Received: 26 September 2017 | Revised: 10 January 2018 | Accepted: 10 February 2018

DOI: 10.1002/da.22740

RESEARCH ARTICLE

WILEY ADAA ANXIETY AND DEPRESSION ASSOCIATION OF AMERICA

Evaluating dimensional models of psychopathology in outpatients diagnosed with emotional disorders: A cautionary tale

Christopher C. Conway PhD[1] | **Timothy A. Brown PsyD**[2]

[1]Department of Psychological Sciences, College of William & Mary, Williamsburg, VA, USA

[2]Center for Anxiety and Related Disorders, Boston University, Boston, MA, USA

Correspondence
Christopher C. Conway, Department of Psychological Sciences, College of William & Mary, 540 Landrum Drive, Williamsburg, VA 23188, USA.
Email: conway@wm.edu

Funding information
National Institute of Mental Health, Grant/Award Number: MH039096

Background: Mental disorders cluster together systematically. Through factor analysis of disorder comorbidity, investigators are establishing the latent dimensions that underlie the development of related syndromes. However, these dimensions have not been validated across diverse patient samples, in which comorbidity patterns vary widely.

Method: We assessed 4,928 outpatients seeking treatment for emotional disorders with a semistructured diagnostic interview. This was the largest patient sample as yet for an evaluation of the latent structure of mental disorders. We compared several competing dimensional models of common mental disorders via confirmatory factor analysis.

Results: The hypothesized confirmatory factor model—anchored by internalizing and externalizing spectra—fit the diagnostic data poorly. Neither a one-factor model, reflecting a unitary liability to all mental disorders, nor a three-factor model, wherein the internalizing dimension bifurcated into distress and fear subfactors, fit appreciably better.

Conclusions: These data provide novel evidence that the internalizing and externalizing spectra are not structurally sound in all clinical contexts. We speculate about the causes of model misfit and advise additional research into the generalizability—with respect to sample, input data, developmental stage, and more—of dimensional models of mental disorder.

KEYWORDS
classification, comorbidity, emotional disorders, hierarchical taxonomy of psychopathology

1 | INTRODUCTION

Co-occurrence of psychiatric disorders is ubiquitous. It is also systematic, in that some conditions are more likely to cluster together than others (Kessler, Chiu, Demler, & Walters, 2005). In recent years, investigators have analyzed these patterns to uncover the latent traits that predispose to groups of disorders. Early studies in this area found that anxiety and depressive disorder diagnoses reflect an overarching internalizing spectrum, whereas antisocial and substance use disorders are indicators of an externalizing spectrum (Krueger, Caspi, Moffitt, & Silva, 1998; Krueger, 1999).

From this perspective, mental disorders are conceptualized as manifestations of pathological dimensions that transcend traditional disorder boundaries. A rapidly growing line of research supports that view. The internalizing and externalizing spectra emerge consistently across populations diverse in age, gender, and country of origin (Lahey, Krueger, Rathouz, Waldman, & Zald, 2016). Further, this original model has been extended to reveal additional dimensions that explain comorbidity. For instance, the internalizing spectrum bifurcates in some datasets into fear and distress subfactors, distinguishing vulnerability to phobic anxiety from depression and generalized anxiety (Clark & Watson, 2006).

This evolving dimensional system has been termed the hierarchical taxonomy of psychopathology (HiTOP[1]; Kotov et al., 2017). Evidence is building for not only the consistency of HiTOP traits across large-scale datasets, but also for concurrent and predictive validity at genetic, physiological, cognitive, and social levels of analysis (Forbes, Tackett, Markon, & Krueger, 2016; Krueger, Tackett, & MacDonald, 2016).

Based on these promising findings, many investigators expect HiTOP to compete with—and eventually supplant—*DSM* as a research and diagnostic tool. However, before HiTOP is disseminated widely, stronger evidence for validity in clinical settings is necessary. To date, the overwhelming majority of relevant data derive from epidemiological and community studies (see Lahey et al., 2016). One prominent exception involved the analysis of clinical and personality disorder comorbidity in a sample of 2,900 outpatients at a general mental health

clinic (Kotov et al., 2011). In that study, the internalizing and externalizing factors emerged alongside three others: somatoform (pain disorder and hypochondriasis), thought disorder (mania, psychosis, and Cluster A personality disorders), and antagonism (Cluster B personality disorders). A limited number of other investigations have examined the internalizing or externalizing dimensions in clinical or forensic samples (e.g., Miller, Fogler, Wolf, Kaloupek, & Keane, 2008; Sellbom, 2016; Wolf et al., 2010).

We argue that more data are needed to evaluate dimensional structures in diverse clinical settings. Variation in diagnostic prevalence and comorbidity rates across community, outpatient, and inpatient samples could lead to different conclusions about the natural organization of mental disorders. That is, as the distribution of disorders available to model (i.e., input data) expands or contracts, the HiTOP factor structure can, and frequently does, change (Kotov et al., 2011; Markon, 2010; Wright et al., 2013). Any such inconsistency in dimensional architecture across patient populations should be understood, we contend, prior to full-scale application of the HiTOP system.

In the present study, we modeled the structure of emotional and substance use disorders in a large outpatient sample. We assessed 4,928 patients presenting at an anxiety disorders clinic using a gold-standard diagnostic interview. Based on prior evidence (Kotov et al., 2017), we expected to find a two-factor model anchored by internalizing and externalizing traits to provide the best fit to our diagnostic correlations.

2 | MATERIALS AND METHODS

2.1 | Participants

Participants were 4,928 adults presenting for assessment and treatment at the Center for Anxiety and Related Disorders (CARD) at Boston University. As a community-based treatment center, CARD receives the vast majority of its patient admissions through referrals from primary care physicians, mental health professionals, and self-referrals. Thus, the breakdown of diagnoses in these 4,928 admissions is apt to be well representative of the patterns of emotional disorder psychopathology seen in treatment-seeking individuals, at least for the geographic region served by CARD. The majority of participants were women (60.2%), and the average age was 32.83 (*SD* = 11.65, range 18–87). The sample was predominantly Caucasian (87.5%; African-American = 4.4%; Asian = 4.8%; Hispanic = 2.6%; Other = 0.7%). The full sample was assessed for current disorders, and a subset of 2,655 patients was additionally assessed for lifetime (i.e., current and past) disorders.[2] The Boston University Charles River Campus institutional review board approved this project ("The Classification of Anxiety Disorders," Protocol #533E).

2.2 | Measures

All participants completed the Anxiety Disorders Interview Schedule for *DSM-IV* (ADIS-IV; Brown, DiNardo, & Barlow, 1994). The ADIS-IV is a semi-structured diagnostic interview designed to determine the presence of current *DSM-IV* emotional and substance use disorders. Diagnosticians were clinical psychologists and advanced clinical doctoral students. Before participating in the study, diagnosticians underwent extensive training and met strict certification criteria in the administration of the ADIS-IV (for a detailed description of these training and certification procedures, see Brown, Campbell, Lehman, Grisham, & Mancill, 2001). To establish consensus diagnoses and to guard against rater drift, all ADIS-IV intakes were presented in weekly staff meetings that entailed the presentation of interviewers' diagnoses and discussion of factors contributing to any diagnostic disagreements (in the case of reliability interviews). A study using a subset of the present study's sample of patients with current diagnoses found that the interrater reliability of ADIS-IV diagnoses was good-to excellent (κs = .67 to .86), except for dysthymia (κ = .31) (Brown et al., 2001). The diagnoses included in the present analyses were mood disorders (major depressive disorder, bipolar disorder, dysthymia), anxiety disorders (generalized anxiety disorder, social phobia, specific phobia, obsessive-compulsive disorder, posttraumatic stress disorder, and panic disorder), and substance use disorders (drug abuse/dependence and alcohol abuse/dependence).

2.3 | Statistical analysis

The raw data were analyzed in Mplus (version 7.11; Muthén & Muthén, 1998–2014) using the WLSMV estimator. Model fit was judged to be adequate when the comparative fit index (CFI) and Tucker-Lewis index (TLI) were .95 or above and when the root mean square error of approximation (RMSEA) was 0.06 or below (Brown, 2015). We also interpreted the Bayesian Information Criterion, lower values of which indicate better fit (Raftery, 1993). Model fit was additionally evaluated on the basis of the size and interpretability of the factor loadings.

We tested a series of confirmatory factor analysis (CFA) models for both current and lifetime diagnoses. First, we evaluated a unidimensional model that represented a general liability to any mental disorder. Second, in a two-factor model, the anxiety and mood disorders were indicators of an internalizing factor, whereas the substance use disorders were indicators of an externalizing factor. Third, in a three-factor model, the internalizing spectrum subdivided into fear (social phobia, specific phobia, panic disorder, and obsessive-compulsive disorder) and distress (generalized anxiety disorder, posttraumatic stress disorder, and mood disorders) subfactors.

3 | RESULTS

3.1 | Descriptive statistics

Table 1 presents current and lifetime disorder prevalence rates and tetrachoric correlations among disorders. Prevalence varied widely, with social phobia (46%) and generalized anxiety disorder (38%) among the most common current conditions and bipolar disorder (2%) and drug misuse (1%) among the least. Visual inspection of the correlation matrix suggested that current alcohol and drug use disorders were more strongly correlated with each other (r = .50)

TABLE 1 Prevalence and tetrachoric correlations of current and lifetime mental disorders

Condition	1	2	3	4	5	6	7	8	9	10	11	Current n	Current %
1. Generalized anxiety disorder	—	.146[c]	−.113[c]	.047	−.056[a]	.237[c]	.445[c]	.189[c]	.147[a]	.097	.126	1859	38
2. Social phobia	.207[a]	—	−.251[c]	−.096[c]	−.445[c]	.042	.200[c]	.265[c]	.070	.128[a]	.154[a]	2255	46
3. Specific phobia	−.037	−.161[a]	—	−.094[b]	−.170[c]	−.043	−..149[c]	−.094[a]	−.029	−.135	−.149	864	18
4. OCD	.166[a]	−.049	.000	—	−.234[c]	.135[b]	.072[a]	−.032	.224[b]	.052	.121	685	14
5. Panic disorder	.024	−.411[a]	−.141[a]	−.150[a]	—	.050	−.015	−.151[c]	−.035	.023	−.010	1471	30
6. Posttraumatic stress disorder	.202[a]	.017	.128[c]	.124[c]	.203[a]	—	.369[c]	.086	.308[b]	.156	.343[c]	178	4
7. Major depressive disorder	.436[a]	.210[a]	−.023	.164[a]	.035	.333[a]	—	.026	−.447[c]	.193[c]	.196[b]	1334	27
8. Dysthymia	.227[a]	.286[a]	−.031	−.084	−.091[c]	.170[b]	.106[b]	—	−.003	.160[a]	.103	334	7
9. Bipolar disorder	.132	.148	.169[c]	.325[a]	−.025	.373[a]	−.207[b]	−.171	—	.004	.350[b]	45	1
10. Alcohol abuse/dependence	.112[b]	.079	−.115[b]	.015	.108[b]	.256[a]	.224[a]	.127[c]	.196[c]	—	.500[c]	187	2
11. Drug abuse/dependence	.171[a]	.115[b]	−.038	−.031	.072	.236[a]	.244[a]	.157[b]	.124	.698[a]	—	45	1
Lifetime *n*	925	1253	623	469	1011	167	1429	297	42	350	270		
Lifetime %	35	47	24	18	38	6	54	11	2	13	10		

Notes. Current diagnosis correlations are on the top diagonal ($N = 4{,}928$), lifetime diagnosis correlations are on the bottom ($N = 2{,}655$). *n* = number of patients diagnosed with each condition. OCD = obsessive-compulsive disorder. [a]$p < .001$; [b]$p < .01$; [c]$p < .05$

than with anxiety and depressive disorders, consistent with a division between internalizing and externalizing dimensions. Correlations among internalizing disorders were generally not as strong, and in some cases they were negative. Panic disorder and specific phobia, in particular, were inversely related with several other internalizing syndromes, including a correlation of −.44 between panic disorder and social phobia. Thus, this pattern of correlations was at odds with the notion that all anxiety and depressive disorders would be strong indicators of a latent internalizing spectrum.

3.2 | Confirmatory factor models

The top half of Table 2 shows the fit indices for the planned CFAs for current disorders. The unidimensional model was clearly a poor fit to the data. The two-factor model also fit inadequately, in large part because several disorders had unexpected loadings on their hypothesized factors. Both specific phobia ($\lambda = -.07$, $p = .07$) and panic disorder ($\lambda = -.12$, $p < .01$) had negative loadings on the internalizing factor, and loadings for bipolar disorder and obsessive-compulsive disorder were not substantial (i.e., λs < .30), albeit statistically significant at a .05 alpha threshold (see Table 3). The three-factor solution was not viable, despite fit indices that approximated acceptable levels, due to an out-of-range parameter estimate (i.e., a negative variance for the fear factor).[3] The model fit comparison and factor loading pattern was very similar for lifetime disorders (Tables 2 and 3).

3.3 | Exploratory factor models

In post hoc supplementary analyses, we computed a series of exploratory factor analyses (EFAs) to search atheoretically for the best-fitting dimensional model in this sample. We tested EFA models that extracted one through four factors. We based the four-factor limit on prior latent variable modeling research with similar input data (Kotov et al., 2017). For current diagnoses, only the unidimensional EFA converged, although it did not fit the data acceptably, $\chi^2(44) = 720.90$, CFI = 0.56, TLI = 0.45, RMSEA = .056. EFAs of two through four factors produced improper solutions due to negative residual variance estimates for social phobia. A very similar pattern of results emerged for lifetime diagnoses.

We also used exploratory structural equation modeling analyses to attempt a data-driven approach to finding a suitable bifactor solution. We sought to estimate models with one general factor and between two and four group factors. However, paralleling our CFA results[3], these models were not interpretable. In the case of current diagnoses, the exploratory bifactor models did not converge, again due to negative indicator residual variance estimates. For lifetime diagnoses, bifactor models with two and four group factors converged, but several factors lacked substantial, statistically significant factor loadings.

4 | DISCUSSION

We tested a promising dimensional model of mental disorders in a sample of outpatients diagnosed with emotional disorders. This was, to date, the largest clinical sample used for such model-fitting. A two-factor model including internalizing and externalizing spectra did not adequately capture the pattern of mental disorder comorbidity in this dataset for current or lifetime diagnoses. This is the first clinical evidence to challenge the reproducibility of the internalizing and externalizing dimensions, which have been widely supported in large-scale community samples (Lahey et al., 2016).

TABLE 2 Fit indices for confirmatory factor analyses of current and lifetime diagnoses

Model	*df*	χ^2	CFI	TLI	RMSEA	BIC
Current[a]						
1-factor	44	720.89	.560	0.450	0.056	39138.29
2-factor	43	696.66	.575	0.456	0.056	39127.09
3-factor[c]	41	323.34	.816	0.754	0.037	38809.88
Lifetime[b]						
1-factor	44	679.98	.579	0.474	0.074	26606.83
2-factor	43	480.90	.713	0.633	0.062	26471.60
3-factor[c]	41	218.14	.884	0.844	0.040	26329.17

Notes. CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square of approximation; BIC = Bayesian information criterion; df = degrees of freedom.
[a]$N = 4{,}928$
[b]$N = 2{,}655$
[c]Model included a negative model-implied variance for the fear dimension.

TABLE 3 Standardized factor loadings for the two-factor model for current and lifetime diagnoses

	Current[a]				Lifetime[b]			
	Internalizing		Externalizing		Internalizing		Externalizing	
Diagnosis	λ	*p*	λ	*p*	λ	*p*	λ	*p*
Generalized anxiety disorder	.531	.000			.604	.000		
Social phobia	.509	.000			.397	.000		
Specific phobia	−.245	.000			−.072	.071		
Obsessive-compulsive disorder	.111	.001			.186	.000		
Panic disorder	−.313	.000			−.121	.001		
Posttraumatic stress disorder	.398	.000			.414	.000		
Major depressive disorder	.579	.000			.644	.000		
Dysthymia	.316	.000			.348	.000		
Bipolar disorder	.149	.068			.175	.028		
Alcohol abuse/dependence			.610	.000			.774	.000
Drug abuse/dependence			.820	.000			.902	.000

Notes. λ = standardized factor loading.
[a]$N = 4{,}928$
[b]$N = 2{,}655$.

This result points to several avenues for future research into the structure of psychopathology. First, inconsistency in disorder prevalence across clinical and community settings might promote discrepant structural findings. Simulation studies could be a useful tool for explicating associations between diagnostic frequency and coherence of latent dimensions of psychopathology. Second, we strictly adhered to *DSM-IV* hierarchical exclusion rules—a standard part of most assessment procedures that mirror *DSM*—during diagnostic assessment. This differential diagnosis practice skews the joint distribution of some closely related conditions. For instance, we observed small, and sometimes negative, associations among major depression, dysthymia, and bipolar disorder, likely straining the fit of all confirmatory models we tested. In contrast, in the other large clinical sample to evaluate the internalizing and externalizing spectra, Kotov et al. (2011) relaxed hierarchical decision-making rules and found much larger, positive correlations. We urge investigators to consider sidestepping, or at least temporarily suspending, these hierarchy conventions when designing new studies to avoid artificial constraints imposed by *DSM* rules, which are unlikely to reflect the true organization of psychopathology. Better still, we advise for future research to make use of more homogeneous symptom components (e.g., anhedonia, traumatic intrusions, hostility) that cut across traditional disorder boundaries to delineate the latent structure of mental disorders (e.g., Markon, 2010). This approach, unconstrained by historical diagnostic norms, may improve the precision of dimensional frameworks and thus more effectively "carve nature at its joints." Third, the diversity of indicators available for structural modeling probably influences the reproducibility of higher-order dimensions of mental disorder. The externalizing spectrum was defined here, as in several other investigations (e.g., Griffith et al., 2010), by substance use diagnoses only (cf. Krueger, Markon, Patrick, Benning, & Kramer, 2007). Future clinical work should aim for a denser representation of externalizing and internalizing liability alike to ensure the generalizability of latent constructs.

5 | CONCLUSION

We conclude that the results of diagnosis-based analyses from community samples may not replicate in all clinical samples, and we therefore urge investigators to examine the generalizability of the HiTOP model. Under what circumstances do the internalizing and externalizing spectra (and other dimensions) best recover the patterns of association among psychiatric problems? We expect there will be meaningful variation across samples, assessment instruments, and coverage of diagnoses. In particular, our data underscore the importance of validation studies in diverse patient groups. This additional research will, we believe, allow investigators to extend the HiTOP model and maximize its utility across research and clinical contexts.

ENDNOTES

[1] http://medicine.stonybrookmedicine.edu/HITOP

[2] Sample sizes for current and past disorders differ because past disorders were not assessed in all 5-year periods of this grant-funded research. The demographic features of the lifetime disorder subsample and the full sample were nearly identical.

[3] In auxiliary analyses, we examined a bifactor model in which all diagnoses loaded on a general factor and either an internalizing or externalizing specific factor. All factors were uncorrelated with one another. This approach was based on some prior evidence for a general factor of psychopathology thought to reflect liability to all mental disorders (Lahey et al., 2012). The bifactor model was not an adequate fit to the current data because the specific factors did not have interpretable patterns of factor loadings. The specific internalizing factor, for instance, had only one statistically significant loading. Full results are available upon request.

ORCID

Christopher C. Conway PhD http://orcid.org/0000-0002-6848-2638

REFERENCES

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. (2nd ed). Guilford.

Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology*, *110*, 585–599.

Brown, T. A., DiNardo, P., & Barlow, D. H. (1994). *Anxiety disorders interview schedule for DSM-IV: Lifetime version (ADIS-IV-L)*, San Antonio, TX: Psychological Corporation.

Clark, L. A., & Watson, D. (2006). Distress and fear disorders: An alternative empirically based taxonomy of the 'mood'and 'anxiety'disorders. *The British Journal of Psychiatry*, *189*, 481–483.

Forbes, M. K., Tackett, J. L., Markon, K. E., & Krueger, R. F. (2016). Beyond comorbidity: Toward a dimensional and hierarchical approach to understanding psychopathology across the life span. *Development and Psychopathology*, *28*, 971–986.

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

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

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., & Bagby, R. M. … Widiger, T. A. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, *126*, 454–477.

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

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

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

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

Krueger, R. F., Tackett, J. L., & MacDonald, A. (2016). Toward validation of a structural approach to conceptualizing psychopathology: A special section of the Journal of Abnormal Psychology. *Journal of Abnormal Psychology*, *125*, 1023–1026.

Lahey, B. B., Applegate, B., Hakes, J. K., Zald, D. H., Hariri, A. R., & Rathouz, P. J. (2012). Is there a general factor of prevalent psychopathology during adulthood? *Journal of Abnormal Psychology*, *121*, 971–977.

Lahey, B. B., Krueger, R. F., Rathouz, P. J., Waldman, I. D., & Zald, D. H. (2016). A hierarchical causal taxonomy of psychopathology across the life span. *Psychological Bulletin*, *143*, 142–186.

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

Miller, M. W., Fogler, J. M., Wolf, E. J., Kaloupek, D. G., & Keane, T. M. (2008). The internalizing and externalizing structure of psychiatric comorbidity in combat veterans. *Journal of Traumatic Stress*, *21*, 58–65.

Muthén, L. K., & Muthén, B. O. (1998–2014). *Mplus version 7*. Los Angeles, CA: Author.

Raftery, A. E. (1993). Bayesian model selection in structural equation models. In K. A. Bollen and J. S. Long (Eds). *Testing structural equation models*. Newbury Park, CA: Sage. (pp. 163–180).

Sellbom, M. (2016). Elucidating the validity of the externalizing spectrum of psychopathology in correctional, forensic, and community samples. *Journal of Abnormal Psychology*, *125*, 1027–1038.

Wolf, E. J., Miller, M. W., Krueger, R. F., Lyons, M. J., Tsuang, M. T., & Koenen, K. C. (2010). Posttraumatic stress disorder and the genetic structure of comorbidity. *Journal of Abnormal Psychology*, *119*, 320–330.

Wright, A., Krueger, R. F., Hobbs, M. J., Markon, K. E., Eaton, N. R., & Slade, T. (2013). The structure of psychopathology: Toward an expanded quantitative empirical model. *Journal of Abnormal Psychology*, *122*, 281–294.

How to cite this article: Conway CC, Brown TA. Evaluating dimensional models of psychopathology in outpatients diagnosed with emotional disorders: A cautionary tale. *Depress Anxiety*. 2018;35:898–902. https://doi.org/10.1002/da.22740