generalized anxiety disorder: Cumulative and sequential comorbidity in a birth cohort followed prospectively to age 32 years. *Archives of General Psychiatry*. 2007; 64:651–660. [PubMed: 17548747]

- Muthén, LK.; Muthén, BO. Mplus user's guide. 5th ed.. Los Angeles, CA: Muthén and Muthén; 1998–2007.
- Myung IJ. The importance of complexity in model selection. *Journal of Mathematical Psychology*. 2000; 44:190–204. [PubMed: 10733864]
- Orvaschel H, Lewinsohn PM, Seeley JR. Continuity of psychopathology in a community sample of adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1995; 34:1525–1535. [PubMed: 8543521]
- Orvaschel H, Puig-Antich J, Chambers WJ, Tabrizi MA, Johnson R. Retrospective assessment of prepubertal major depression with the Kiddie-SADS-E. *Journal of the American Academy of Child Psychiatry*. 1982; 21:392–397. [PubMed: 7119313]
- Procidano ME, Heller K. Measures of perceived social support from friends and from family. *American Journal of Community Psychology*. 1983; 11:1–24. [PubMed: 6837532]
- Rapaport MH, Clary C, Fayyad R, Endicott J. Quality-of-life impairment in depressive and anxiety disorders. *American Journal of Psychiatry*. 2005; 162:1171–1178. [PubMed: 15930066]
- Rohde P, Lewinsohn PM, Seeley JR. Comorbidity of unipolar depression: II. Comorbidity with other mental disorders in adolescents and adults. *Journal of Abnormal Psychology*. 1991; 100:214–222. [PubMed: 2040773]
- Rohde P, Lewinsohn PM, Seeley JR, Klein DN, Andrews JA, Small JW. Psychosocial functioning of adults who experienced substance use disorders as adolescents. *Psychology of Addictive Behavior*. 2007; 21:155–164.
- Silberg JL, Rutter M, Eaves L. Genetic and environmental influences on the temporal association between earlier anxiety and later depression in girls. *Biological Psychiatry*. 2001; 49:1040–1049. [PubMed: 11430845]
- Slade T, Watson D. The structure of common DSM-IV and ICD-10 mental disorders in the Australian general population. *Psychological Medicine*. 2006; 36:1593–1600. [PubMed: 16882356]
- South SC, Krueger RF. Marital quality moderates genetic and environmental influences on the internalizing spectrum. *Journal of Abnormal Psychology*. 2008; 117:826–837. [PubMed: 19025229]
- Verduin TL, Kendall PC. Differential occurrence of comorbidity within childhood anxiety disorders. *Journal of Clinical Child and Adolescent Psychology*. 2003; 32:290–295. [PubMed: 12679288]
- Vollebergh WAM, Iedema J, Bijl RV, de Graaf R, Smit F, Ormel J. The structure and stability of common mental disorders: The NEMESIS study. *Archives of General Psychiatry*. 2001; 58:597–603. [PubMed: 11386990]
- Watson D. Rethinking mood and anxiety disorders: A quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology*. 2005; 114:522–536. [PubMed: 16351375]
- Watson D. Differentiating the mood and anxiety disorders: A quadripartite model. *Annual Review of Clinical Psychology*. 2009; 5:221–247.
- Watson D, O'Hara MW, Stuart S. Hierarchical structures of affect and psychopathology and their implications for the classification of emotional disorders. *Depression and Anxiety*. 2008; 25:282–288. [PubMed: 18415951]
- Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Archives of General Psychiatry*. 1976; 33:1111–1115. [PubMed: 962494]
- Wells JE, Horwood LJ. How accurate is recall of key symptoms of depression? A comparison of recall and longitudinal reports. *Psychological Medicine*. 2004; 34:1001–1011. [PubMed: 15554571]
- Wittchen H-U, Beesdo-Baum K, Gloster AT, Höfler M, Klotzsch J, Lieb R, et al. The structure of mental disorders re-examined: Is it developmentally stable and robust against additions? *International Journal of Methods in Psychiatric Research*. 2009; 18:189–203. [PubMed: 20033884]
- Yang C-C. Evaluating latent class analyses in qualitative phenotype identification. *Computational Statistics & Data Analysis*. 2006; 50:1090–1104.
- Yu, C-Y. Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. Los Angeles: Unpublished doctoral dissertation, University of California; 2002.
- Zinbarg R, Barlow DH. Structure of anxiety and the anxiety disorders: A hierarchical model. *Journal of Abnormal Psychology*. 1996; 105:181–193. [PubMed: 8722999]

Internalizing Model 1 (INT1): Single-Factor *Internalizing Model 2a (INT2a): DSM-IV Model*

Internalizing Model 2b (INT2b): Krueger–Watson model

Figure 1. Path diagram for Single-Factor, *DSM-IV*, and Krueger–Watson models (Internalizing Models INT1, INT2a, and INT2b, respectively). All parameter estimates are standardized and significant at $p < .01$. PTSD = post-traumatic stress disorder.

Table 1

Tetrachoric Correlations (above Diagonal with Standard Errors) and Odds Ratios (below Diagonal with 95% Confidence Intervals) between Internalizing Disorder Diagnostic Categories

	MDD	DYS	GAD	PTSD	PHOB	SOC	PAN	OCD	SAD
MDD	—	0.63 (0.09)	0.49 (0.12)	0.37 (0.08)	0.32 (0.10)	0.35 (0.10)	0.47 (0.08)	0.34 (0.16)	0.48 (0.10)
DYS	23.2 (4.3–124.5)	—	0.01 (0.22)	0.47 (0.10)	0.33 (0.14)	–0.10 (0.21)	0.48 (0.10)	0.31 (0.20)	0.32 (0.14)
GAD	10.2 (1.9–55.9)	1.0 (0.1–11.4)	—	0.30 (0.15)	0.03 (0.23)	0.45 (0.14)	0.08 (0.18)	0.44 (0.20)	0.02 (0.23)
PTSD	3.7 (2.0–7.1)	6.6 (3.0–14.7)	3.8 (1.2–12.7)	—	0.20 (0.13)	0.25 (0.13)	0.43 (0.10)	0.49 (0.15)	0.36 (0.12)
PHOB	3.5 (1.5–8.4)	4.4 (1.5–13.0)	1.2 (0.1–13.2)	2.4 (0.8–6.9)	—	0.25 (0.13)	0.43 (0.13)	0.49 (0.15)	0.29 (0.14)
SOC	4.1 (1.6–10.2)	0.6 (0.1–6.1)	8.3 (2.4–28.3)	2.9 (1.1–7.9)	6.4 (2.2–18.0)	—	0.37 (0.12)	0.33 (0.20)	0.42 (0.13)
PAN	6.1 (2.7–13.6)	7.0 (3.1–16.0)	1.5 (0.3–8.3)	5.2 (2.6–10.7)	3.4 (1.2–9.3)	4.8 (1.9–12.1)	—	0.44 (0.16)	0.29 (0.13)
OCD	4.8 (0.8–29.2)	4.8 (0.8–29.9)	10.4 (1.6–67.1)	10.1 (2.5–40.3)	14.7 (3.4–63.8)	5.5 (0.9–34.7)	8.1 (1.9–34.5)	—	0.54 (0.14)
SAD	7.8 (2.6–23.5)	4.1 (1.4–12.0)	1.1 (0.1–12.3)	4.3 (1.8–10.5)	3.8 (1.2–12.1)	2.8 (0.8–10.2)	6.8 (2.0–15.7)	5.2 (0.8–32.2)	—

Note.

MDD = major depressive disorder, DYS = dysthymic disorder, GAD = generalized anxiety disorder, PTSD = posttraumatic stress disorder, PHOB = specific or simple phobias, SOC = social phobia, PAN = panic disorder, OCD = obsessive-compulsive disorder, SAD = separation anxiety disorder.

Table 2

Internalizing Disorders: Model Fit Indices and Selection Criteria

Model	Indicators of Model Fit						Model Selection Criteria		
	X ² value	X ² df	X ² p-value	CFI	TLI	RMSEA	WRMR	Adj. BIC	AIC
<i>7 Variables:</i>									
1-Factor	21.57	12	0.043	0.957	0.938	0.031	0.874	2768.94	2747.53
DSM	19.36	11	0.055	0.962	0.941	0.031	0.832	2769.95	2747.01
Krueger-Watson	19.37	11	0.055	0.962	0.941	0.031	0.828	2769.78	2746.85
<i>9 Variables:</i>									
1-Factor	27.64	19	0.091	0.966	0.960	0.024	0.838	3098.38	3070.86
DSM	23.96	18	0.156	0.976	0.971	0.020	0.783	3098.48	3069.43
Krueger-Watson	24.70	18	0.134	0.973	0.967	0.021	0.795	3098.88	3069.83

Note.

CFI=comparative fit index; TLI = Tucker-Lewis fit index; RMSEA = root mean square error of approximation; WRMR = root mean square residual; BIC = Bayesian information criterion; AIC = Akaike information criterion. The “best” of the competing models as implied by the AIC statistic is indicated by a value of 0.00 under the AIC column.

Table 3
Measurement Model Latent Factors as Predictors of T₃ Family Density of Psychiatric Disorders: Unadjusted Analyses

Psychiatric Disorders Among First-Degree Relatives	1-Factor Model (INT1)			2-Factor DSM-IV Model (INT2a)			2-Factor Krueger-Watson model (INT2b)		
	Coeff	p	R ²	Internalizing Disorders	Mood Disorders	Anxiety Disorders	Distress Disorders	Fear Disorders	R ²
Major depressive disorder	.24	<.001	.058	.24	<.001	.057	.24	<.001	.055
Dysthymia	.12	<.001	.014	.13	<.001	.017	.12	<.001	.013
Specific phobia	.09	.027	.008	.09	.028	.008	.09	.029	.008
Social phobia	.04	.333	.001	.03	.403	.001	.03	.362	.002
Panic disorder	.12	.003	.014	.09	.015	.008	.13	.005	.017
Substance use disorder	.18	<.001	.034	.18	<.001	.032	.18	<.001	.034

Note.

All regression coefficients are standardized; significant ($p < .05$) values are bolded. INT = internalizing model; Coeff. = coefficient.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript