



Published in final edited form as:

J Abnorm Psychol. 2013 February ; 122(1): 86–92. doi:10.1037/a0029598.

The Structure and Predictive Validity of the Internalizing Disorders

Nicholas R. Eaton,

Department of Psychology, University of Minnesota

Robert F. Krueger,

Department of Psychology, University of Minnesota

Katherine M. Keyes,

Department of Epidemiology, Mailman School of Public Health, Columbia University

Melanie Wall,

Department of Psychiatry, Columbia University College of Physicians and Surgeons, and Department of Biostatistics, Mailman School of Public Health, Columbia University

Deborah S. Hasin,

Department of Psychiatry, Columbia University College of Physicians and Surgeons, and Department of Epidemiology, Mailman School of Public Health, Columbia University

Kristian E. Markon,

Department of Psychology, University of Iowa

Andrew E. Skodol, and

Department of Psychiatry, University of Arizona College of Medicine and Columbia University College of Physicians and Surgeons

Bridget F. Grant

Laboratory of Epidemiology and Biometry, Division of Intramural Clinical and Biological Research, National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD.

Abstract

Multivariate comorbidity research indicates mood and anxiety (internalizing) disorders share one or more common liabilities, but categorical, dimensional, and hybrid accounts of these liabilities have not been directly compared. We modeled seven internalizing disorders in a nationally representative sample of 43,093 individuals via confirmatory factor, latent class, exploratory factor mixture, and exploratory structural equation modeling analyses. A two-dimensional (distress-fear) liability structure fit best and replicated across gender, assessment waves, and lifetime/12-month diagnoses. These liabilities, not disorder-specific variation, predicted future internalizing pathology, suicide attempts, angina, and ulcer.

Keywords

internalizing; dimensions; classes; psychopathology; depression; anxiety

Multivariate research indicates a latent general liability—*internalizing*—accounts for higher-than-chance levels of mood and anxiety disorder comorbidity (Krueger & Markon, 2006; Watson, 2005). Internalizing liability is largely stable over time (Eaton, Krueger, & Oltmanns, 2011; Fergusson, Horwood, & Boden, 2006; Krueger et al., 1998; Vollebergh et al., 2001), explains group-level differences in prevalence rates (Eaton et al., 2012), and serves as the primary pathway for the continuity/development of lifetime internalizing disorders (Kessler et al., 2011). Internalizing liability appears to reflect a coherent heritable risk structure for many putatively distinct disorders, highlighting its potential etiological informativeness (Kendler et al., 2011). Such results have prompted some to suggest the organizational “metastructure” of the upcoming fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (*DSM-5*) should group internalizing disorders together (vs. “mood disorders” and “anxiety disorders”; Andrews et al., 2009; Regier et al., 2011).

Internalizing Liability Structure

The nature of internalizing liability has not been completely characterized, despite empirical and nosological efforts highlighting its importance. Many studies assume internalizing liability is continuous: either unidimensional (Eaton, Krueger, & Oltmanns, 2011; Fergusson, Horwood, & Boden, 2006; Krueger et al., 1998) or two-dimensional (Eaton et al., 2012; Krueger, 1999; Slade & Watson, 2006; Vollebergh et al., 2001). In the two-dimensional conceptualization, a higher-order internalizing dimension subsumes two low-order dimensions of distress (liability to depression, dysthymia, and generalized anxiety disorder [GAD]) and fear (liability to social phobia, specific phobia, and panic disorder). These competing dimensional internalizing models have rarely been compared directly, yielding equivocal results (Seeley et al., 2011). Thus, the dimensional structure of internalizing liability is unclear.

Some studies have alternatively conceptualized internalizing as a set of liability classes (W. W. Eaton et al., 1989; Vaidyanathan, Patrick, & Iacono, 2011). For instance, Ferdinand and colleagues (2005) identified two classes of individuals who experienced severe anxiety symptoms: one class with severe depressive symptoms and another with mild depressive symptoms. To our knowledge, no studies have directly compared dimensional and class conceptualizations.

A third compelling conceptualization of internalizing liability posits a hybrid dimension-class structure. In a hybrid structure, each individual would be a member of a liability class, and, within each class, a dimension would allow class members to vary in liability severity. For instance, a high mood/low anxiety liability class might emerge, and a severity dimension would allow individuals in the class to show varying levels of (high) mood pathology and (low) anxiety pathology. To our knowledge, no studies have tested hybrid internalizing liability structures nor compared hybrid possibilities to dimension- and class-only model. This is in contrast to dimension/class/hybrid comparisons of *externalizing*—the liability for antisocial behaviors and substance use disorders (Walton, Ormel, & Krueger, 2011).

These issues are significant given the relevance of the internalizing liability spectrum for nosology (*DSM-5*), assessment, and treatment. Further, different structures support different inferences about lifetime internalizing comorbidity and continuity/development (e.g., whether continuity/development represents trait-like stability of liability dimensions or transitions between liability classes). Simply assuming a liability structure may produce misleading results.

Predictive Validity of Liability Versus Disorder-Specific Variation

An accurate liability structure characterization is the first step in determining how internalizing pathology predicts future outcomes of interest. The liability represents the *common* variance across internalizing disorders; each disorder also has *specific* (unique) variance (e.g., what makes depression different from GAD). Internalizing disorders have been prospectively linked to mental disorders (Kessler et al., 2011), suicide, and important health outcomes such as angina and ulcer (Folks & Kinney, 1992; Smith et al., 1996). (Although most ulcer cases are associated with *Helicobacter pylori* infection, only a small percentage of infected patients develop ulcers, suggesting roles for anxiety and neuroticism [Goodwin & Stein, 2003].). It is unclear whether these associations reflect common liability, or disorder-specific, variation.

We addressed these gaps in the literature by comparing dimensional, class, and hybrid models. We conducted these analyses separately by gender and on each of two data waves. We then investigated the links of internalizing liability/disorders with future internalizing pathology. Finally, we examined the prospective predictive validity of internalizing liability versus disorder-specific variance for three health outcomes.

Method

Participants

We analyzed two longitudinal data waves from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC), a representative sample of the civilian, noninstitutionalized U.S. population: Wave-1 ($n = 43,093$; 2001–2002; response rate: 81% of those eligible) and Wave-2 ($n = 34,653$; 2004–2005; 86.7% of eligible original sample; 70.2% cumulative response rate). Wave-1 was 57% ($n = 24,575$) female; ages ranged from 18–98 years. Hispanic/Latino, Black, and young adults were oversampled. Design variables ensured age, racial/ethnic, and gender representativeness of the U.S. based on the 2000 Census. The research protocol, including informed consent, received full ethical review and approval from the U.S. Census Bureau and Office of Management and Budget.

Assessment

The Alcohol Use Disorder and Associated Disabilities Interview Schedule—*DSM-IV* Version (AUDADIS-IV; Grant et al., 2003), a structured interview designed for administration by experienced lay interviewers, assessed *DSM-IV* disorders. We examined lifetime and 12-month major depressive disorder, dysthymic disorder, GAD, panic disorder (with agoraphobia), social phobia, specific phobia, and bipolar I disorder diagnoses. AUDADIS-IV diagnostic reliabilities were: depression ($\kappa = .65$), dysthymia (.58), GAD (.42), panic (.42), social phobia (.46), specific phobia (.48), and bipolar (.59; Grant et al., 2003, 2005; Hasin et al., 2005). Test-retest reliabilities are similar to those for other structured interviews (e.g., the DIS, CIDI; Wittchen, 1994). The AUDADIS-IV has marked advantages over some interviews, including assessment of clinically significant distress and impairment (Hasin et al., 2005). Although rarely included in early structural models, recent research indicates bipolar disorder shows notable relations with internalizing liability (Kessler et al., 2011; Vaidyanathan, Patrick, & Iacono, 2011; Watson, 2005).

Statistical Analyses

Analyses were conducted in Mplus version 6 (Muthén & Muthén, 2011) using the default delta parameterization for structural analyses and theta parameterization for validity analyses. We used a robust maximum likelihood estimator (MLR), treated diagnoses as categorical, and included the NESARC's design variables. To evaluate model fit with MLR,

we used the Bayesian information criterion (BIC), which balances model fit and parsimony. Smaller BIC values indicate better performance. BIC allows for the comparison of non-nested models and outperforms other tests—including the Vuong-Lo-Mendell-Rubin test and the bootstrapped likelihood ratio test (not available with design features)—and fit indices in large samples (Nylund, Asparouhov, & Muthén, 2007). BIC differences of 10 between two models indicate posterior odds of 150:1 favoring the model selected by BIC, providing “very strong” evidence for it (Raftery, 1995). For multigroup analyses, numerical integration is not available, which precluded maximum likelihood estimation and BIC computation; thus, we used a weighted least squares (WLSMV) estimator and evaluated the comparative fit index (CFI), Tucker-Lewis index (TLI), and root mean squared error of approximation (RMSEA). CFI/TLI values $> .95$ and RMSEA values $< .06$ suggest good model fit (Hu & Bentler, 1999).

Confirmatory factor analysis

Confirmatory factor analysis (CFA) models latent dimensions underlying multivariate diagnostic comorbidity (Brown, 2006). CFA requires parameterization of which diagnoses load on which factor(s). CFA analyses were guided by the literature and exploratory factor analyses (EFA; not reported for brevity). In two-factor CFAs, depression, dysthymia, GAD, and bipolar disorder indicated distress liability; social phobia, specific phobia, and panic disorder indicated fear liability. That said, EFAs indicated consistent nontrivial cross-loadings for GAD and bipolar disorder but no other disorders. To investigate the possible improvement/detriment in model fit with additions of cross-loadings, we fit two-factor CFAs with no cross-loadings and with cross-loadings for (1) GAD, (2) bipolar, and (3) GAD and bipolar.

Latent class analysis

Latent class analysis (LCA) posits the existence of a finite number of mutually exclusive groups (classes) of individuals defined by patterns of disorder endorsement (McCutcheon, 1987). Individuals are assigned to the most probable class based on their endorsement patterns. LCA assumes “conditional independence,” requiring comorbidity be completely accounted for by class membership (e.g., no within-class diagnosis correlations, all class members have the same liability).

Exploratory factor mixture analysis

Exploratory factor mixture analysis (EFMA) is a hybrid class-dimension approach (Muthén, 2006; Muthén & Muthén, 2011). EFMA posits a set of latent classes and one or more factors. This allows for severity differences among class members and permits all diagnoses to cross-load. We fit EFMA models up to the number of classes identified by LCA.¹

Exploratory structural equation modeling

Exploratory structural equation modeling (ESEM) combines EFA with structural equation modeling (Asparouhov & Muthén, 2009). All diagnoses load on all factors, which are rotated. Factors can then be linked to other variables via structural equations.

Predictive validity

Our longitudinal design allowed us to test the prospective predictive validity of latent internalizing liability versus disorder-specific variation. To explore these issues, we

¹We also tested two hybrid latent class factor analytic models, allowing for non-parametric factor distributions. These models, not reported for brevity, provided inferior fit to the optimal two-factor models. See models FMM-1 and FMM-2 in Clark and colleagues (2012) for more information. Results are available from the corresponding author.

predicted future internalizing disorders and three health outcomes from Wave-2 that have previously been linked to internalizing psychopathology: lifetime history of at least one suicide attempt, past-year angina pectoris or chest pain, and past-year stomach ulcer.

Results

Internalizing Structure

Wave-1—We compared latent dimensional (CFA), class (LCA), and hybrid (EFMA) models of the seven internalizing disorders separately by gender (see Table 1). The best fitting model in both genders was a two-factor (distress and fear) CFA where GAD and bipolar cross-loaded. Rather than modeling the multitude of remaining cross-loading combinations, we modeled internalizing liability via ESEM, allowing all disorders to cross-load. A single ESEM, fit simultaneously to women and men in multigroup analysis, provided very close fit (CFI = .995; TLI = .992; RMSEA = .014). An ESEM fit to the total Wave-1 sample provided excellent fit to the data (CFI = .999; TLI = .998; RMSEA = .007; loadings in Table 3).

Wave-2—We reconducted the analyses above using 12-month diagnoses from Wave-2. Two-dimensional models provided the best fit (see Table 2). In women, a two-dimensional model with GAD and bipolar cross-loadings was optimal; in men, a twodimensional model without cross-loadings was optimal. Multi-Group ESEM analyses indicated excellent fit when the same ESEM was parameterized in each gender (CFI = .998; TLI = .996; RMSEA = .007). An ESEM in the total Wave-2 sample provided excellent fit (CFI = .999; TLI = .998; RMSEA = .006; loadings in Table 3).

Prediction of Future Internalizing

To examine prospective prediction of Wave-2 disorders and liabilities by Wave-1 disorders and liabilities, we conducted an ESEM using the full Wave-1 and Wave-2 samples, invariant parameters across waves, correlated factors within waves, and regressions of Wave-2 distress and fear liabilities on both Wave-1 liabilities. The model provided excellent fit (CFI = .990, TLI = .987, RMSEA = .010) and was used for further analyses. Wave-2 distress was significantly predicted by Wave-1 distress ($p < .001$) but not Wave-1 fear ($p = .213$); Wave-2 fear was significantly predicted by Wave-1 fear ($p < .001$) but not Wave-1 distress ($p = .658$). Overall, 39.1% of Wave-2 distress variance, and 52.1% of Wave-2 fear variance, were accounted for.

We tested whether disorder-specific correlated residuals (e.g., Wave-1 depression disorder-specific variance correlated with Wave-2 depression disorder-specific variance) predicted future internalizing disorders. We used an effect size benchmark of at least $r = .316$ (indicating 10% of a disorder's Wave-2 residual variance was accounted for by Wave-1 residual variance). Only social phobia met our benchmark ($r = .417$; accounting for 17.4% of the variation). No other disorder closely approached the benchmark: depression <0.01%, dysthymia 6.81%, GAD 1.35%, bipolar 8.64%, specific phobia 7.73%, and panic disorder 0.09%. In contrast, the liability factors accounted for between 48.0% (specific phobia) and 79.6% (depression) of disorder variance. We therefore included a correlated residual for social phobia in the final model (see Figure 1).

Predictive Validity

Because latent liability dimensions, rather than disorder-specific variance, were the primary pathways for internalizing disorders over time, we investigated whether this pattern would replicate with prospective prediction of other important health outcomes. Wave-2 suicide attempt history, angina, and ulcer were regressed on Wave-1 distress and fear liability

dimensions. Distress significantly predicted suicide attempt history ($\beta = .566, p < .001$), but fear did not ($\beta = .021, p = .684$), and together accounted for 33.7% of suicide variance. Distress ($\beta = .169, p < .001$) and fear ($\beta = .105, p = .009$) significantly predicted angina, accounting for 6.5% of angina variance. Distress ($\beta = .196, p < .001$) and fear ($\beta = .169, p < .001$) significantly predicted ulcer, accounting for 6.0% of ulcer variance. We relaxed our benchmark for determining the meaningfulness of the relations between disorder-specific variance (residuals) and the residual variance in suicide attempt history, angina, and ulcer to 5%, because prediction below this mark would likely not hold much clinical meaning and be inferior to the liability factors. No disorder-specific variance reached this benchmark. The best disorder-specific variance predictors were: bipolar for suicide attempt history (accounting for 1.0%), depression for angina (1.8%), and panic disorder for ulcer (3.5%).

Discussion

Internalizing Structure

Our primary goal was to compare various possible structures of internalizing disorder liability in a large, nationally representative sample. Internalizing liability was best conceptualized as dimensional (vs. classes or class-dimension hybrid). Findings replicated in women and in men, across two assessment waves, and in lifetime and 12-month diagnoses. These replications provide evidence for the fundamentally dimensional nature of internalizing liability.

In all analyses, a two-dimension liability model was optimal. This finding resolves the outstanding question of whether internalizing liability is best conceptualized as a unitary dimension or as correlated distress and fear subdimensions. Distress and fear liabilities were critically important to capturing internalizing disorder comorbidity. Together, they accounted for between 48% and 80% of internalizing disorders' diagnostic variance.

Our findings clarify why different research teams may have had mixed results regarding the relative superiority of one- or two-dimension models: the CFA imposition of simple structure. Models with GAD and bipolar cross-loadings fit better than simple structure models in three of four cases, highlighting the importance of cross-loadings. This indicated that some disorders reflected liability to *both* distress and fear. These results support the utility of approaches such as ESEM, which have many strengths of CFA without CFA's stricter limitations on the number of permissible cross-loadings.

Finally, our results extend internalizing liability to include bipolar disorder (see also Kessler et al., 2011; Vaidyanathan, Patrick, & Iacono, 2011). Nearly 50% of bipolar's diagnostic variance was accounted for by internalizing liability. This percentage was notably lower than that of most other internalizing disorders, but higher than that of specific phobia, suggesting other liabilities may be at play as well (Watson, 2005).

Prediction of Future Outcomes

A second goal was to determine how internalizing liability and disorder-specific variation fared in prospective prediction of important outcomes: internalizing liability/disorders, suicide attempts, angina, and ulcer.

Internalizing outcomes

The prediction of future internalizing liability/disorders requires a clear conceptualization of the nature of this liability to determine the most appropriate statistical approach and pathway for potential liability continuity/development. For instance, if liability classes composed the variation in internalizing liability, researchers would be required to use methods that reflect

the likelihood of transitioning between liability classes over time. Because our results indicated internalizing liability was dimensional, trait-like stability analyses were an appropriate means to gauge internalizing liability continuity/development over time.

On the point of continuity/development mechanisms, several findings emerged. First, the liability factors were notably stable over time, indicated by 39% of Wave-2 distress variance, and 51% of Wave-2 fear variance, being accounted for by Wave-1 liabilities. Interestingly, Wave-2 distress and fear were not significantly predicted by Wave-1 fear and distress, respectively. Given our very large sample size, this nonsignificance is striking. These results highlight that distress and fear liabilities, although significantly correlated, are unique and stable pathways for the continuity/development of internalizing psychopathology.

What role does disorder-specific variation play in internalizing psychopathology continuity/development over time? The answer appears to be very little. Only social phobia showed a meaningful correlated residual across waves, accounting for 17% of diagnostic variance. By comparison, distress and fear liabilities accounted for 64% of the Wave-1 social phobia variance and 65% of the Wave-2 variance. Thus, even in this case most favorable to disorder-specific variation, the liability factors accounted for around 3.75 times more variance in a particular disorder than did its previous disorder-specific variation.

Suicide/Health outcomes

Our third goal was to examine the predictive utility of the liability dimensions versus disorder-specific variation for noninternalizing outcomes. Again, prospective prediction of future important health outcomes—suicide attempt history, angina, and ulcer—was better through liability factors rather than disorder-specific variation. While angina and ulcer were predicted weakly by both liability dimensions and disorders, liability dimensions predicted these outcomes much better. In the case of suicide attempts, the difference was marked, with Wave-1 liability dimensions capturing a great deal (34%) of Wave-2 suicide attempt history, while Wave-1 disorder-specific variation captured between 0% and 1%. Combining these results with those of predicting future internalizing disorders above, our findings unequivocally supported the notion that latent internalizing liability dimensions, not disorder-specific variation, accounted for links to important outcomes.

Implications

Our results support the recent call (Andrews et al., 2009; Regier et al., 2011) for the *DSM-5* metastructure to place internalizing disorders into a single group (albeit with distress and fear subheadings), rather than into putatively distinct mood and anxiety disorders chapters. If the majority of variance of most of these disorders is accounted for by shared liabilities to distress and fear, however, we must ask if disorder-specific variation contributes meaningfully to understanding psychopathology. In this study, the *unique* piece of each disorder did not predict future disorders or health outcomes meaningfully—liability factors did. Etiological (Kendler et al., 2011; Kessler et al., 2011) and intervention-oriented (Barlow et al., 2011; Goldberg et al., 2011) research has also supported the importance of these liability factors. If, in accumulated future studies, disorder-specific variation does not contribute meaningfully to understanding and ameliorating internalizing psychopathology, it must be characterized as unimportant (error) variance. Are separate diagnoses of these highly comorbid internalizing disorders truly of benefit to mental health efforts or are they primarily indicators of constructs such as broad liabilities? At minimum, our results highlight the pitfalls of reifying internalizing disorder diagnoses (Hyman, 2010). Further, they suggest that clinicians might consider novel treatment modalities that target internalizing liability (e.g., Barlow et al., 2011).

Limitations

This study is not without its limitations. First, it considered seven internalizing disorders, and inclusion of additional disorders/syndromes, including internalizing, externalizing, and psychotic disorders, would be valuable in extending the findings. Second, our use of lifetime diagnoses required retrospection, which could be subject to memory bias; however, our results replicated using 12-month diagnoses. Third, because of space limitations, we did not test factorial invariance formally across gender, wave, and diagnosis. Fourth, AUDADIS-IV reliabilities ranged across diagnoses, with some diagnoses having lower reliabilities than was preferable; in addition, we chose to model both lifetime and 12-month diagnoses for analytic and interpretive purposes, but this yields different prevalence rates and reliabilities across waves. Both of these limitations could have impacted our results. Finally, this study focused on *diagnostic* comorbidity rather than symptom-level or -count analyses, which address a related, but different, set of questions.

Acknowledgments

This work was supported by the National Institute on Alcohol Abuse and Alcoholism grants U01AA018111, K05AA014223, and the National Institute on Drug Abuse grants R01DA018652, and F31DA026689. NESARC was sponsored by National Institute on Alcohol Abuse and Alcoholism with support from National Institute on Drug Abuse. The views and opinions expressed in this report are those of the authors and do not necessarily represent the views of the sponsoring agencies or the U.S. government.

References

- Andrews G, Goldberg DP, Krueger RF, Carpenter WT Jr, Hyman SE, Sachdev P, Pine DS. Exploring the feasibility of a meta-structure for DSM-V and ICD-11: Could it improve utility and validity? *Psychological Medicine*. 2009; 39:1993–2000. doi: 10.1017/S0033291709990250. [PubMed: 19796425]
- Asparouhov T, Muthén B. Exploratory structural equation modeling. *Structural Equation Modeling*. 2009; 16:397–438. doi:10.1080/10705510903008204.
- Barlow, DH.; Farchione, TJ.; Fairholme, CP.; Ellard, KK.; Boisseau, CL.; Allen, LB.; Ehrenreich-May, J. Unified protocol for transdiagnostic treatment of emotional disorders: Therapist guide. Oxford University Press; New York, NY: 2011.
- Brown, TA. Confirmatory factor analysis for applied research. Guilford Press; New York, NY: 2006.
- Clark, SL.; Muthén, B.; Kaprio, J.; D’Onofrio, BM.; Viken, R.; Rose, RJ.; Smalley, SL. Models and strategies for factor mixture analysis: Two examples concerning the structure underlying psychological disorders. 2012. Manuscript submitted for publication available at <http://statmodel.com/papers.shtml>
- Eaton NR, Keyes KM, Krueger RF, Balsis S, Skodol AE, Markon KE, Hasin DS. An invariant dimensional liability model of gender differences in mental disorder prevalence: Evidence from a national sample. *Journal of Abnormal Psychology*. 2012; 121:282–288. doi:10.1037/a0024780. [PubMed: 21842958]
- Eaton NR, Krueger RF, Oltmanns TF. Aging and the structure and long-term stability of the internalizing spectrum of personality and psychopathology. *Psychology and Aging*. 2011; 26:987–993. doi:10.1037/a0024406. [PubMed: 21728443]
- Eaton WW, McCutcheon A, Dryman A, Sorenson A. Latent class analysis of anxiety and depression. *Sociological Methods & Research*. 1989; 18:104–125. doi:10.1177/0049124189018001004.
- Ferdinand RF, de Nijs PFA, van Lier P, Verhulst FC. Latent class analysis of anxiety and depressive symptoms in referred adolescents. *Journal of Affective Disorders*. 2005; 88:299–306. doi:10.1016/j.jad.2005.08.004. [PubMed: 16182373]
- Fergusson DM, Horwood LJ, Boden JM. Structure of internalizing symptoms in early adulthood. *The British Journal of Psychiatry*. 2006; 189:540–546. doi:10.1192/bjp.bp.106.022384. [PubMed: 17139039]

- Folks DG, Kinney FC. The role of psychological factors in gastrointestinal conditions: A review pertinent to DSM-IV. *Psychosomatics: Journal of Consultation Liaison Psychiatry*. 1992; 33:257–270. doi: 10.1016/S0033-3182(92)71964-1.
- Goldberg, D.; Simms, LJ.; Gater, R.; Krueger, RF.; Regier, DA.; Narrow, WE.; Kuhl, EA.; Kupfer, DJ. The conceptual evolution of DSM-5. American Psychiatric Publishing, Inc; Arlington, VA: 2011. Integration of dimensional spectra for depression and anxiety into categorical diagnoses for general medical practice; p. 19-35.
- Goodwin RD, Stein MB. Peptic ulcer disease and neuroticism in the United States adult population. *Psychotherapy and Psychosomatics*. 2003; 72:10–15. doi:10.1159/000067184. [PubMed: 12466633]
- Grant BF, Dawson DA, Stinson FS, Chou PS, Kay W, Pickering R. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): Reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug and Alcohol Dependence*. 2003; 71:7–16. doi:10.1016/S0376-8716(03)00070-X. [PubMed: 12821201]
- Grant BF, Stinson FS, Hasin DS, Dawson DA, Chou SP, Ruan WJ, Huang B. Prevalence, correlates, and comorbidity of bipolar I disorder and axis I and II disorders: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry*. 2005; 66:1205–1215. doi:10.4088/JCP.v66n1001. [PubMed: 16259532]
- Hasin DS, Goodwin RD, Stinson FS, Grant BF. Epidemiology of major depressive disorder: Results from the National Epidemiologic Survey on Alcoholism and Related Conditions. *Archives of General Psychiatry*. 2005; 62:1097–1106. doi:10.1001/archpsyc.62.10.1097. [PubMed: 16203955]
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6:1–55. doi: 10.1080/10705519909540118.
- Hyman SE. The diagnosis of mental disorders: The problem of reification. *Annual Review of Clinical Psychology*. 2010; 6:155–179. doi: 10.1146/annurev.clinpsy.3.022806.091532.
- Kendler KS, Aggen SH, Knudsen GP, Røysamb E, Neale MC, Reichborn-Kjennerud T. The structure of genetic and environmental risk factors for syndromal and subsyndromal common DSM-IV axis I and all axis II disorders. *The American Journal of Psychiatry*. 2011; 168:29–39. doi:10.1176/appi.ajp.2010.10030340. [PubMed: 20952461]
- Kessler RC, Ormel J, Petukhova M, McLaughlin KA, Green JG, Russo LJ, Ustün TB. Development of lifetime comorbidity in the World Health Organization World Mental Health Surveys. *Archives of General Psychiatry*. 2011; 68:90–100. doi:10.1001/archgenpsychiatry.2010.180. [PubMed: 21199968]
- 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. doi:10.1037/0021-843X.107.2.216. [PubMed: 9604551]
- Krueger RF, Markon KE. Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology*. 2006; 2:111–133. doi: 10.1146/annurev.clinpsy.2.022305.095213.
- Krueger RF. The structure of common mental disorders. *Archives of General Psychiatry*. 1999; 56:921–926. doi:10.1001/archpsyc.56.10.921. [PubMed: 10530634]
- McCutcheon, AL. Latent class analysis. Sage Publications, Inc; Newbury Park, CA: 1987.
- Muthén B. Should substance use disorders be considered as categorical or dimensional? *Addiction*. 2006; 101:6–16. doi:10.1111/j.1360-0443.2006.01583.x. [PubMed: 16930156]
- Muthén, LK.; Muthén, BO. Mplus user's guide. 6th ed.. Author; Los Angeles, CA:
- Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*. 2007; 14:535–569. doi:10.1080/10705510701575396.
- Raftery AE. Bayesian model selection in social research. *Sociological Methodology*. 1995; 25:111–163. doi:10.2307/271063.
- Regier, DA.; Narrow, WE.; Kuhl, EA.; Kupfer, DJ., editors. The conceptual evolution of DSM-5. American Psychiatric Publishing, Inc; Arlington, VA: 2011.

- Seeley JR, Kosty DB, Farmer RF, Lewinsohn PM. The modeling of internalizing disorders on the basis of patterns of lifetime comorbidity: Associations with psychosocial functioning and psychiatric disorders among first-degree relatives. *Journal of Abnormal Psychology*. 2011; 120:308–321. doi:10.1037/a0022621. [PubMed: 21401216]
- 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. doi:10.1017/S0033291706008452. [PubMed: 16882356]
- Smith DF, Sternorff B, Røpcke G, Gustavsen EM, Hansen JR. Prevalence and severity of anxiety, depression and Type A behaviors in angina pectoris. *Scandinavian Journal of Psychology*. 1996; 37:249–258. doi:10.1111/j.1467-9450.1996.tb00657.x. [PubMed: 8856997]
- Vaidyanathan U, Patrick CJ, Iacono WG. Patterns of comorbidity among mental disorders: A person-centered approach. *Comprehensive Psychiatry*. 2011; 52:527–535. doi:10.1016/j.comppsy.2010.10.006. [PubMed: 21111407]
- Vollebergh WAM, Iedema J, Bijl RV, de Graaf R, Smit F, Ormel J. The structure and stability of common mental disorders. *Archives of General Psychiatry*. 2001; 58:597–603. doi:10.1001/archpsyc.58.6.597. [PubMed: 11386990]
- Walton KE, Ormel J, Krueger RF. The dimensional nature of externalizing behaviors in adolescence: Evidence from a direct comparison of categorical, dimensional, and hybrid models. *Journal of Abnormal Child Psychology*. 2011; 39:553–561. doi:10.1007/s10802-010-9478-y. [PubMed: 21210204]
- Watson D. Rethinking the mood and anxiety disorders: A quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology*. 2005; 114:522–536. doi:10.1037/0021-843X.114.4.522. [PubMed: 16351375]
- Wittchen H-U. Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): A critical review. *Journal of Psychiatric Research*. 1994; 28:57–84. doi:10.1016/0022-3956(94)90036-1. [PubMed: 8064641]

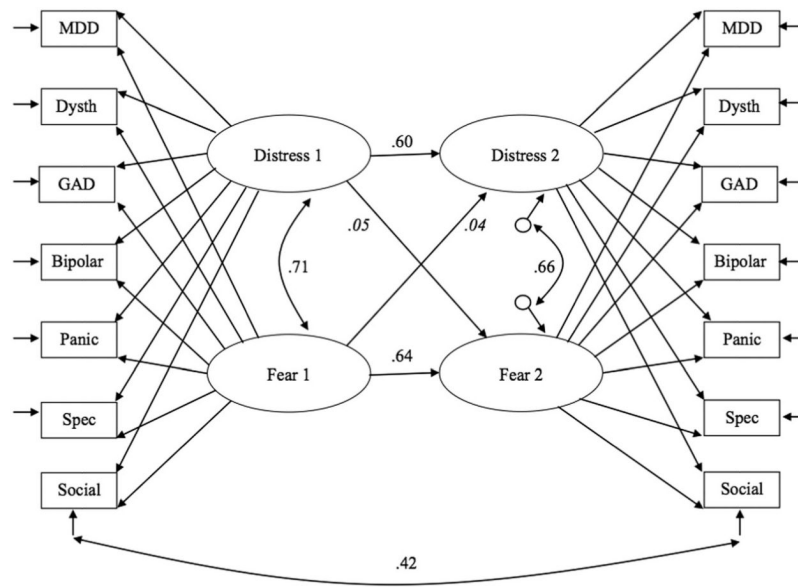


Figure 1. Structure and continuity/development of internalizing
 Values represent standardized coefficients, significant ($p < .001$) unless italicized. Wave-1 diagnoses at left and Wave-2 diagnoses at right. See Table 3 for loadings. MDD = major depressive disorder; Dysth = dysthymic disorder; GAD = generalized anxiety disorder; Spec = specific phobia. Short arrows indicate disorder-specific and residual variances. Long arrows connecting factors to disorders are loadings.

Table 1
Wave-1 Model Fits

	Women			Men		
	<i>LL</i>	<i>k</i>	BIC	<i>LL</i>	<i>k</i>	BIC
CFA						
1-f	-39,473.726	14	79,088.986	-20,956.659	14	42,050.888
2-f ¹	-39,243.365	15	78,638.372	-20,784.331	15	41,716.060
2-f ²	-39,211.432	16	78,584.615	-20,806.671	16	41,770.567
2-f ³	-39,239.157	16	78,640.066	-20,826.664	16	41,810.552
2-f ⁴	-39,184.476	17	78,540.813	-20,255.719	17	40,678.488
LCA						
1-c	-44,867.488	7	89,805.743	-23,785.132	7	47,639.050
2-c	-39,666.783	15	79,485.207	-20,992.054	15	42,131.506
3-c	-39,346.174	23	78,924.866	-20,789.956	23	41,805.921
4-c	-39,145.776	31	78,604.946	-20,706.558	31	41,717.736
5-c	-39,128.158	39	78,650.586	-20,694.366	39	41,771.966
EFMA						
2-c, 1-f	-39,135.081	29	78,563.337	-20,706.991	29	41,698.950
3-c, 1-f	-39,107.236	44	78,659.289	-20,685.525	44	41,803.415
4-c, 1-f	-39,097.900	59	78,792.260	-20,678.159	59	41,936.081
2-c, 2-f	-39,117.660	41	78,649.809	-20,685.964	41	41,774.815
3-c, 2-f	-39,090.439	62	78,807.667	-20,668.788	62	41,946.820
4-c, 2-f	-39,078.154	83	78,995.396	-20,659.646	83	42,134.891

Note. Optimal BIC values are bolded. f = factor(s); c = class(es); *LL* = log-likelihood; *k* = number of parameters. CFA subscripts:

- ¹No cross-loadings
- ²GAD cross-loading
- ³bipolar cross-loading
- ⁴GAD and bipolar cross-loadings

Table 2
Wave-2 Model Fits

	Women			Men		
	<i>LL</i>	<i>k</i>	BIC	<i>LL</i>	<i>k</i>	BIC
CFA						
1-f	-22,428.539	14	44,995.789	-10,120.117	14	20,374.443
2-f ¹	-22,343.517	15	44,835.653	-10,053.557	15	20,250.908
2-f ²	-22,336.262	16	44,831.051	-10,050.110	16	20,253.601
2-f ³	-22,334.909	16	44,828.345	-10,053.502	16	20,260.385
2-f ⁴	-22,317.35	17	44,803.135	-10,049.262	17	20,261.492
LCA						
1-c	-25,051.100	7	50,171.555	-11,460.735	7	22,988.574
2-c	-22,546.723	15	45,242.064	-10,139.881	15	20,423.556
3-c	-22,378.572	23	44,985.025	-10,058.199	23	20,336.882
4-c	-22,282.130	31	44,871.406	-10,028.789	31	20,354.753
5-c	-22,270.535	39	44,927.478	-10,018.032	39	20,409.929
EFMA						
2-c, 1-f	-22,281.485	29	44,850.300	-10,027.226	29	20,332.456
3-c, 1-f	-22,264.700	44	44,965.348	-10,006.757	44	20,435.312
4-c, 1-f	-22,259.478	59	45,103.524	-9,991.828	59	20,549.249
2-c, 2-f	-22,265.622	41	44,937.469	-10,009.302	41	20,411.643
3-c, 2-f	-22,255.839	62	45,125.971	-9,996.391	62	20,587.133
4-c, 2-f	-22,243.250	83	45,308.859	-9,983.135	83	20,761.935

Note. Optimal BIC values are bolded. f = factor(s); c = class(es); *LL* = log-likelihood; *k* = number of parameters. CFA subscripts:

¹No cross-loadings

²GAD cross-loading

³bipolar cross-loading

⁴GAD and bipolar cross-loadings.

ESEM Results

Table 3

	Wave-1		Wave-2		Total	
	1	2	1	2	1	2
Depression	.92	.01	.91	-.02	.88	.01
Dysthymia	.87	-.04	.74	.03	.92	-.10
Generalized anxiety	.54	.31	.56	.28	.57	.26
Bipolar	.53	.21	.49	.29	.52	.23
Social phobia	.11	.67	.09	.76	.12	.67
Specific phobia	-.02	.72	-.15	.78	-.17	.84
Panic disorder	.00	.90	.00	.92	.00	.88
Factor correlation		.64		.71		—

Note. Loadings > .4 are bolded. “Wave-1” and “Wave-2” represent solutions for each wave. “Total” represents total sample (Waves 1 and 2) solution. Figure 1 gives total factor correlations.