HHS Public Access

Author manuscript

Clin Psychol Sci. Author manuscript; available in PMC 2023 November 17.

Published in final edited form as:

Clin Psychol Sci. 2023 November ; 11(6): 1044–1063. doi:10.1177/21677026221119483.

The dimensional structure of internalizing psychopathology: Relation to diagnostic categories

Hannah R. Snyder[1], **Rebecca L. Silton**[2], **Benjamin L. Hankin**[3], **Harry R. Smolker**[4], **Roselinde H. Kaiser**[4], **Marie T. Banich**[4,6], **Gregory A. Miller**[3,5], **Wendy Heller**[3]

[1]Department of Psychology, Brandeis University

[2]Department of Psychology, Loyola University Chicago

[3]Department of Psychology, University of Illinois at Urbana-Champaign

[4]Department of Psychology and Neuroscience, University of Colorado Boulder

[5]Department of Psychology and Department of Psychiatry and Biobehavioral Sciences, UCLA

[6]Institute of Cognitive Science, University of Colorado Boulder

Abstract

Recent approaches aim to represent the dimensional structure of psychopathology, but relatively little research has rigorously tested sub-dimensions within internalizing psychopathology. This study tests pre-registered models of the dimensional structure of internalizing psychopathology, and their relations with current and lifetime depressive and anxiety disorders diagnostic data, in adult samples harmonized across three sites (n=427). Across S-1 bifactor and hierarchical models, we found converging evidence for both general and specific internalizing dimensions. Depression, generalized anxiety disorder (GAD), social anxiety disorder (SAD), and panic attacks were all associated with a general internalizing factor that we posit primarily represents motivational anhedonia. GAD was also associated with a specific anxious apprehension factor, and SAD with specific anxious apprehension and low positive affect factors. We suggest that dimensional approaches capturing shared and specific internalizing symptom facets more accurately describe the structure of internalizing psychopathology and provide useful alternatives to categorical diagnoses to advance clinical science.

Keywords

internalizing; latent variable; anxious apprehension; anxious arousal; low positive affect; loss of interest

Concerns with a categorical approach to psychiatric classification have long been noted (e.g., Brown & Barlow, 2005; Coghill & Sonuga-Barke, 2012; Kendler et al., 2011; Widiger & Samuel, 2005). To address these limitations, recent efforts have examined affective, physiological, or symptom dimensions associated with psychopathology, as reflected in the Research Domain Criteria initiative of the National Institute of Mental Health (RDoC;

Insel & Cuthbert, 2009; Kozak & Cuthbert, 2016; Miller et al., 2016; Yee et al., 2015), p factor models (e.g., Lahey et al., 2017), and the Hierarchical Taxonomy of Psychopathology (HiTOP) approach (e.g., Conway et al., 2019; Kotov et al., 2017). While differing in methods and included dimensions, these approaches all aim to represent the structure of psychopathology as multiple continuous dimensions along which individuals vary, as a more accurate model than traditional category-oriented approaches, such as the DSM (American Psychiatric Association, 2013). Such approaches show promise for clarifying risk factors and etiology and for predicting outcomes (e.g., Lahey et al., 2017).

Moving toward validated dimensional models of internalizing psychopathology is particularly critical given the extremely high comorbidity among categorically-defined internalizing disorders and the lack of evidence for one-to-one correspondence of causal risk mechanisms to specific internalizing disorders (e.g., Hankin et al., 2016). This pattern of strong co-occurrence suggests that traditional psychiatric nosologies (e.g., DSM, ICD) do not accurately represent the structure of internalizing psychopathology, impeding efforts to understand risk mechanisms, predict outcomes, and target interventions. However, efforts to date have been focused mainly on the highest levels of the hierarchy– general psychopathology, and broad externalizing and internalizing dimensions, with relatively little research that rigorously tests sub-dimensions within externalizing and internalizing psychopathology in a way that allows associations with higher-order, broad dimensions to be statistically differentiated from those with specific sub-dimensions. Here, we test models that include specific dimensions within internalizing psychopathology.

Conceptual models of internalizing psychopathology

Several proposed dimensional models of internalizing psychopathology account for comorbidity across categorically-defined disorders by positing both broad (higher-order or shared) and specific internalizing dimensions. The tripartite model proposed that internalizing disorders are best modeled by a common component characterized by distress and negative affect, along with specific anhedonic depression/low positive affect and anxious arousal components. Data in both community and clinical samples are consistent with this model (e.g., Clark & Watson, 1991; Joiner, 1996; Teachman et al., 2007). An extension of this model proposed a broad internalizing spectrum with further division into distress and fear subfactors (e.g., Watson, 2005), a structure that has subsequently been incorporated into the HiTOP conceptual model (e.g., Kotov et al., 2017). However, these models both omit a key, distinct, internalizing dimension– anxious apprehension (worry)–and do not capture the full dimensional space along which forms of psychopathology vary. For example, by grouping depression and generalized anxiety disorder (GAD) together under “distress”, they miss the hallmark distinguishing feature of GAD — uncontrollable and excessive worry, which is not common in MDD absent comorbid GAD — as well as the more prominent role of low positive affect in anhedonic depression.

To overcome this limitation, Heller and colleagues proposed an expanded model (Nitschke et al., 2001), based on a neuropsychological model of brain function (Heller, 1990) and other theoretical conceptualizations regarding symptom structure (e.g., the tripartite model) and risks for depression and anxiety disorders (Barlow, 1991). This model

identified symptom dimensions corresponding to anxious arousal, anxious apprehension, and anhedonic depression (which was further deconstructed into low positive affect [LPA] and a dimension reflecting anhedonia, loss of interest and other symptoms of clinical depression; Nitschke et al., 2001). This model was found to have good fit to the data by previous factor-analytic investigation (Nitschke et al., 2001). These symptom dimensions have subsequently been associated with distinct physiological functions (Sharp et al., 2015), cognitive processes (Crocker et al., 2012; Guha et al., 2020; Sharp et al., 2015; Silton et al., 2011; Warren et al., 2013), and patterns of regional brain activity (Engels et al., 2007, 2010; Heller & Nitschke, 1998; Heller et al., 1997; Nitschke et al., 1999; Nitschke & Heller, 2005) and connectivity (Burdwood et al., 2016), providing initial criterion and divergent validity of these dimensions.

In addition to distinct cognitive, affective and physiological profiles, these symptom dimensions have been linked to different DSM diagnoses, but generally not with one-to-one mappings. LPA was originally theorized to be specific to depression and anxious arousal to be specific to anxiety (Watson et al., 1988). Supporting this theory, LPA has been associated with unremitting depression (Kumari et al., 2003) as well as risk for future depressive episodes (Pine et al., 1999; Wilcox & Anthony, 2004). Furthermore, trait LPA predicts depressive symptoms even after accounting for high trait negative affect, which suggests a distinct contribution of LPA to depressive symptoms (Loh et al., 2014; Verstraeten et al., 2009). However, LPA has subsequently also been linked to social anxiety, although potentially not as strongly as to depression (Watson & Naragon-Gainey, 2010).

Evidence is less clear that anxious arousal relates specifically to anxiety disorders. Although anxious arousal symptoms clearly occur *during* panic attacks, surprisingly little research has tested whether anxious arousal is also higher on average in individuals with panic disorder or who have a history of panic attacks. Some studies have found that although anxious arousal symptoms are elevated in those with panic attacks, they are not higher than in those with other internalizing disorders (Grisanzio et al., 2018; Williamson et al., 2005), whereas other research suggests a more specific link between anxious arousal and panic disorder (Wardenaar et al., 2012).

Regarding anxious apprehension symptoms, they are higher in GAD than in other anxiety disorders, but elevated in all anxiety disorders compared to participants with no diagnosis (Olatunji et al., 2010), leaving specificity unclear. Some evidence suggests that, controlling for comorbidity, only GAD and SAD are related to anxious apprehension (Starcevic et al., 2007). In sum, extant research supports some aspects of theories that particular dimensions relate relatively specifically to certain depression and anxiety disorders, although there are gaps and inconsistencies that preclude a full understanding of connections among dimensional symptom profiles and psychiatric disorders.

Statistical model-based approaches to analyze structure of internalizing psychopathology

Critically, however, research to date has largely not used analytic approaches that could statistically separate broad, general internalizing psychopathology from what is specific

to LPA, anxious arousal, or anxious apprehension. In particular, because depression and anxiety symptoms commonly co-occur, associations of risk factors and outcomes with each symptom dimension could be driven by what they share (general internalizing symptoms) or by what is specific to each (e.g., symptoms specific to anxious arousal). Thus, it remains unclear where DSM depression and anxiety diagnoses are located within this multidimensional space and whether cognitive, psychosocial, and biological risk factors and outcomes implicated in these disorders confer broad or dimension-specific vulnerability to internalizing psychopathology.

Recent approaches model the structure of psychopathology in ways that allow broad vs. dimension-specific associations to be differentiated, using latent variable models that statistically separate general internalizing and specific (e.g., anxious-arousal specific, anxious-apprehension specific and LPA-specific) internalizing dimensions. Bifactor and hierarchical latent variable models are the most common approaches. There are strengths and limitations to each of these modeling approaches (Bader & Moshagen, 2021; Greene et al., 2019; see also Analysis Methods), and they are thus best seen as providing complementary tools.

Bifactor models capture the general component with a factor on which all indicators (e.g., symptoms) load and specific factors that represent the unique variance associated with sub-dimensions not accounted for by the general factor (e.g., what is specific to anxious arousal symptoms after accounting for general internalizing symptoms). Numerous studies use bifactor models to clarify risk factors (e.g., stress, cognitive function, genetic risk) and outcomes (e.g., adaptive functioning) for broad psychopathology dimensions (p factor, internalizing-specific and externalizing-specific; for review, see Lahey et al., 2017). Minimal research to date has employed this approach to investigate risk for more fine-grained dimensions of internalizing symptoms. One study of the tripartite model dimensions (measured via the Mood and Anxiety Symptom Questionnaire [MASQ]) found that they were well represented by a bifactor model with a general factor and anxious-arousal and anhedonia-specific factors and that these factors had distinct patterns of correlations with different symptoms of depression (Lin et al., 2014). Using more general depression and anxiety symptom measures, several studies found that a bifactor model with general internalizing, anxiety-specific, and depression-specific components fit well in treatment-seeking college students (Fassett-Carman et al., 2020) and primary care patients (Simms et al., 2012) and that these factors are differentially associated with risk factors (stress exposure and appraisals; Fassett-Carman et al., 2020) and outcomes (psychosocial dysfunction; Simms et al., 2012). However, none of these studies included a specific measure of anxious apprehension. One study in an outpatient clinical sample that did include anxious apprehension and anxious arousal along with a broader pool of dimensions (fatigue, phobic fear, worthlessness, and tension) chose to omit LPA from the model (den Hollander-Gijsman et al., 2012). This study found that a bifactor model including a general factor and all the specific factors for the dimensions above fit well; collapsing anxious apprehension into the general factor resulted in a poorly fitting model, supporting it as a separable dimension.

Hierarchical models represent the same structure with a different parameterization: instead of all indicators (e.g., symptoms) loading onto a general factor and their respective specific

factor as in a bifactor model, indicators load onto subfactors, and these subfactors then load onto a higher-order factor (i.e., the higher-order factor is defined by the covariance of the subfactors, not the covariance of the indicators). Several studies have tested hierarchical models of internalizing psychopathology in which the indicators were DSM diagnostic categories (e.g., Krueger, 1999; Vollebergh et al., 2001; Watson, 2005); however, this recapitulates the problems with diagnostic categories noted above, in that it does not allow consideration of how internalizing dimensions *cut across and vary within* diagnostic categories. This problem can be addressed by testing hierarchical models at the symptom rather than disorder level. Several of the studies above testing bifactor models also tested hierarchical models at the symptom level, which generally also fit well (den Hollander-Gijsman et al., 2012; Lin et al., 2014; Simms et al., 2012). Both approaches thus have the potential to clarify how general and specific internalizing dimensions characterize different mood and anxiety disorders and ultimately to clarify risk and outcome pathways. In addition, testing associations with both bifactor and hierarchical models of the same data is important for demonstrating robustness of findings across modeling approaches.

Current Study, Goals, and Hypotheses

The current study aims to enhance understanding of internalizing psychopathology dimensions by addressing several important gaps and limitations in prior research. First, there are important limitations to both fear/distress models (e.g., failing to distinguish the hallmark symptoms of depression vs. GAD) and the tripartite model (e.g., omitting the dimension of anxious apprehension), which our hypothesized model addresses. Second, few prior studies have tested dimensional internalizing latent variable models that include both general and specific factors, which can be a valuable tool for clarifying the generality vs. specificity of links between internalizing dimensions and risk and outcome variables. Third, these dimensional models have not been examined in relation to DSM diagnostic data, which is an important step in bridging previous diagnostic category-based research to current dimensional approaches. Making connections between dimensional and categorical diagnostic systems is essential because a vast body of prior research has used diagnostic categories. By making this mapping we can draw on this literature to consider theories and predictions regarding dimensional constructs where less direct evidence exists as of yet. In addition, diagnostic categories are still, and are likely to remain for some time, the primary unit of analysis in clinical settings (indeed, the RDoC initiative was not conceived to replace DSM in clinical practice; Cuthbert et al., in press). Thus, better understanding the relation of these diagnoses to symptom dimensions is critical for clinical translational work.

To address these gaps, the current study models the dimensional structure of internalizing psychopathology (ascertained with self-report questionnaires) and tests their relation to lifetime diagnostic data on unipolar depressive disorders (major depressive disorder [MDD] and persistent depressive disorder [PDD]), generalized anxiety disorder (GAD), social anxiety disorder (SAD), and panic disorder (PD) (ascertained with semi-structured diagnostic interviews) in a combined sample of adults across three sites. The hypotheses and data analysis plan were preregistered prior to data analysis.[1]

As noted above, past studies evaluating bifactor or hierarchical latent variable models of internalizing have not included specific measures of all four dimensions: anxious arousal, anxious apprehension, LPA, and general internalizing (Nitschke et al. [2001] tested correlated factor models with no general internalizing factor). Moreover, available studies have primarily used treatment-seeking samples, where the structure of psychopathology may differ markedly from that in the general population. Thus, the first goal of the current study was to fit bifactor and hierarchical models to fully test Heller and colleagues' four-dimensional model in community samples to capture the full dynamic range of symptoms dimensionally, in accord with the RDoC approach (Kozak & Cuthbert, 2016). The second goal of the present study was to more rigorously investigate how DSM-defined depression and anxiety disorders relate to both general and specific (anxious-arousal, anxious-apprehension, and LPA) internalizing dimensions. Identifying the unique variance associated with all dimensions of interest is essential to successfully relate these specific dimensions to particular disorders.

We hypothesized that a bifactor model, with a general internalizing factor and specific anhedonia (LPA; based on previous analyses we expect loss of interest [LI][2] items to load only on the general factor; see Analysis Methods), anxious arousal, and anxious apprehension factors would fit the data well. We also tested a hierarchical (second-order) model, with a higher-order general internalizing factor and LPA, loss of interest, anxious arousal, and anxious apprehension sub-factors (consistent with the bifactor model, we anticipated that the LI first-order factor might load almost perfectly on the general factor, with little residual variance specific to LI). We also expected this model to fit well, given that hierarchical and bifactor models often fit equally well (Bader & Moshagen, 2021; Mansolf & Reise, 2017). Thus, we did not aim to compare the fit of these models but rather used them as complementary ways of modeling the same hypothesized structure, to evaluate robustness of results to alternative modeling approaches. We also evaluated the reliability of factors in both models (Rodriguez et al., 2016).

We then tested associations of these dimensional factors with DSM diagnostic interview data, coded on a 4-point severity scale (see SCID section of Measures). Primary analyses used lifetime history of disorder, given that individuals who currently do not meet diagnostic criteria frequently continue to have elevated symptoms and risk for recurrence (e.g., Bystritsky, 2006; Fava et al., 2007), but we also included analyses with only current diagnoses.

[1]Since the present study is a secondary data analysis, hypotheses and the data analysis plan were preregistered before data analysis but after data collection. We also conducted two additional types of analyses that were not preregistered. First, for the hierarchical models we used indirect effects to quantify the proportion of the total effects mediated by the general factor, after we became aware of the recommendation to do so (Conway et al., 2021). Note that estimating the indirect effects does not change the model at all, but simply outputs this additional information. Second, upon reviewer request, we additionally calculated correlations among factor scores across models to better compare the measurement models. The preregistration, along with the data and Mplus analysis scripts and full analysis outputs are available here: https://doi.org/10.17605/OSF.IO/XTD34

[2]In addition to "loss of interest", this subscale has been given a variety of labels in the literature, including "anhedonia", "depression" or "depressed mood", "low positive emotions", "negative affectivity" and simply the "MASQ AD-8". We use "loss of interest" (LI) throughout for simplicity and clarity, although some of its items are narrower than others, and some may tap distress as well. See the Discussion section for further consideration of the constructs this subscale may be capturing.

We hypothesized that, compared to participants with no lifetime history of mood or anxiety disorders on the SCID, individuals with a lifetime history of (1) all depressive and anxiety disorders would have elevated levels on the general internalizing factor, (2) depressive disorders (MDD and persistent depressive disorder [PDD, which includes persistent dysthymia]) would additionally have high levels of LPA, (3) GAD would additionally have higher levels of anxious apprehension, (4) PD would additionally have higher levels of anxious arousal, and (5) SAD might have both higher levels of LPA and anxious apprehension. Since these models are novel, we did not have hypotheses about the relative strengths of associations with general internalizing vs. the specific factors.

Methods

Participants

Data were harmonized across three samples to increase sample size for latent variable analyses and increase geographic and demographic diversity. The combined sample included 427 adults ages 18–65 (M = 29.43 years, SD=6.30; 75% female, 23% male, 2% other/neither). Participants identified their race as 72% white, 15% Asian, 7% Black/African American, 4% other/multiracial, 1% American Indian/Alaska Native, and >1% Native Hawaiian/Pacific Islander; 8% identified their ethnicity as Hispanic/Latino.

Samples consisted of unselected community participants (to capture the full dimensional spectrum of symptoms) and had been given the same questionnaires (MASQ and PSWQ) and diagnostic measures (depression and anxiety disorders assessed with the Structured Clinical Interview for the DSM). Sample one included 117 community adults ages 19–51 (M = 34.8, SD = 9.37; 67.5% female) from Champaign-Urbana, Illinois. Sample two included 86 community adults ages 35–65 (M = 48.3, SD = 8.55; 93.0% female) from the Denver, Colorado, metro area. Seven participants from this sample were missing SCID diagnostic data due to scheduling difficulties. Sample three were 224 undergraduate students age 18–23 (M = 19.29, SD = 1.27; 71% female, 25% male, 4% other/neither) at Brandeis University in Waltham, Massachusetts. One participant from this sample was missing questionnaire data. See online supplemental materials for full descriptions of each sample. All participants provided informed consent, and all study procedures were approved by the IRB at their respective institutions.

Measures

Mood and Anxiety Symptoms Questionnaire (MASQ).—The MASQ assesses anxious arousal and anhedonic depression symptom severity in the past week (Watson et al., 1995). It has good internal consistency, test-retest reliability, and convergent and discriminant validity in relation to depression and anxiety disorders in relevant samples (e.g., Bredemeier et al., 2010; Nitschke et al., 2001; Watson et al., 1995). There are two anhedonic depression subscales, low positive affect (LPA; 14 items, e.g., “Felt like I was having a lot of fun” – reverse coded) and loss of interest (LI; 8 items; e.g., “Felt really bored”; Nitschke et al., 2001). The anxious arousal subscale (AAr; 17 items) assesses somatic symptoms, mainly of physiological hyperarousal (e.g., “Hands were shaky”).

Penn State Worry Questionnaire (PSWQ).—Anxious apprehension was assessed with the PSWQ (Meyer et al., 1990), which asks about (trait) tendency to worry (16 items, e.g., “My worries overwhelm me.”). It has good internal consistency, test-retest reliability, and convergent and discriminant validity in relation to anxiety disorders in relevant samples (e.g., Brown et al., 1992; Kertz et al., 2014; Meyer et al., 1990).

Structured Clinical Interview for the DSM (SCID).—The Structured Clinical Interview for DSM (SCID, American Psychiatric Association, 2015) was used to assess diagnoses of depressive and anxiety disorders in each sample (sample 1 used the SCID for the DSM-IV, and samples 2 and 3 used the SCID for the DSM-5; all symptom-level data were recoded following DSM-5 diagnostic criteria). In addition, the psychosis and mania screens were used to rule out participants with a psychotic or bipolar disorder, but none met criteria. The SCID is the most frequently used and studied diagnostic interview for adults, with strong evidence of reliability for evaluating categorical diagnoses (e.g., Osório et al., 2019), distinct from widespread reservations about the validity of DSM diagnoses themselves (e.g., Hyman, 2010; Yee, Javitt, & Miller, 2015). Diagnostic interviewers at each site were graduate students and staff who completed intensive training on conducting the interviews and determining diagnoses following DSM rules and who received ongoing supervision. In sample 1, both current and past depression was covered, but only current (past six months) anxiety disorders. In samples two and three, interviews covered both past and current depressive and anxiety disorders.

The following diagnoses were assessed in all samples, and coding was harmonized across sites for the present study using DSM-5 criteria: major depressive disorder (MDD), persistent depressive disorder (PDD, which includes chronic dysthymia), generalized anxiety disorder (GAD), social anxiety disorder (SAD), and panic disorder (PD). In addition, as many individuals have panic attacks but do not meet criteria for panic disorder (because diagnosable attacks have an identifiable trigger and/or are not accompanied by at least one month of worry about future attacks or behavioral change), we also separately coded panic attacks (PA) regardless of PD status. Since many individuals who do not meet full criteria for a disorder nonetheless have clinically elevated symptoms, we coded SCID data in a more dimensional manner as follows (see supplementary materials for analyses with traditional binary coding):

3 = Meets full DSM-5 criteria for disorder.

2 = Meets threshold (rating of 3) on hallmark symptom (depressed mood and/or anhedonia for MDD, depressed mood for PDD, persistent uncontrollable worry about 2+ domains for GAD, marked fear or anxiety about one or more social situations for SAD, recurrent and unexpected panic attacks for PD, one or more panic attack for PA regardless of PD status), but the total number of symptoms scored at threshold was one fewer than the required number for the diagnosis.

1 = Meets threshold rating on hallmark symptom but two or more fewer other symptoms at threshold than required for the diagnosis.

0 = No threshold hallmark symptom.

Analytic Approach

Measurement models.—Confirmatory factor analyses (CFA) were conducted in Mplus version 8 using robust full information maximum likelihood estimation (MLR, which is robust to non-normality). As χ^2 is sensitive to sample size, good model fit is defined as: root mean square error of approximation (RMSEA) <.06 good, <.08 acceptable; comparative fit index (CFI) >.95 good, >.90 acceptable; and standardized root mean square residual (SRMR) <.08 good, acceptable <.10 (Hu & Bentler, 1999).

One Factor and Correlated Factor models.: We first examined fit and modification indices for a model with correlated MASQ anxious arousal [AAr], low positive affect [LPA] and loss of interest [LI], and PSWQ [AAp]) factors, and where residual correlations between indicators were suggested (estimated correlations ≥ .3) they were added if they were justified based on similar item content/wording; these residual correlations were then added to all other models for fair comparison of fit. We compared the model fit of the bifactor and hierarchical factor models to this correlated factor model and a one-factor model with all indicators across scales loading onto a single factor (using AIC, as not all models are nested). However, we emphasize that model fit is not our primary criteria for choosing between models, given controversy regarding whether fit criteria are biased toward bifactor models (Greene et al., 2019) or not (Bader & Moshagen, 2021). Rather, we selected bifactor and hierarchical factor models a priori based on their suitability for addressing the research questions as described above. Statistical indices of factor reliability and validity (omega, ECV, H) were also calculated and evaluated (Rodriguez et al., 2016).

Bifactor CFA Model.: All items were specified to load onto the general internalizing factor that captures covariance across all items and additionally onto their respective specific factors that represent the remaining unique covariance in each scale/subscale (MASQ AAr, LPA and LI, and PSWQ AAp), after accounting for the general factor. Factors were constrained to be orthogonal to one another because what is shared between factors is already captured by the general factor (e.g., Chen et al., 2012).

We have previously tested a version of this model, using item parcels rather than individual items, in a subsample of participants in the current analyses (Samples 1 and 2 combined), for use in neuroimaging analyses (Banich et al., 2020). In this subsample, in an initial model including all four specific factors, LI parcels had strong loadings only on the general internalizing factor, and there was no significant variance associated with the LI-specific factor, indicating that the general factor fully accounted for covariance among the LI parcels. Thus, the LI-specific factor was eliminated, and LI parcels loaded only onto the general factor. Here we again initially tested a model with all four specific factors along with the general factor but hypothesized that evidence would be lacking for an LI-specific factor, in which case it would again be eliminated, with LI items loaded only on the general factor, thus specifying an S-1 bifactor model (i.e., one fewer specific factors than dimensions), which has been advocated as providing a more robust and interpretable factor structure (e.g., Eid, 2020).

Hierarchical CFA Model.: In the hierarchical (two-level) CFA model, item indicators load onto first-level factors (AAr, LPA, LI, and AAp), and these factors then load onto a second-level general internalizing factor. By default, the first-order factors are not correlated, as their covariances are now accounted for by the general factor.

SEMs testing relations to diagnostic interview measures.—We estimated SEM models with diagnostic variables (see SCID methods) predicting factors in the CFA models, controlling for age and gender. We ran these analyses first with each diagnosis tested in a separate model, then with all diagnoses included as predictors, to test specificity given comorbidity across disorders. Bifactor model SEMs were tested using the residual method (Koch et al., 2018), which provides unbiased estimates of the relation between the predictor and the general factor, free of influences of the specific factors, and vice versa. For hierarchical model SEMs, we tested for mediation of effects on the lower-order factors through the higher-order general internalizing factor using "model indirect" syntax in Mplus, to quantify the degree to which the general internalizing factor explains relations with the lower-order factors (total effects, which are identical to those in the correlated factors model; Conway et al., 2021).

We corrected for multiple comparisons using FDR correction with two thresholds: within each of the models and across all models (excluding exploratory models; study-wise correction). For *a priori* hypothesized effects, as preregistered, we considered effects with uncorrected $p < .05$ significant, whereas for any additional effects not hypothesized *a priori* here, we only considered effects significant if they have FDR corrected $p < .05$ for at least the model-wise correction. Standardized βs are reported to allow comparison of effect sizes.

Results

The data, along with all Mplus files, are available on OSF (link in Footnote 1). Table 1 provides descriptive statistics for the full sample, and Supplemental Materials Table S1 does so for each sample separately. Table 2 provides bivariate correlations among manifest variables. Diagnostic rates were generally in line with population estimates, adjusting for gender balance (Kessler et al., 2012; Vandeleur et al., 2017), except for GAD, which was more prevalent in the current sample than in epidemiological studies (Kessler et al., 2012) but equivalent to a large-scale longitudinal birth cohort study (Moffitt et al., 2007). As expected based on population prevalence rates, the rate of panic disorder (PD) was too low (4.3%, $n = 18$) to provide an adequate sample size for analysis; thus, we conducted all analyses with panic attacks rather than PD. In total, 69% of the sample had a lifetime history of at least one criterial symptom for at least one diagnosis (Table S2).

Measurement Models

Correlated factor comparison model.—We planned to compare the bifactor and hierarchical models to simpler one-factor and correlated-factor models. A one-factor model fit very poorly (CFI=.485, RMSEA=.094, SRMR=.212) and thus was not considered further, and only the correlated factor model was used for comparison.

Initial examination of the correlated-factors model indicated that 10 pairs of items with closely-related content (e.g., “Felt faint” and “Felt dizzy or light headed”) had substantial residual correlations ($r \geq .30$, Table S3); all but two residual correlations were within-factor (3 for AAp, 2 for AAr, 2 for LI, 1 for LPA), and the remaining two were between AAr and LI items that both refer to death and LPA and LI items that both refer to self-image. Since these residual correlations were all theoretically justified based on the similarity of the items, they were included in all models (correlated, bifactor, hierarchical) for head-to-head model comparison.[3]

The final correlated factor model with these residual correlations included had good fit (Table 3). All items loaded significantly on their respective factors (standardized loadings median (SD): AAp= 0.76 (.15), AAr=0.52 (.15), LPA=0.73 (.05), LI=0.58 (.11)) and factors were moderately to strongly correlated (Table S3). Metrics of factor reliability and replicability were good for all factors (ω > .8, H >.8; Table 2). Although the χ^2 difference test for metric invariance across sites was significant (Satorra-Bentler scaled χ^2 difference test = 141.02 (102), $p = .006$), the effect size was very small (Cohen’s $w = .057$, where .1 is considered small, Cheung & Rensvold, 2002), suggesting that loading differences across sites are minor.

Bifactor model.—As predicted based on prior analysis of a subsample of this data (Banich et al., 2020), an initial model including an LI-specific factor found non-significant and negative loadings on this specific factor, and strong loadings of the LI items only on the general factor, suggesting that the general factor fully accounts for covariance among the LI items. Thus, the LI-specific factor was eliminated, yielding an S-1 bifactor model with the LI items loading only on the general factor (Table S4). The meaning of the general factor is thus defined by the LI items. This model had good fit by RMSEA and SRMR (Table 2). All items loaded significantly on their respective factors (standardized loadings median (SD): general factor= 0.41 (.12), AAp-specific=0.63 (.12), AAr-specific=0.40 (.14), LPA-specific= 0.62 (.06)), with the exception of two AAr-specific loadings (these items, “I felt like I was choking” and “I was afraid I was going to die” had floor effects and also had standardized loadings <.3 in the correlated factor model).

Total reliability (ω) was good for all factors (>.85). The percentage of reliable variance in total scores attributed to individual differences in the general factor (ωH) was 68%, indicating that the general factor accounted for the majority of the variance, but that the measures cannot be considered unidimensional (Rodriquez et al., 2016). The percentage of reliable variance in subscale scores due to specific factors after accounting for the general factor (ωHS) was 65% for LPA-specific, 67% for AAp-specific, and 46% for AAr-specific. The general factor accounted for 42% of the explained common variance (ECV) across all questionnaire items, the LPA-specific factor accounted for 69% of the ECV in the MASQ LPA subscale, the AAr-specific factor accounted for 56% of the ECV in the MASQ AA subscale, and the AAp-specific factor accounted for 69% of the ECV in the PSWQ.

[3]To check robustness of the results to including these residual correlations, we ran the main SEM analyses (lifetime history of all diagnostic predictors simultaneously) for the correlated factor, bifactor and hierarchical models with no residual correlations included. Results were nearly identical, with no difference in the significance of associations vs. the analyses with residual correlations included. Full results output for the models with no residual correlations are available on OSF (link in Footnote 1).

Factor replicability (H) was good for all factors (>.86, Table 3) except the AAr-specific factor, which was slightly lower (.77). Thus, both the general and specific factors captured substantial percentages of the questionnaire variance, indicating they are best represented as having both shared and unique components as in the bifactor model and are generally reliable and replicable per the ω and H indices. The model demonstrated metric invariance across sites (Satorra-Bentler scaled χ^2 difference test = 201.47 (196), $p = .893$, $w = .049$ [very small effect]).

Hierarchical model.—This model had good fit by RMSEA and SRMR (Table 3). All items loaded significantly on the lower-level factors (with loadings nearly identical to those for the correlated factors model, as expected), and all lower-level factors loaded significantly onto the general (second-order) factor. Consistent with the bifactor model, the LI factor loaded very strongly on the general factor (β=.996) with no significant residual variance, indicating that the general factor fully accounts for the covariance among the LI items. The other three factors had loadings between .56-.65 (Table S5). Reliability (ω) and replicability (H) were good for all factors (≥.799, Table 2). Metric invariance of the first-order loadings was tested with the correlated factors model as described above, with a very small effect size w indicating that differences across sites were minor. Loadings on the second-order general internalizing factor demonstrated metric invariance across sites (Satorra-Bentler scaled χ^2 difference test = 11.38 (8), $p = .181$, $w = .058$ [very small effect]).

Correspondence across measurement models.—To test the degree to which the factors from the three measurement models captured the same constructs, we tested correlations among the extracted factor scores (Table S21). As expected, the factor scores for the general factors from the hierarchical and bifactor models and the LI factor from the correlated factors model were nearly identical. All factor scores from the correlated factors model correlated strongly with both the general factors from the bifactor and hierarchical models (rs= .61-.73, commensurate with their loadings on the general factor in the hierarchical model), and with their corresponding specific factor from the bifactor model (rs=.76-.84), but were only weakly correlated with the other specific factors from the bifactor model (rs=−.15-.13). This demonstrates good convergent and divergent relations of factors across models, indicating that results are robust across these different ways of modeling the data.

Structural Equation Models: Lifetime History

Models included age and gender as covariates, and exploratory analyses tested for age and gender moderation. Only results surviving at least within-model FDR correction are reported here; full results are reported in the Supplemental Materials tables.

Correlated factor model (Figure 1a, Figure 2a Table S6).—Because AIC favored the correlated factors model (although the bifactor and hierarchical models also had acceptable to good fit), we tested SEMs with the correlated factor model in addition to the preregistered tests with the bifactor and hierarchical models.

History of MDD/PDD was associated with all four factors, with the strongest association with LI (β=.353, p<.001) and smaller associations for LPA (β=.258, p<.001; difference from LI Z=2.17, p=.030), AAp (β=.227, p<.001; difference from LI Z=2.30, p=.022), and AAr (β=.136, p=.008; difference from LI Z=4.27 p<.001), which did not significantly differ from one another. Controlling for all other SCID predictors, history of MDD/PDD was still associated with LI (β=.285, p<.001), LPA (β=.212, p<.001), and AAp (β=.118, p=.008) but not AAr.

History of GAD was also associated with all four factors, with the strongest association for AAp (β=.370, p<.001) and weaker associations for AAr (β=.191, p=.003; difference from AAp Z=2.91, p=.004), LI (β=.192, p=.001; difference from AAp Z=3.47, p=.001), and LPA (β=.154, p=.001; difference from AAp Z=4.67, p<.001), which did not differ significantly from one another. Controlling for all other SCID predictors, history of GAD was still associated with AAp (β=.291, p<.001) but not with any of the other factors.

History of SAD was also associated with all four factors, with none of the associations differing significantly from one another: AAp (β=.241, p<.001), AAr (β=.147, p=.010), LPA (β=.251, p<.001), and LI (β=.208, p<.001). Controlling for all other SCID predictors, history of SAD was still associated with AAp (β=.187, p<.001), LPA (β=.221, p<.001), and LI (β=.160, p=.002), but not AAr.

History of panic attacks (PA) was also associated with all four factors: AAp (β=.238, p<.001), AAr (β=.194, p<.001), LPA (β=.172, p<.001), and LI (β=.281, p<.001); only LI and LPA significantly differed (Z=2.08, p = .038). Controlling for all other SCID predictors, history of PA was still associated with LI (β=.160, p=.010) but no other factors.

S-1 bifactor model (Figure 1b, Figure 2b, Table S7).—History of MDD/PDD was associated with the general internalizing factor, both when tested individually (β=.327, p<.001), and when controlling for all other SCID predictors (β=.212, p<.001), but was not associated with any of the specific factors. When tested individually, history of GAD was associated with the general internalizing factor (β=.241, p<.001) and the AAp-specific factor (β=.313, p<.001) but not the other specific factors. Controlling for all other SCID predictors, it was only associated with the AAp-specific factor (β=.293, p<.001). Both tested alone and controlling for all other SCID predictors, history of SAD was associated with the general factor (alone: β=.212, p<.001; all: β=.166, p=.001), and the AAp-specific (alone: β=.170, p=.001; all: β=.149, p=.002) and LPA-specific (alone: β=.173, p=.001; all: β=.176, p=.001)) but not AAr-specific factors. Both tested alone and controlling for all other SCID predictors, history of PA was associated with the general factor (alone: β=.296, p<.001; all: β=.161, p=.007), but not the specific factors.

Hierarchical model (Figure 1c, Figure 2c, Table S8).—We did not test associations with the lower-level LI factor because it had no significant residual variance to predict, as it loaded almost perfectly on the higher-level general internalizing factor (see Hierarchical Measurement Model above). History of MDD/PDD was associated with the general internalizing factor both tested alone (β=.353, p<.001), and controlling for all other SCID predictors (β=.284, p<.001), but not with the lower-level factors. Tests of mediation showed

that indirect effects through the general factor were all significant and accounted for large proportions of the total effects. History of GAD was associated with the general factor when tested alone only (β=.192, p=.001) and the lower-level AAp factor both when tested alone (β=.285, p<.001) and controlling for all other SCID predictors (β=.273, p<.001), but was not associated with the other lower-level factors. Indirect effects through the general factor were significant for all three lower-level factors, but with a larger proportion mediated for LPA and AAr than AAp. Both tested alone and controlling for all other SCID predictors, history of SAD was associated with the general factor (alone: β=.209, p=.001; all: β=.160, p=.002) and the lower-level AAp (alone: β=.142, p=.001; all: β=.121, p=.003) and LPA (alone: β=.138, p=.003; all: β=.139, p=.002) but not AAr factors. Indirect effects through the general factor were significant for all three lower-level factors, with a larger proportion mediated for AAr than LPA and AAp. Both tested alone and controlling for all other SCID predictors, history of PA was associated with the general factor (alone: β=.281, p<.001; all: β=.160, p=.010) but not the lower-level factors. Indirect effects through the general factor were significant for all three lower-level factors, with a larger proportion mediated for AAr and LPA than AAp.

Exploratory Supplemental Analyses

To test robustness of our results to different ways of modeling diagnostic data, and across ages and genders, we conducted exploratory analyses with only current rather than lifetime diagnoses (thus mapping more closely onto the timeframe of the MASQ, which assesses past week symptoms), binary (yes/no) coding of lifetime diagnoses, and age and gender moderation of lifetime dimensional diagnoses. To limit multiple comparisons, we tested these exploratory analyses only with all predictors together in the same model. Only effects that survived within-model FDR correction are noted (see Supplemental Materials Tables S9–20 for full results, and S21 for a summary of results across analyses).

Current diagnoses (Figure 2, Tables S9–11).—Results for current diagnoses were similar to the main lifetime dimensional models. The significance of all effects remained the same as in the lifetime models with the following exceptions. In the correlated factors model, MDD/PDD with AAp was no longer significant, whereas PA with AAp and AAr became significant (Table S9). In the bifactor model, GAD with the general factor and MDD/PDD with LPA-specific became significant, albeit with a small effect size (Table S10). In the hierarchical model, SAD with the general and lower-level LPA factors were no longer significant with FDR correction (although they were significant uncorrected, Table S11).

Binary diagnoses (Tables S12–14).—The significance of all effects remained the same as in the lifetime models with the 4-point severity scoring, except that in the bifactor (Table S13) and hierarchical (Table S14) models, SAD was no longer significantly associated with the AAp or LPA-specific/lower-level factors.

Age and gender moderation.—Across most models, younger participants had higher levels of the general, AAp and Aar factors and female gender predicted higher AAp and LPA (Tables S6–S8). However, there was little evidence of age or gender moderation of associations with diagnostic variables. In the correlated factors model, the associations

between GAD and AAr and between SAD and LPA were stronger in younger participants (Table S16). With FDR correction, there was no significant age moderation in the bifactor or hierarchical models (Tables S17–18). With FDR correction, there was no significant gender moderation in any model (Tables S19–21).

Discussion

Three complementary ways of representing internalizing symptoms dimensions (correlated factor, S-1 bifactor, and hierarchical models) found converging evidence for both general and specific internalizing dimensions, which were differentially associated with history of, or current, depressive and anxiety disorders. The latter indicates that these scales have distinct relations to aspects of psychopathology.

Structure of Internalizing Psychopathology

Correlated factor, bifactor, and hierarchical models all had good fit, and there were strong correlations between corresponding factor scores across models, indicating that the three models capture similar constructs. Importantly, factors in all models had good reliability. Thus, we view these models not as competitors but rather as three viable tools that can optimally address different research questions and provide converging evidence to test the robustness of results across alternative approaches.

Moreover, they provided converging evidence that, in this sample, MASQ LI items were central to the general internalizing factor, as defined with the four dimensions assessed in the current study, as these items loaded exclusively on the general factors in the bifactor and hierarchical models, with no evidence of a residual LI factor in either case, and near perfect correlations of the general factors with the LI factor in the correlated factors model. Thus, the general factors in both models were largely defined by the LI items. Likewise, in the correlated factor model, the LI factor had the strongest correlations with the other factors. The items in this subscale are somewhat mixed but include a focus on feeling slowed down, bored, or apathetic. These items appear to capture an aspect of anhedonia that has been termed motivational anhedonia (Treadway & Zald, 2011), hypothesized to reflect decreased motivation, distinct from a reduction in the capacity to experience pleasure. As such, the LI subscale may primarily capture the anticipatory or “wanting” component of reward processing, distinguished from the consummatory or “liking” component (primarily captured by the LPA subscale), which are behaviorally and neurally dissociable (e.g., Husain & Roiser, 2018; Nguyen et al., 2021).

Our results suggest that, over the span of symptoms captured by the questionnaires in the present study, these symptoms of motivational anhedonia are shared across internalizing dimensions. This is consistent with models that have implicated anhedonia/apathy as a transdiagnostic factor across multiple psychiatric and neurological conditions, associated with dysfunction in neural mechanisms supporting effort-based decision-making for reward (e.g., Husain & Roiser, 2018). If so, this would differ somewhat from the tripartite model, which conceptualized the general factor as distress (e.g., Clark & Watson, 1991; Joiner, 1996; Teachman et al., 2007). However, given the mix of items in the MASQ LI subscale, it may assess a mix of motivational anhedonia and distress. It is important to note that

the factor structure in a given dataset is entirely dependent on the measures included in any given model and relations among elevations on those factors (which may differ across populations), as they simply represent the covariance of the indicators. Thus, in models including other types of internalizing symptoms, beyond the four dimensions captured here, or in different populations, the general factor may not capture the same constructs.

Associations with Diagnostic Variables

As hypothesized, history of MDD/PDD, SAD and panic attacks (and GAD when tested individually, but not controlling for all other disorders) were all associated with the general internalizing factor, in both the bifactor and hierarchical models. This suggests that the general factors in both models are indeed capturing internalizing psychopathology shared across these depressive and anxiety disorders. The robust association of history of MDD/PDD with the general factor is particularly notable, as few participants had current depression (6% with at least one criterial symptom vs. 45% lifetime history). This association is consistent with evidence that individuals who no longer meet criteria for a depressive disorder nonetheless often continue to experience elevated symptoms (e.g., Judd, 2012) and suggests that although motivational anhedonia/distress symptoms may remain elevated, positive affect may not remain low (as discussed below). When all diagnostic variables were tested together to control for comorbidity, GAD was no longer associated with the general factor, while MDD/PDD, SAD and panic attacks remained significant. This may be because GAD is highly comorbid with the other disorders: only 11% of those with a history of at least the criterial symptom of GAD had no history of criterial symptoms on any other disorder (vs. 30% for MDD/PDD, 36% for SAD and 17% for panic attacks, see Table S2), consistent with prior research (e.g., Newman et al., 2013).

Second, our hypothesis that MDD/PDD would additionally be associated with LPA-specific symptoms was not supported for lifetime history, but was supported, albeit with a small effect size, for current MDD/PDD in the bifactor model (and without FDR correction in the hierarchical model). LPA may be more of a state effect of current depression, whereas motivational anhedonia may be more trait-like in those with a history of depression. Future research with higher levels of current MDD/PDD is needed to further test this possibility. It is also possible that previously observed associations between depression and LPA were driven at least in part by general internalizing, rather than being specific to LPA.

Third, as hypothesized, GAD was associated with anxious apprehension specific factors in both the bifactor and hierarchical models, even after controlling for all other diagnoses. Indeed, these associations were among the strongest, second only to MDD/PDD with the general factors. This finding is perhaps not surprising given that anxious apprehension (worry) is the defining symptom of GAD and that GAD tends to be quite chronic in the absence of effective treatment (Newman et al., 2013). On the other hand, some have suggested that GAD reflects general distress with little specificity (e.g., Watson, 2005). The present results are not consistent with this view and rather suggest that there is something specific about worry that is elevated in those with a history of GAD. We posit that this anxious-apprehension-specific factor may reflect a tendency to engage in repetitive negative

thinking that is shared with rumination, potentially with a future orientation component that is specific to worry (Hur et al., 2017; Taylor & Snyder, 2021).

Fourth, contrary to our hypothesis, although panic attacks were associated with anxious arousal in the correlated factors model, the bifactor and hierarchical models showed that this association was due to the general factor, not what was specific to anxious arousal. Although the MASQ anxious arousal items overlap with the somatic/bodily symptoms that typically occur *during* panic attacks (e.g., racing heart, difficulty breathing, shaking), these results suggest that those with a history of panic attacks do not necessarily experience sustained increased anxious arousal outside the context of the panic attacks. Evidence is mixed regarding whether autonomic arousal is higher in individuals with panic disorder who are not currently having a panic attack (Johnson et al., 2014). Future research assessing *peak* intensity in addition to overall frequency of anxious arousal symptoms would clarify this point. However, the lack of association between panic attacks and factors specific to anxious arousal in the present study should be interpreted with caution, as the present sample did not have a prevalence sufficiently high to test associations with panic *disorder* and thus tested associations with panic attacks that occurred in any context. Future research in samples with higher prevalence of panic disorder is needed to address this issue. In addition, many participants may have had only a few prior panic attacks, rather than more severe, repeated panic attacks. However, the associations of panic attacks with the general internalizing factors suggest that they are reliably associated with current internalizing symptoms, just not those specific to anxious arousal.

Finally, we hypothesized that SAD would be associated with both specific LPA and anxious apprehension. This hypothesis was supported across both the bifactor and hierarchical models. The association of LPA with SAD but not GAD or panic attacks is consistent with evidence that, unique among the anxiety disorders, LPA is involved in the etiology and maintenance of SAD, and that adults with SAD have social anhedonia (Richey et al., 2019). Our results were also consistent with evidence that anxious apprehension is elevated in SAD, albeit not to the same extent as in GAD (Olatunji et al., 2010). It is likely that those with SAD worry primarily about social situations, versus those with GAD who worry about multiple topics (e.g., Olatunji et al., 2010), but this potential distinction cannot be ascertained in the current study because the PSWQ focuses mainly on the frequency and uncontrollability of worry rather than the focus of worry. Future research probing the contents of worries may further differentiate SAD and GAD on this dimension.

Limitations and Future Directions

In addition to those noted above, the present study had a number of limitations that should be addressed in future research. First, although the combined sample was reasonably large, was geographically diverse, and spanned early through middle adulthood, it is not representative of the US population as a whole, nor certainly of the world population. Although we did not find evidence of gender moderation or age moderation within the age range of the present sample, future research is needed to test the generalizability of the findings to youth and older adults, populations with different ethnicities, socioeconomic statuses, and cultures, and in clinical and treatment-seeking samples.

Second, there were some limitations due to the nature of the secondary data analyzed and the need to harmonize data sets that were originally collected for other primary analyses. Importantly, because, as is typical, data collection used skip logic, such that the interview for a disorder continued only if its criterial symptom(s) were endorsed, we could not assess disorder severity in a fully dimensional manner, which may have provided more power and sensitivity. In addition, because one site ascertained only current, and not remitted, anxiety disorders, the lifetime anxiety analyses more weighted toward current diagnoses. However, we also analyzed current diagnoses only, for which all sites used the same past six months assessment. Although most results regarding lifetime history and current diagnoses converged, there were also interesting differences in that current depressive disorders were associated with (past week) low positive affect, whereas lifetime history of depressive disorders was not. Questionnaires and diagnostic interviews also covered different time frames from one another, with SCIDs covering lifetime history and the past six months (for current diagnoses), the PSWQ asking about general (trait) tendencies, and the MASQ asking about symptoms during the past week. Future research would benefit from assessing diagnostic status and symptom dimensions over the same timeframe.

We also did not have interview data across sites for specific phobias or other disorders that are considered to fall within the internalizing spectrum in some models (e.g., PTSD, OCD, and eating disorders in the HiTOP model; Kotov et al., 2021). Future research is needed to test internalizing symptom dimension factors in relation to a wider range of diagnoses. Likewise, our symptom measures were not exhaustive of all possible internalizing symptoms. Although the MASQ and PSWQ are useful, and they capture the dimensions proposed by Heller and colleagues, which were the focus of the current investigation, future research would benefit from including measures of other internalizing symptoms. The factor model and associations might shift with inclusion of symptoms such as negative affectivity, as results of such analyses, and even suitability of various analytic strategies, depend substantially on the base rates and pairwise coincidence of individual symptoms in the sample analyzed (Fried et al., 2021; Watts et al., 2020; Watts et al., 2019). Scope of symptoms, in turn, is critical for translation to clinical practice. In addition, future research is needed to determine how dimensional models can be used for within-person inference to support clinical decision-making.

Finally, although examining these factors in relation to DSM diagnostic variables was an important step in characterizing them, there is some irony in using a diagnostic system that is widely believed to be deeply flawed as a way of evaluating dimensional approaches (Yee et al., 2015). The present study employed SCID-based diagnoses as a way of bridging recent dimensional approaches and traditional categorical psychiatric diagnoses as well as to inform interpretation of past research. Future research is needed to examine the nomological networks and neural and environmental correlates of these empirically-derived symptom dimensions, and research has begun to test the bifactor internalizing model presented here in relation to neurobiology (e.g., Banich et al., 2020).

Conclusions

In sum, we found that internalizing psychopathology symptoms are well-represented as having both a shared, general component, and distinct anxious arousal, anxious apprehension, and low positive affect components, consistent with Heller and colleagues' model (e.g., Nitschke et al., 2001). Bifactor and hierarchical models performed equally well and yielded highly similar results. All disorders assessed were associated with elevated general internalizing, and after accounting for the general internalizing factor, each disorder had a unique profile of associations with the specific symptom dimensions. Altogether, this pattern suggests that traditional DSM diagnoses are an admixture of multiple symptom dimensions which partially overlap and partially differentiate them.

These findings contribute to growing evidence that DSM-defined disorders are too heterogenous within disorders and too overlapping across disorders to provide an adequate basis for etiological and treatment research and that dimensional approaches more accurately characterize psychopathology (e.g., Kotov et al., 2021; Ruscio, 2019). We thus believe it is a mistake to rely on disorder diagnoses (or questionnaires aligned with DSM disorder criteria) as the "ground floor" in hierarchical models (or indicators in any latent variable model), as this bakes into the models the very problems with traditional diagnostic categories these new taxonomies attempt to remedy. New dimensional measures that assess broader arrays of symptoms across different dimensions hold promise for more empirically-driven assessment (e.g., Simms et al., 2021; Watson et al., 2021). We suggest that dimensional approaches that capture both shared and specific internalizing symptom facets can provide a useful alternative to advance clinical science.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

Portions of this research were supported by R01MH105501 (to Banich and Hankin, MPIs), T32MH016880 (for Smolker), R21DA14111 and T32MH19554 (to Miller), R01MH61358 (to Heller), R01MH117131 (to Kaiser), and P50MH079485 (to Banich).

References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.).

Bader M, & Moshagen M (2021). No probifactor model fit index bias, but a propensity toward selecting the best model. Journal of Abnormal Psychology, in press.

Banich MT, Smith LL, Smolker HR, Hankin BL, Silton RL, Heller W, & Snyder HR (2020). Common and specific dimensions of internalizing disorders are characterized by unique patterns of brain activity on a task of emotional cognitive control. International Journal of Psychophysiology, 151, 80–93.

Barlow DH (1991). Disorders of emotion. Psychological Inquiry, 2, 58–71.

Bredemeier K, Spielberg JM, Silton RL, Berenbaum H, Heller W, & Miller GA (2010). Screening for depressive disorders using the Mood and Anxiety Symptoms Questionnaire Anhedonic Depression Scale: A receiver-operating characteristic analysis. Psychological Assessment, 22, 702–710.

Brown TA, Antony MM, & Barlow DH (1992). Psychometric properties of the Penn state Worry Questionnaire in a clinical anxiety disorders sample. Behaviour Research and Therapy, 30, 33–37.

Brown TA, & Barlow DH (2005). Dimensional versus categorical classification of mental disorders in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders and beyond: Comment on the special section. Journal of Abnormal Psychology, 114, 551–556

Burdwood EN, Infantolino ZP, Crocker LD, Spielberg JM, Banich MT, Miller GA, & Heller W (2016). Resting-state functional connectivity differentiates anxious apprehension and anxious arousal. Psychophysiology, 53, 1451–1459.

Bystritsky A (2006). Treatment-resistant anxiety disorders. Molecular Psychiatry, 11, 805–814.

Chen FF (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural Equation Modeling, 14, 464–504.

Chen FF, Hayes A, Carver CS, Laurenceau J-P, & Zhang Z (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. Journal of Personality, 80, 219–251.

Cheung GW, & Rensvold RB (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling, 9, 233–255.

Clark LA, & Watson D (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. Journal of Abnormal Psychology, 100, 316–336.

Coghill D, & Sonuga-Barke EJ (2012). Categories versus dimensions in the classification and conceptualization of child and adolescent mental disorders: Implications of recent empirical study. Journal of Child Psychology and Psychiatry, and Allied Disciplines, 53, 469–489

Conway CC, Forbes MK, Forbush KT, Fried EI, Hallquist MN, Kotov R, Mullins-Sweatt SN, Shackman AJ, Skodol AE, South SC, Sunderland M, Waszczuk MA, Zald DH, Afzali MH, Bornovalova MA, Carragher N, Docherty AR, Jonas KG, Krueger RF, … Eaton NR (2019). A Hierarchical taxonomy of psychopathology can transform mental health research. Perspectives on Psychological Science, 14, 419–436.

Conway CC, Forbes MK, & South SC (2021). A hierarchical taxonomy of psychopathology (HiTOP) primer for mental health researchers. Current Directions in Psychological Science, 10, 236–258.

Crocker LD, Heller W, Spielberg JM, Warren SL, Bredemeier K, Sutton BP, Banich MT, & Miller GA (2012). Neural mechanisms of attentional control differentiate trait and state negative affect. Frontiers in Psychology, 3, 298.

Cuthbert BN, Miller GA, Sanislow C, & Vaidyanathan U (in press). The Research Domain Criteria Project: Integrative translation for psychopathology. Blaney PH & Krueger RF (Eds.), Oxford Textbook of Psychopathology, Third Edition. New York: Oxford University Press.

den Hollander-Gijsman ME, Wardenaar KJ, de Beurs E, van der Wee NJA, Mooijaart A, van Buuren S, & Zitman FG (2012). Distinguishing symptom dimensions of depression and anxiety: An integrative approach. Journal of Affective Disorders, 136, 693–701.

Eid M (2020). Multi-faceted constructs in abnormal psychology: Implications of the difactor S1 model for individual clinical assessment, Journal of Abnormal Child Psychology, 48, 895–900.

Engels AS, Heller W, Mohanty A, Herrington JD, Banich MT, Webb AG, & Miller GA (2007). Specificity of regional brain activity in anxiety types during emotion processing. Psychophysiology, 44, 352–363.

Engels AS, Heller W, Spielberg JM, Warren SL, Sutton BP, Banich MT, & Miller GA (2010). Co-occurring anxiety influences patterns of brain activity in depression. Cognitive, Affective & Behavioral Neuroscience, 10, 141–156.

Fassett-Carman AN, DiDomenico GE, von Steiger J, & Snyder HR (2020). Clarifying stress-internalizing associations: Stress frequency and appraisals of severity and controllability are differentially related to depression-specific, anxiety-specific, and transdiagnostic internalizing factors. Journal of Affective Disorders, 260, 638–645.

Fava GA, Ruini C, & Belaise C (2007). The concept of recovery in major depression. Psychological Medicine, 37, 307–317.

Fried EI, Greene AL, & Eaton NR (2021). The p factor is the sum of its parts, for now. World Psychiatry, 20, 69–70.

Greene AL, Eaton NR, Li K, Forbes MK, Krueger RF, Markon KE, Waldman ID, Cicero DC, Conway CC, Docherty AR, Fried EI, Ivanova MY, Jonas KG, Latzman RD, Patrick CJ, Reininghaus U, Tackett JL, Wright AGC, & Kotov R (2019). Are fit indices used to test psychopathology structure biased? A simulation study. Journal of Abnormal Psychology, 128, 740–764.

Grisanzio KA, Goldstein-Piekarski AN, Wang MY, Rashed Ahmed AP, Samara Z, & Williams LM (2018). Transdiagnostic symptom clusters and associations with brain, behavior, and daily function in mood, anxiety, and trauma disorders. JAMA Psychiatry, 75, 201–209.

Guha A, Spielberg JM, Lake J, Popov T, Heller W, Yee CM, & Miller GA (2020). Effective connectivity between Broca's area and amygdala as a mechanism of top-down control in worry. Clinical Psychological Science, 8, 84–98. [PubMed: 32983628]

Hankin BL, Snyder HR, Gulley LD, Schweizer TH, Bijttebier P, Nelis S, Toh G, & Vasey MW (2016). Understanding comorbidity among internalizing problems: Integrating latent structural models of psychopathology and risk mechanisms. Development and Psychopathology, 28, 987–1012. [PubMed: 27739389]

Heller W (1990). The neuropsychology of emotion: Developmental patterns and implications for psychopathology. In Stein NL, Leventhal B, & Trabasso TR (Eds.), Psychological and Biological Approaches to Emotion (pp. 67–211). Taylor & Francis.

Heller W, & Nitschke JB (1998). The puzzle of regional brain activity in depression and anxiety: The importance of subtypes and comorbidity. Cognition and Emotion, 12, 421–447.

Heller W, Nitschke JB, Etienne MA, & Miller GA (1997). Patterns of regional brain activity differentiate types of anxiety. Journal of Abnormal Psychology, 106, 376–385. [PubMed: 9241939]

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

Hur J, Heller W, Kern JL, & Berenbaum H (2017). A bi-factor approach to modeling the structure of worry and rumination. Journal of Experimental Psychopathology, 8, 252–264.

Husain M, & Roiser JP (2018). Neuroscience of apathy and anhedonia: A transdiagnostic approach. Nature Reviews Neuroscience, 19, 470–484. [PubMed: 29946157]

Hyman SE (2010). The diagnosis of mental disorders: the problem of reification. Annual Review of Clinical Psychology, 6, 155–179.

Insel TR, & Cuthbert BN (2009). Endophenotypes: bridging genomic complexity and disorder heterogeneity. Biological Psychiatry, 66, 988–989. [PubMed: 19900610]

Joiner TE (1996). A confirmatory factor-analytic investigation of the tripartite model of depression and anxiety in college students. Cognitive Therapy and Research, 20, 521–539.

Johnson PL, Federici LM, & Shekhar A (2014). Etiology, triggers and neurochemical circuits associated with unexpected, expected, and laboratory-induced panic attacks. Neuroscience and Biobehavioral Reviews, 46, 429–454. [PubMed: 25130976]

Judd LL (2012). Dimensional paradigm of the long-term course of unipolar major depressive disorder. Depression and Anxiety, 29, 167–171. [PubMed: 22511337]

Kendler KS, Zachar P & Craver C (2011). What kinds of things are psychiatric disorders? Psychological Medicine, 41, 1143–1150. [PubMed: 20860872]

Kertz SJ, Lee J, & Björgvinsson T (2014). Psychometric properties of abbreviated and ultra-brief versions of the Penn State Worry Questionnaire. Psychological Assessment, 26, 1146–1154. [PubMed: 24932640]

Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, & Wittchen HU (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. International Journal of Methods in Psychiatric Research, 21, 169–184. [PubMed: 22865617]

Koch T, Holtmann J, Bohn J, & Eid M (2018). Explaining general and specific factors in longitudinal, multimethod, and bifactor models: Some caveats and recommendations. Psychological Methods, 23, 505–523. [PubMed: 28737413]

Kotov R, Krueger RF, Watson D, Achenbach TM, Althoff RR, Bagby RM, Brown TA, Carpenter WT, Caspi A, Clark LA, Eaton NR, Forbes MK, Forbush KT, Goldberg D, Hasin D, Hyman SE, Ivanova MY, Lynam DR, Markon K, … Zimmerman M (2017). The Hierarchical Taxonomy

of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. Journal of Abnormal Psychology, 126, 454–477. [PubMed: 28333488]

Kozak MJ, & Cuthbert BN (2016). The NIMH Research Domain Criteria Initiative: Background, issues, and pragmatics. Psychophysiology, 53, 286–297. [PubMed: 26877115]

Krueger RF (1999). The structure of common mental disorders. Archives of General Psychiatry, 56, 921–926. [PubMed: 10530634]

Kumari V, Mitterschiffthaler MT, Teasdale JD, Malhi GS, Brown RG, Giampietro V, Brammer MJ, Poon L, Simmons A, Williams SCR, Checkley SA, & Sharma T (2003). Neural abnormalities during cognitive generation of affect in treatment-resistant depression. Biological Psychiatry, 54, 777–791. [PubMed: 14550677]

Lahey BB, Krueger RF, Rathouz PJ, Waldman ID, & Zald DH (2017). Validity and utility of the general factor of psychopathology. World Psychiatry, 16, 142–144. [PubMed: 28498590]

Lin A, Yung AR, Wigman JTW, Killackey E, Baksheev G, & Wardenaar KJ (2014). Validation of a short adaptation of the Mood and Anxiety Symptoms Questionnaire (MASQ) in adolescents and young adults. Psychiatry Research, 215, 778–783. [PubMed: 24387880]

Loh JMI, Schutte NS, & Thorsteinsson EB (2014). Be happy: The role of resilience between characteristic affect and symptoms of depression. Journal of Happiness Studies,1125–1138.

Mansolf M, & Reise SP (2017). When and why the second-order and bifactor models are distinguishable. Intelligence, 61,120–129.

Meyer TJ, Miller ML, Metzger RL, & Borkovec TD (1990). Development and validation of the Penn State Worry Questionnaire. Behaviour Research and Therapy, 28, 487–495. [PubMed: 2076086]

Miller GA, Rockstroh BS, Hamilton HK, & Yee CM (2016). Psychophysiology as a core strategy in RDoC. Psychophysiology, 53, 410–414. [PubMed: 26877134]

Moffitt TE, Harrington H, Caspi A, Kim-Cohen J, Goldberg D, Gregory AM, & Poulton R (2007). Depression and generalized anxiety disorder: Cumulative and sequential comorbidity in a birth cohort followed prospectively to age 32 years. Archives of General Psychiatry, 64, 651–660. [PubMed: 17548747]

Newman MG, Llera SJ, Erickson TM, Przeworski A, & Castonguay LG (2013). Worry and generalized anxiety disorder: A review and theoretical synthesis of evidence on nature, etiology, mechanisms, and treatment. Annual Review of Clinical Psychology, 9, 275–297.

Nguyen D, Naffziger EE, & Berridge KC (2021). Positive affect: Nature and brain bases of liking and wanting. Current Opinion in Behavioral Sciences, 39, 72–78. [PubMed: 33748351]

Nitschke JB, & Heller W (2005). Distinguishing neural substrates of heterogeneity among anxiety disorders. International Review of Neurobiology, 67, 1–42. [PubMed: 16291018]

Nitschke JB, Heller W, Imig JC, McDonald RP, & Miller GA (2001). Distinguishing dimensions of anxiety and depression. Cognitive Therapy and Research, 25, 1–22.

Nitschke JB, Heller W, Palmieri PA, & Miller GA (1999). Contrasting patterns of brain activity in anxious apprehension and anxious arousal. Psychophysiology, 36, 628–637. [PubMed: 10442031]

Olatunji BO, Wolitzky Taylor KB, Sawchuk CN, & Ciesielski BG (2010). Worry and the anxiety disorders: A meta-analytic synthesis of specificity to GAD. Applied and Preventive Psychology, 14, 1–24.

Osório FL, Loureiro SR, Hallak JEC, Machado-de-Sousa JP, Ushirohira JM, Baes CVW, Apolinario TD, Donadon MF, Bolsoni LM, Guimarães T, Fracon VS, Silva-Rodrigues APC, Pizeta FA, Souza RM, Sanches RF, Dos Santos RG, Martin-Santos R, & Crippa JAS (2019). Clinical validity and interrater and test-retest reliability of the Structured Clinical Interview for DSM-5 - Clinician Version (SCID-5-CV). Psychiatry and Clinical Neurosciences, 73, 754–760. [PubMed: 31490607]

Pine DS, Cohen E, Cohen P, & Brook J (1999). Adolescent depressive symptoms as predictors of adult depression: Moodiness or mood disorder? American Journal of Psychiatry, 156, 133–135. [PubMed: 9892310]

Richey JA, Brewer JA, Sullivan-Toole H, Strege MV, Kim-Spoon J, White SW & Ollendick TH (2019). Sensitivity shift theory: A developmental model of positive affect and motivational deficits in social anxiety disorder. Clinical Psychology Review, 72, 101756. [PubMed: 31351312]

Rodriguez A, Reise SP, & Haviland MG (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. Psychological Methods, 21, 137–150. [PubMed: 26523435]

Ruscio AM (2019). Normal versus pathological mood: Implications for diagnosis. Annual Review of Clinical Psychology, 15, 179–205.

Sharp PB, Miller GA, & Heller W (2015). Transdiagnostic dimensions of anxiety: Neural mechanisms, executive functions, and new directions. International Journal of Psychophysiology, 98, 365–377. [PubMed: 26156938]

Silton RL, Heller W, Engels AS, Towers DN, Spielberg JM, Edgar JC, Sass SM, Stewart JL, Sutton BP, Banich MT, & Miller GA (2011). Depression and anxious apprehension distinguish frontocingulate cortical activity during top-down attentional control. Journal of Abnormal Psychology, 120, 272–285. [PubMed: 21553941]

Simms LJ, Prisciandaro JJ, Krueger RF, & Goldberg DP (2012). The structure of depression, anxiety and somatic symptoms in primary care. Psychological Medicine, 42, 15–28. [PubMed: 21682948]

Simms LJ, Wright AG, Cicero D, Kotov R, Mullins-Sweatt SN, Sellbom M, ... & Zimmermann J (2021). Development of measures for the Hierarchical Taxonomy of Psychopathology (HiTOP): A collaborative scale development project. Assessment, 29, 3–16. [PubMed: 34013772]

Starcevic V, Berle D, Milicevic D, Hannan A, Lamplugh C, & Eslick GD (2007). Pathological worry, anxiety disorders and the impact of co-occurrence with depressive and other anxiety disorders. Journal of Anxiety Disorders, 21, 1016–1027. [PubMed: 17270391]

Taylor MM, & Snyder HR (2021). Repetitive negative thinking shared across rumination and worry predicts symptoms of depression and anxiety. Journal of Psychopathology and Behavioral Assessment, 43, 904–915.

Teachman BA, Siedlecki KL, & Magee JC (2007). Aging and symptoms of anxiety and depression: Structural invariance of the tripartite model. Psychology and Aging, 22, 160–170. [PubMed: 17385992]

Treadway MT, & Zald DH (2011). Reconsidering anhedonia in depression: Lessons from translational neuroscience. Neuroscience & Biobehavioral Reviews, 35, 537–555. [PubMed: 20603146]

Vandeleur CL, Fassassi S, Castelao E, Glaus J, Strippoli M-PF, Lasserre AM, Rudaz D, Gebreab S, Pistis G, Aubry J-M, Angst J, & Preisig M (2017). Prevalence and correlates of DSM-5 major depressive and related disorders in the community. Psychiatry Research, 250, 50–58. [PubMed: 28142066]

Verstraeten K, Vasey MW, Raes F, & Bijttebier P (2009). Temperament and risk for depressive symptoms in adolescence: Mediation by rumination and moderation by effortful control. Journal of Abnormal Child Psychology, 37, 349–361. [PubMed: 19107592]

Vollebergh WA, Iedema J, Bijl RV, de Graaf R, Smit F, & Ormel J (2001). The structure and stability of common mental disorders: The NEMESIS study. Archives of General Psychiatry, 58, 597–603. [PubMed: 11386990]

Wardenaar KJ, Giltay EJ, van Veen T, Zitman FG, & Penninx BWJH (2012). Symptom dimensions as predictors of the two-year course of depressive and anxiety disorders. Journal of Affective Disorders, 136, 1198–1203. [PubMed: 22177741]

Warren SL, Crocker LD, Spielberg JM, Engels AS, Banich MT, Sutton BP, Miller GA, & Heller W (2013). Cortical organization of inhibition-related functions and modulation by psychopathology. Frontiers in Human Neuroscience,7.

Watson D (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for DSM-V. Journal of Abnormal Psychology, 114, 522–536. [PubMed: 16351375]

Watson D, Clark LA, & Carey G (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. Journal of Abnormal Psychology, 97, 346–353. [PubMed: 3192830]

Watson D, Forbes MK, Levin-Aspenson HF, Ruggero CJ, Kotelnikova Y, Khoo S, ... & Kotov R (2021). The development of preliminary HiTOP internalizing spectrum scales. Assessment, 29, 17–33. [PubMed: 33794667]

Watson D, & Naragon-Gainey K (2010). On the specificity of positive emotional dysfunction in psychopathology: Evidence from the mood and anxiety disorders and schizophrenia/schizotypy. Clinical Psychology Review, 30, 839–848. [PubMed: 19939530]

Watson D, Weber K, Assenheimer JS, Clark LA, Strauss ME, & McCormick RA (1995). Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. Journal of Abnormal Psychology, 104, 3–14. [PubMed: 7897050]

Watts AL, Lane SP, Bonifay W, Steinley D, & Meyer FA (2020). Building theories on top of, and not independent of, statistical models: The case of the p-factor. Psychological inquiry, 31, 310–320. [PubMed: 33510565]

Watts AL, Poore HE, & Waldman ID (2019). Riskier tests of the validity of the bifactor model of psychopathology. Clinical Psychological Science, 7, 1285–1303.

Widiger TA, & Samuel DB (2005). Diagnostic categories or dimensions? A question for the Diagnostic and Statistical Manual Of Mental Disorders--fifth edition. Journal of Abnormal Psychology, 114, 494–504 [PubMed: 16351373]

Wilcox HC, & Anthony JC (2004). Child and adolescent clinical features as forerunners of adult-onset major depressive disorder: Retrospective evidence from an epidemiological sample. Journal of Affective Disorders, 82, 9–20. [PubMed: 15465572]

Williamson RJ, Neale BM, Sterne A, Prince M, & Sham P (2005). The value of four mental health self-report scales in predicting interview-based mood and anxiety disorder diagnoses in sibling pairs. Twin Research and Human Genetics, 8, 101–107. [PubMed: 15901472]

Yee CM, Javitt DC, & Miller GA (2015). Replacing DSM categorical analyses with dimensional analyses in psychiatry research. JAMA Psychiatry, 72, 1159–1160. [PubMed: 26559005]

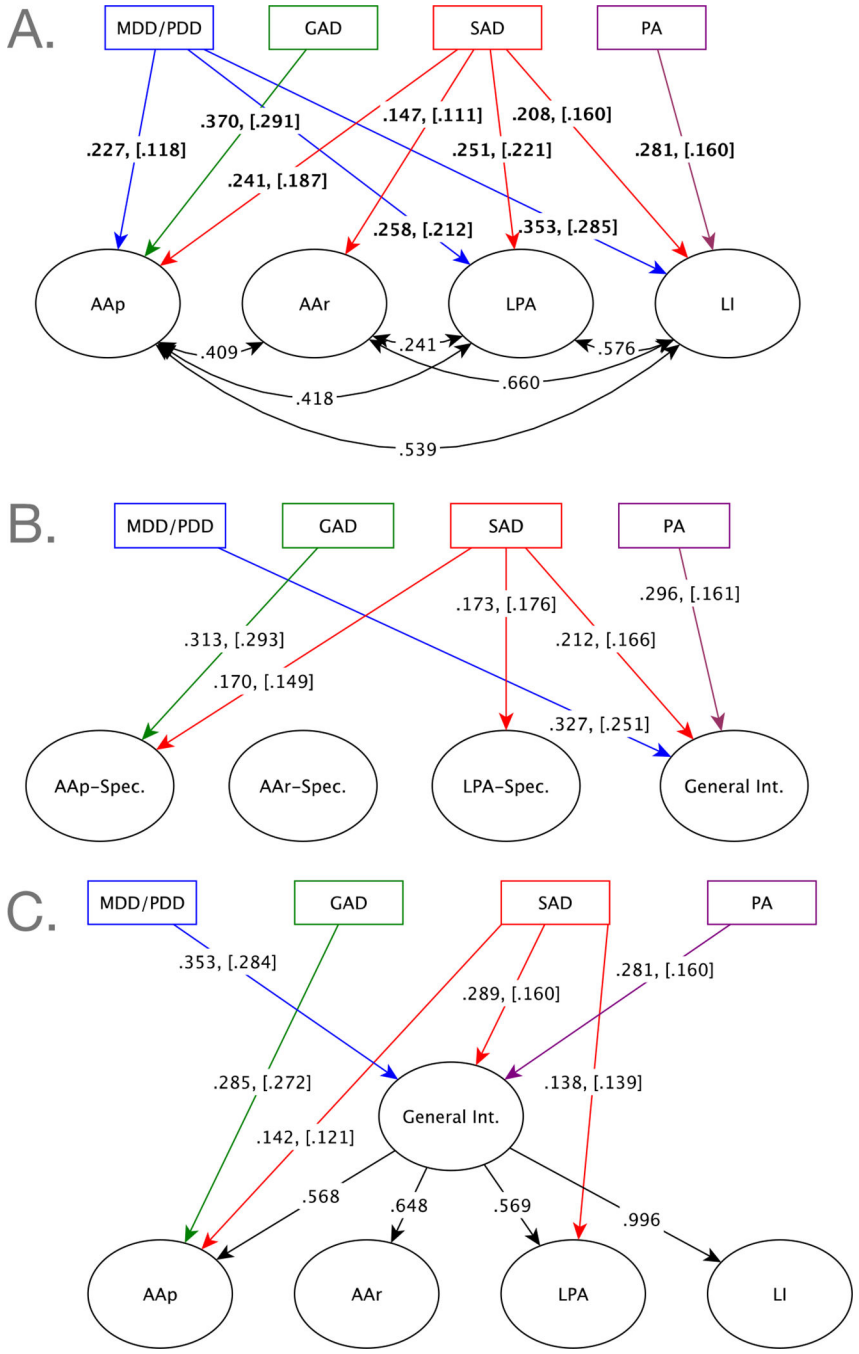

Figure 1. SEMs with lifetime history SCID predictors.

The first number on each line is the standardized coefficient from the model with that disorder alone, and the following number in [brackets] is from the model with all disorders simultaneously. Only paths significant with FDR correction in both analyses are shown here; for full results, see Supplemental Materials Tables S6–S8. A. Correlated factors model. B. S-1 bifactor model. C. Hierarchical model.

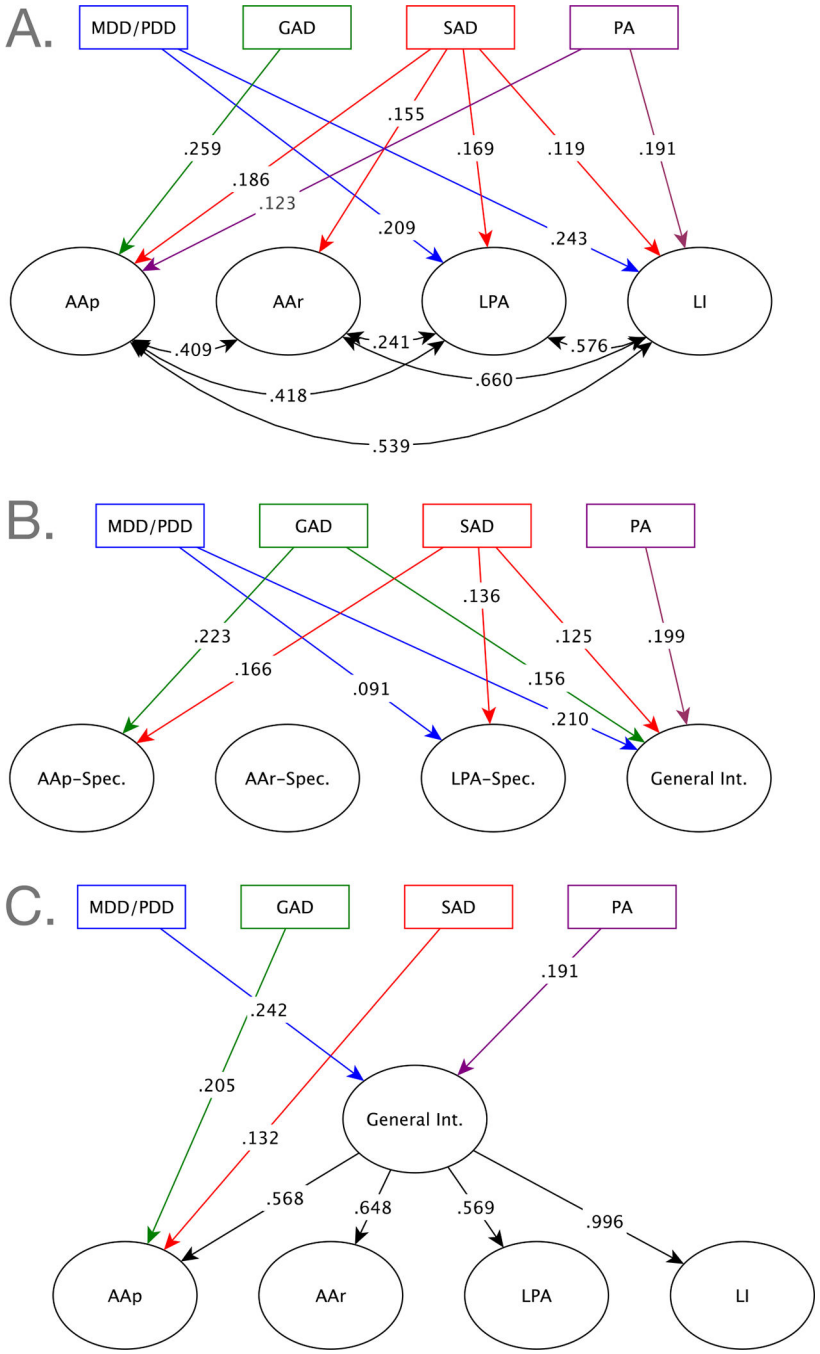

Figure 2. SEMs with all current SCID predictors simultaneously.

A. Correlated factors model. B. S-1 bifactor model. C. Hierarchical model.

Table 1

Descriptive statistics

Questionnaires	Mean		SD		Range		α	
PSWQ	50.17		14.61		18–80		.945	
MASQ AA	23.54		6.71		15–54		.836	
MASQ LPA	40.86		12.00		14–70		.927	
MASQ LI	15.27		5.20		8–37		.809	
Diagnostic Coding	3 (full criteria)		2 (−1 symptom)		1 (− 2+ symptoms)		0 (no threshold hallmark symptom)	
	Lifetime	Current	Lifetime	Current	Lifetime	Current	Lifetime	Current
MDD/PDD	30.2%	4.3%	5.5%	0.2%	10.2%	1.7%	54.0%	93.8%
	(*n* =127)	(*n* =18)	(*n* =23)	(*n* =1)	(*n* =43)	(*n* =7)	(*n* =227)	(*n* =394)
GAD	16.4%	11.2%	1.2%	1.2%	0.2%	0.2%	82.1%	87.4%
	(*n* =69)	(*n* =47)	(*n* =5)	(*n* =5)	(*n* =1)	(*n* =1)	(*n* =345)	(*n* =367)
SAD	16.9%	12.6%	6.0%	5.5%	10.2%	7.6%	66.9%	74.0%
	(*n* =71)	(*n* =54)	(*n* =25)	(*n* =23)	(*n* =43)	(*n* =32)	(*n* =281)	(*n* =311)
PD	4.3%	2.4%	0%	0%	0%	0%	95.7%	97.6%
	(*n* =18)	(*n* =10)					(*n* = 402)	(*n* =420)
PA	19.8%	11.2%	1.9%	1.0%	0.9%	0%	77.4%	87.9%
	(*n* =83)	(*n* =47)	(*n* =8)	(*n* =4)	(*n* =4)		(*n* =325)	(*n* =369)

Note. PSWQ= Penn State Worry Questionnaire; MASQ= Mood and Anxiety Symptom Questionnaire; AA= anxious arousal; LPA= low positive affect; LI= loss of interest; MDD/PDD= major depressive disorder/persistent depressive disorder; GAD= generalized anxiety disorder; SAD= social anxiety disorder; PD= panic disorder; PA= panic attacks.

Table 2

Manifest Variable Bivariate Correlations for Lifetime Diagnostic Variables and Questionnaires

	PSWQ	MASQ-AA	MASQ-LPA	MASQ-LI	MDD/PDD	GAD	SAD	PA	Age
PSWQ	-								
MASQ-AA	.357**	-							
MASQ-LPA	.407**	.198**	-						
MASQ-LI	.466**	.558**	.455**	-					
MDD/PDD	.171*	.032	.246**	.229**	-				
GAD	.395**	.121*	.294**	.163**	.242**	-			
SAD	.273**	.173**	.198**	.214**	.063	.159**	-		
PA	.275**	.202**	.203**	.258**	.236**	.359**	.180**		
Age	−.329**	−.341**	−.142*	−.299**	.175**	−.129*	−.117*	−.202**	
Gender	.083	−.138*	.079	−.115*	.096	.107*	.033	.060	.119*

Note. Correlations involving diagnostic variables and/or gender are Spearman's rho. Correlations among questionnaires are Pearson's r. PSWQ= Penn State Worry Questionnaire; MASQ= Mood and Anxiety Symptom Questionnaire; AA= anxious arousal; LPA= low positive affect; LI= loss of interest. MDD/PDD= major depressive disorder/persistent depressive disorder; GAD= generalized anxiety disorder; SAD= social anxiety disorder; PA= panic attacks. Gender coded 1= male, 2=female.

** $p < .001$

* $p < .05$

Table 3

Model fit and factor reliability

	Model fit					**Factor reliability metrics**			
Model	Satorra-Bentler Scaled χ^2	RMSEA (90% CI)	SRMR	CFI[1]	AIC, ΔAIC		ω	H	FD
Correlated factors	2486.18 (1414), p<.001, SCR=1.11	.042 (.039–.045)	.053	.899	AIC=54442.56	LPA	.940	.943	.971
						LI	.806	.826	.924
						AAr	.869	.875	.934
						AAp	.949	.959	.978
Bifactor (S-1)	2411.23 (1373) p<.001, SCR=1.10	.042 (.039–.045)	.057	.902	AIC =54463.03 ΔAIC Vs. correlated factor=20.47	General Int.	ω =.956 ωH=.679	.930	.934
						LPA-spec	ω =.941 ωHS=.647	.893	.934
						AAr-spec	ω =.852 ωHS=.462	.768	.863
						AAp-spec	ω =.945 ωHS=.667	.911	.944
Hierarchical	2509.04 (1416), p<.001, SCR=1.11	.043 (.040–.045)	.061	.897	AIC=54464.68 ΔAIC Vs. correlated factor= 22.12	General Int.	.799	.992	.921
						LPA	.940	.943	.971
						LI	.806	.826	.923
						AAr	.868	.876	.934
					ΔAIC Vs. bifactor=1.65	AAp	.949	.959	.978

Note. The MLR estimator produces Satorra-Bentler scaled χ^2 values, where the normal-theory χ^2 statistic is divided by a scaling correction to better approximate χ^2 under normal conditions. SCR= scaling correction factor; CD= difference test scaling correction.

[1] In cases where the RMSEA of the baseline (null) model is < .158, incremental fit indices such as CFI are artificially low and thus not considered informative; instead, absolute fit indices should be used to evaluate model fit (Kenny, 2015). RMSEA for the baseline model was .129. Thus, we rely on absolute measures of fit (RMSEA and SRMR). CFI is reported for completeness.