HHS Public Access

Author manuscript

Clin Psychol Sci. Author manuscript; available in PMC 2020 May 01.

Published in final edited form as:

Clin Psychol Sci. 2019 May ; 17(3): 566–581. doi:10.1177/2167702618813989.

Decoupling Personality and Acute Psychiatric Symptoms in a Depressed Sample and a Community Sample

Jay C. Fournier,
Department of Psychiatry, University of Pittsburgh School of Medicine

Aidan Wright,
Department of Psychology, University of Pittsburgh

Jennifer L. Tackett,
Department of Psychology, Northwestern University

Amanda Uliaszek,
Department of Psychology, University of Toronto Scarborough

Paul A. Pilkonis,
Department of Psychiatry, University of Pittsburgh School of Medicine

Stephen B. Manuck,
Department of Psychology, University of Pittsburgh

R. Michael Bagby
Department of Psychiatry, University of Toronto.

Abstract

The association between depression and neuroticism is complex, but due to the difficulty in assessing neuroticism during mood episodes, the mechanisms underlying this relationship remain poorly understood. In this study, we sought to decompose neuroticism into finer-grained elements that were uncorrelated with psychiatric symptoms and to examine the incremental validity of those elements in explaining deficits in interpersonal functioning. A bifactor model with one general factor and six specific factors fit the data well in both a depressed (N=807) and a community (N=1,284) sample, and the specific factors were relatively independent of acute symptoms. Moreover, two specific factors (Angry Hostility and Self-Consciousness) accounted for incremental variance in interpersonal functioning problems in the community sample and in a subgroup of depressed participants. The results demonstrate that neuroticism can be decomposed

Correspondence concerning this article should be addressed to: Jay Fournier, Department of Psychiatry, University of Pittsburgh School of Medicine, 3811 O'Hara Street, Pittsburgh, PA 15213. fournierjc@upmc.edu.

Authors' Contributions

J.C.F developed the study concept. P.A.P., R.M.B., and S.B.M. were responsible for the design and implementation of the studies contributing data to the current study. J.C.F. performed the data analysis in consultation with A.W. All authors contributed to the interpretation of the results. J.C.F drafted the paper. All authors provided critical revisions and approved the final version of the paper for submission.

The findings reported in this article have not been published elsewhere. A portion of the findings have been presented at the Annual Meeting of the Society for Interpersonal Theory and Research (Pittsburgh, PA, 2017) and at the Annual Meeting of the Society for Research in Psychopathology (Denver, CO, 2017).

into components that are distinct from symptoms and that are incrementally associated with deficits in interpersonal functioning.

Keywords

Neuroticism; Depression; Bifactor measurement models; Interpersonal functioning; Replication

The relationship between neuroticism and depression is complex. Neuroticism is often correlated with depressive symptoms (Rosellini & Brown, 2011) and frequently identified as a vulnerability factor for the development of depressive episodes (see, Bagby, Psych, Quilty, & Ryder, 2008; Kotov, Gamez, Schmidt, & Watson, 2010). At the same time, substantial individual differences in neuroticism exist among depressed people, and these differences have consequences. Bagby and colleagues (2008), for example, observed that higher baseline neuroticism among depressed adults was associated with superior response to selective serotonin reuptake inhibitor (SSRI) medication compared to cognitive therapy (CT) (a pattern that was conceptually replicated by Fournier and colleagues [2008] using frank personality disorder diagnosis rather than neuroticism to index personality pathology). Furthermore, there is evidence that SSRIs reduce levels of neuroticism itself, independent of changes in psychiatric symptoms (Quilty, Meusel, & Bagby, 2008; Tang et al., 2009). The mechanisms through which maladaptive personality features like neuroticism affect and are affected by treatments for depression, however, are currently unknown.

The core challenge in examining the mechanisms through which neuroticism and its components impact treatment stems from the difficulty in assessing personality functioning during an acute depressive episode. Without information about an individual's personality prior to the onset of a depressive episode, it is difficult to separate the unique effects of trait-like personality functioning from state increases in psychiatric symptoms and distress when both are measured concurrently. This issue has long burdened the field (Hirschfeld et al., 1983; Watson & Clark, 1995; Widiger & Trull, 1992), and it reflects the fact that scores on personality measures can appear more pathological during a depressive episode than prior to the episode’s onset or following remission. This is due in part to the substantial conceptual and item overlap between measures of symptom severity and measures of neuroticism. Furthermore, the negative information processing biases associated with depression (Gotlib & Joormann, 2010) may color responses on self-report measures of personality and distort genuine trait-like patterns of functioning. At the same time, the fact that levels of neuroticism, and not symptom severity, predict differential response between distinct modalities of treatment, combined with findings suggesting that neuroticism itself can be differentially impacted by particular treatments over and above changes in symptoms, signals that measures of neuroticism are capturing something important for understanding underlying pathology, even when assessed during an acute depressive episode (Santor, Bagby, & Joffe, 1997). The goal of the current study was to identify components of neuroticism that can be separated from concurrent symptom severity and to examine whether those components are incrementally associated with meaningful domains of functioning over and above current symptoms.

The meaning of neuroticism has evolved from early accounts by Eysenck and colleagues who used the term to reflect a propensity for emotional instability, mediated by hyperactivity in limbic regions in response to stress (Eysenck, 1967), to include the tendency to experience negative affect more broadly and to have difficulty regulating emotional experience (e.g., Barlow, Ellard, Sauer-Zavala, Bullis, & Carl, 2014; Costa & McCrae, 1992; Ormel et al., 2013). Costa and McCrae (1992), the developers of the popular family of NEO instruments for measuring neuroticism (i.e., NEO PI, NEO PI-R and the NEO-3), suggest that it is composed of six lower-order facets: Anxiety, Angry hostility, Depression, Self-consciousness, Impulsiveness, and Vulnerability. Their proposed structure, however, has received only limited support when examined by independent researchers (Endler, Rutherford, & Denisoff, 1997). Recognizing the lack of agreement regarding the facet structure of the Five-Factor Model (FFM) of personality, Naragon-Gainey and Watson (2014) used exploratory factor analyses on a number of personality measures to derive a consensual definition of the facets of the FFM. For the Neuroticism domain, they identified three facets: Anxiety, comprised of the Anxiety, Self-consciousness, and Vulnerability facets of the NEO PI, along with factors from other measures that tapped anxiety, cooperativeness, and stress reactivity; Anger, comprised of the Angry hostility and Impulsiveness facets of the NEO PI, along with negative loadings from factors from other instruments assessing even-temperedness, empathy, and calmness; and Depression, comprised of the Depression facet from the NEO PI along with depression-related factors from other instruments. They argued that the Depression facet operates as an additional measure of symptomatology, suggesting that neuroticism may best be understood as having two components, one associated with anxiety and vulnerability to emotional stress and the other representing irritability, anger, and hostility. The notion that neuroticism is divisible into two components is consistent with work identifying two genetic factors associated with neuroticism (Jang, Livesley, Angleitner, Riemann, & Vernon, 2002) and with the results of DeYoung and colleagues (2007) who argued that the structure of neuroticism includes two intermediate aspects existing at a middle level between the lower order facets and the higher order personality domains. The two aspects they identified, Withdrawal and Volatility, are similar to the Anxiety and Anger factors described above (although the Withdrawal aspect contained items representing depression whereas the Anxiety factor did not).

To understand the mechanisms through which neuroticism affects and is affected by treatments, there is a pressing need to identify precisely what it is that measures of neuroticism are capturing during an acute depressive episode, to disentangle the components of neuroticism from symptom-related distress, and to determine which specific features are influencing the observed treatment effects. Such work would not only provide valuable information about the nature of personality function when assessed during acute depressive episodes, but it would also allow researchers to examine causal relationships between the personality and symptom dimensions as each changes over time, to examine the independent influences of personality and symptom dimensions on broader domains of functioning, and to examine the independent biological and neurofunctional correlates of the personality and symptom dimensions so that relevant treatment targets might be identified.

Bifactor measurement models offer one way to identify unique and distinctive components of neuroticism that may be independent of depressive symptoms, even when measures of

personality and psychiatric symptoms are administered concurrently (Uliaszek et al., 2009). Bifactor models decompose the variance associated with a measure such that a general factor accounts for the common variance that underlies responses to all of the items, and additional specific factors (also called grouping factors) account for the remaining variance attributable to subsets of the items (Reise, 2012; Reise, Moore, & Haviland, 2010). As such, bifactor models make different assumptions about the structure of a domain compared to more standard factor models. For many psychological constructs, like neuroticism, there is a tension between whether the domain in question represents one, unidimensional entity or whether there are separable facets of the domain that can provide important incremental information over and above its most central feature (Reise et al., 2010). In a standard confirmatory factor model, the common variance across a measure is spread across all of the factors in the structure, and a higher-order factor can be included to represent the underlying general domain that gives rise to those correlations. By contrast, bifactor models assume that both the general factor and the specific factors have direct and independent effects on the observed data (Murray & Johnson, 2013). Higher-order and bifactor models share the feature that both contain a representation of a general domain, but in higher-order models, the effects of the higher-order general factor on the observed item scores are fully mediated by the lower-order factors. In bifactor models, these relationships are direct (Beaujean, Parkin, & Parker, 2014; Murray & Johnson, 2013; Reise et al., 2010). One consequence of this distinction is that for the higher-order models, the lower-order factors represent a mixture of common and specific variance (Beaujean et al., 2014; Reise et al., 2010). As such, features of these factors, such as their reliability and their correlations with external variables, reflect an admixture of the general domain and the specific variance due to the circumscribed feature in question (Reise et al., 2010). By contrast, bifactor models instantiate a complete decoupling of the specific components of a domain from whatever is common across the observed responses. As such, bifactor models provide researchers a set of conservative tests not available with standard higher-order factor models. First, they can be used to determine what, if anything, is left over beyond the common variance that runs across a measure of a domain. That is, they can be used to determine whether there are nonignorable subcomponents of the domain that cannot be captured by a unidimensional score representing what all of the items on the measure have in common. Second, they can identify which items are strongly associated with the common element of the domain and which are more strongly associated with a specific subcomponent. Finally, and most notably, they can be used to determine whether the specific subcomponents provide additional information about an external variable, e.g., an individual's functioning, over and above the common element representing the domain's most central feature (Benson, Kranzler, & Floyd, 2016; Chen, West, & Sousa, 2006).

The goals of the current study were (a) to determine whether meaningful subcomponents of neuroticism could be separated from concurrent symptoms of depression and anxiety and (b) to determine whether those subcomponents were associated with functioning, over and above acute symptoms. Given the associations between total neuroticism scores and psychiatric symptoms (Rosellini & Brown, 2011), we accomplished the first goal by fitting bifactor models to a popular measure of neuroticism, the NEO PI-R, to determine whether sufficient variance existed to form specific factors once the common variance across the

measure had been taken into account. We predicted that the general neuroticism factor from these models, representing the common variance, would capture the associations with symptoms of depression and anxiety, whereas the specific factors would be relatively unrelated to symptoms. To address the second goal, we examined whether the specific factors that we identified were incrementally associated with deficits in interpersonal functioning. We chose to examine interpersonal functioning as our validity criterion given the pivotal role that interpersonal functioning plays in theories about both the liability to depression and the course of the illness once it develops (Aaron T. Beck, Epstein, & Harrison, 1983; Blatt, Shahar, & Zuroff, 2001). Problems in interpersonal functioning are well-documented among depressed individuals (Barrett & Barber, 2007; Cain et al., 2012; Dawood, Thomas, Wright, & Hopwood, 2013; Hames, Hagan, & Joiner, 2013), and one of the empirically supported psychotherapies for depression, interpersonal psychotherapy, specifically targets interpersonal functioning problems to improve depressive symptoms (Klerman, Weissman, Rounsaville, & Chevron, 1984).

In the current study we tested three bifactor models: one based on the original six facet proposal of Costa and McCrae (1992) and two representations of the two-factor model described above (DeYoung et al., 2007; Naragon-Gainey & Watson, 2014). We hypothesized 1) that the portion of the structure representing the common variance across the items, the general factor, would correlate with standard symptom measures. By contrast, we hypothesized 2) that sufficient residual variance would exist to form coherent specific factors, 3) that those specific factors would be relatively independent of current psychiatric symptoms, and finally 4) that the specific factors would be incrementally associated with measures of interpersonal functioning deficits. Given recent concerns regarding replicability in social science research (e.g., Pashler & Wagenmakers, 2012; Tackett et al., 2017), our secondary aim was to replicate the observed findings across two separate, independent samples – one comprised of depressed adults and the other a community sample of individuals displaying the full range of symptom presentations from psychological health to acute mental illness.

Method

Participants

Depression Sample.—Data for the depression sample were collected at two sites, the University of Toronto/Centre for Addiction and Mental Health and the University of Pittsburgh. All participants (N=807) were diagnosed using the Structured Clinical Interview for DSM-IV Axis I Disorders with a current depression spectrum disorder (major depressive disorder, dysthymic disorder, or depressive disorder not-otherwise-specified). The Toronto sample (N=748) was part of a larger research database comprised of data aggregated across multiple, smaller studies (see, Tackett, Quilty, Sellbom, Rector, & Bagby, 2008 for additional details). The majority (n=551) entered the database as part of treatment trials; a smaller number (n=147) entered from other clinical research studies; and a small number (n=50) entered following clinical assessments. Participants from the Pittsburgh site (n=59) were a sub-sample from a larger study of emotional and interpersonal functioning among individuals across the spectrum of severity of borderline personality disorder (Scott et al.,

2013). Formal borderline personality diagnosis was not required for entry into the larger study, and only 17% of the larger sample received that diagnosis (Wright, Scott, Stepp, Hallquist, & Pilkonis, 2015). Thus, the combined sample represents a diverse collection of participants with depression from a variety of clinical research settings.

The majority of participants were female (62%), and the average age was 40 (*SD*=11). Most carried diagnoses of major depressive disorder (97%), whereas smaller numbers were diagnosed with dysthymic disorder (2.6%), and depressive disorder not-otherwise-specified (0.4%). The majority of participants were unmarried (63%), and most had received some form of post-secondary education (76%). See Supplemental Table 1 for additional information.

Community Sample.—The second sample was derived from the University of Pittsburgh Adult Health and Behavior (AHAB) project (Manuck, Phillips, Gianaros, Flory, & Muldoon, 2010). The AHAB project represents a registry of 1,295 midlife (30-54 years of age) community volunteers who provided interview, questionnaire, behavioral, and biological data. AHAB participants were recruited through mass mailings in Southwestern Pennsylvania. Exclusion criteria for the larger AHAB project have been described in detail elsewhere (Erickson et al., 2013; Manuck et al., 2010). The only psychiatric exclusion for the AHAB project was the presence of a psychotic-spectrum disorder. For the current study, we included all individuals from the AHAB project who provided data on the NEO PI-R (*N*=1,284). The gender ratio in the AHAB sample was roughly balanced (53% female), and the average age was 45 (*SD*=7). The majority of participants were married (65%), and most had received some form of post-secondary education (87%). The majority of participants (80%) did not meet criteria for any Axis-I diagnosis, and only a small percentage (1%) were diagnosed with a current depression spectrum disorder. Thus, the sample represented a heterogeneous mixture of psychiatrically healthy individuals as well as individuals with a variety of psychiatric disorders. See Supplemental Table 1 for additional information.

Procedures

Participation was voluntary, and all participants provided written informed consent. Research methods were approved by the relevant Institutional Review Board at each site and were carried out in accordance with the World Medical Association Declaration of Helsinki.

Clinical and Personality Measures.—In the depression sample, depression symptom severity was assessed using the 17-item version of the Hamilton Rating Scale for Depression (HRSD, Hamilton, 1960). Across the separate smaller studies aggregated at the Toronto site, evaluations were conducted by individuals with a variety of roles on the respective studies, including clinical evaluators, trained research coordinators, and in some cases, research clinicians. At the Pittsburgh site, evaluations were conducted by blinded clinical interviewers. In the Pittsburgh depression sample, anxiety was assessed using the Hamilton Rating Scale for Anxiety (HRSA, Hamilton, 1959); at the Toronto site, anxiety was assessed using the total count of threshold-level symptoms from the anxiety disorders subsection (Section F) of the Structured Clinical Interview for DSM-IV Disorders (SCID, First, Spitzer, Gibbon, Willians, & Benjamin, 1995). In the community sample, depression symptoms were

assessed using the Beck Depression Inventory (BDI, A.T. Beck, Steer, & Brown, 1996). Concurrent state anxiety was not assessed in the community sample. All participants in both samples completed the Revised NEO Personality Inventory (NEO-PI-R, Costa & McCrae, 1992), a 240-item self-report inventory that measures the constructs associated with the Five-Factor Model of personality. Analyses focused on the 48 items in the neuroticism scale. Negatively keyed items were reverse coded.

Statistical Approach.—Ratings on NEO PI-R items were treated as ordered categories, and polychoric correlations were used to estimate models (Holgado-Tello, Chacón-Moscoso, Barbero-García, & Vila-Abad, 2010). All measurement models were confirmatory, such that all cross-loadings were set to zero, and all models were estimated using Mplus (Version 7.2) with the robust weighted least squares with mean and variance adjustment (WLSMV) estimator (Muthen & Muthen, 2012). Five models of the structure of neuroticism in depression were examined - two standard confirmatory factor analytic (CFA) models and three confirmatory bifactor models. The standard CFA models were examined to provide a basis for comparison and are reported for descriptive purposes. The first (Model 1) examined a simple one-factor solution in which all 48 items of the neuroticism scale loaded on a single latent variable. The second (Model 2) examined the original six-facet structure of neuroticism proposed by Costa and McCrae (1992), with the facets freed to correlate. To examine the study's primary hypotheses, we estimated three fully orthogonal bifactor models. The first of these (Model 3) tested a version of the two-factor conceptualization of neuroticism described above. It incorporated a general factor that we labeled Negative Affectivity to capture the common variance across all items, and two specific factors, one representing Emotional Vulnerability (comprised of items from the Anxiety, Depression, Self-consciousness, and Vulnerability facets) and the other representing Volatility (containing items from the Angry-hostility and Impulsiveness facets). The second bifactor model (Model 4) was similar to Model 3, but, in accordance with the findings of Naragon-Gainey and Watson (2014), items from the original Depression subscale were separated to form a third specific factor. The final bifactor model (Model 5) contained a general Negative Affectivity factor and six specific factors, one each for the six facets of neuroticism. For the sake of comparisons with prior literature, we retained the names of the original facets for the six specific factors in this model.

Model fit was assessed using the comparative fit index (CFI, Bentler, 1990), the Tucker-Lewis index (TLI, Tucker & Lewis, 1973) and the root mean square error of approximation (RMSEA, Steiger & Lind, 1980). Values grater than 0.90 for the CFI and TLI (Bentler, 1990; Hu & Bentler, 1999) and RMSEA values less than 0.08 are commonly taken to reflect adequate fit (Marsh, Hau, & Wen, 2004). To ensure an interpretable model, we also examined the magnitude of factor loadings. Once the best-fitting bifactor model was selected, we assessed the proportion of the total common variance that could be explained by the general factor and the proportion remaining that could be explained by the specific factors (Reise, Scheines, Widaman, & Haviland, 2012). Factor score determinacy, representing the correlation between a factor score estimate and the latent construct it assesses (Gorsuch, 1983), was also calculated (Rodriguez, Reise, & Haviland, 2015. See the Supplement for a discussion of additional bifactor indices.). In order to evaluate the

generalizability of the best-fitting bifactor model, the same statistical approach described above was used in both the depression and the community samples. Finally, factor congruencies (Lorenzo-Seva & ten Berge, 2006) were calculated to evaluate the similarity between the best-fitting bifactor solutions from the two samples.

Criterion Validity.—Primary validity analyses of the best-fitting bifactor model were conducted using data from the community sample, which completed the 32-item Inventory of Interpersonal Problems (IIP) short form (Soldz, Budman, Demby, & Merry, 1995). To provide preliminary evidence about whether the patterns observed in the community sample replicated in a clinical sample, additional analyses, reported in the Supplement, were conducted with a small subset of data from the Pittsburgh site of the depression sample (n=59) who completed the 127-item version of the IIP (Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988). The IIP tools are well-validated, self-report measures of distress resulting from interpersonal causes. Validity analyses focused on the five scales from the IIP identified by Pilkonis and colleagues (1996) as having particular relevance to patients with personality pathology: *Interpersonal Sensitivity*, reflecting emotional reactivity to criticism; *Interpersonal Ambivalence*, representing difficulty in joining with others to advance a goal; *Aggression*, tapping hostility towards others; *Need for Social Approval*, reflecting anxiety about others' evaluations; and *Lack of Sociability*, representing reticence and distress when interacting with others. Separate structural equation models were used to estimate associations with each of the IIP factors in the community sample. Each model simultaneously examined the contribution of the general factor and all of the specific factors, while controlling for concurrently assessed depression symptoms. We chose this modelling strategy in order to provide a strict test of the incremental validity of each of the specific factors. To ensure that that any observed relationships were not simply the result of residual correlations with the remaining domains of the Five Factor Model of personality, we added the remaining four domains of the Five Factor Model (Extraversion, Conscientiousness, Openness, and Agreeableness) as additional covariates in secondary analyses of data from the community sample. Finally, tertiary analyses examined the interpersonal circumplex scales from the IIP in the community sample (see the Supplement).

Results

Depression Sample

As expected, neuroticism levels were high in the depression sample, M=119.55, SD=23.73 (approximately 2 SD above the population average), but the variability and range (50-181) were substantial, suggesting large individual differences. Average depression symptoms (HRSD) were in the moderate-to-severe range, M=18.67 SD=6.54, average HRSA anxiety symptoms at the Pittsburgh site were in the mild-to-moderate range, M=21.93 SD=8.30, and the mean number of anxiety disorder criteria endorsed at the Toronto site was M=6.28 SD=7.02 (this variable was non-normally distributed with *Median*=4 and *Mode*=0).

Model fit.—The fit statistics for each model are displayed in Table 1. For comparative purposes, we first display the fit statistics for the standard CFA models and note that these failed to meet standard thresholds for adequate fit. Next, we display the fit indices of the

three orthogonal bifactor models. The fit of the first two bifactor models (Models 3 & 4), which included a general factor and two specific factors representing Emotional Vulnerability and Volatility (Model 4 additionally included a separate Depression specific factor), likewise failed to meet standard thresholds for adequate fit. Moreover, the loadings of the relevant items on the Emotional Vulnerability specific factors for each of these bifactor models were generally quite small, with few items loading > |0.30| (see Supplemental Tables 2-5 for the factor structures of Models 1-4). By contrast, the final bifactor model, which included one general factor and six specific factors for the original six neuroticism facets (Model 5, Table 2), had good fit to the data.

Salient factor loadings.—The factor loadings of the final model are displayed in Table 2. All but three items from the original Anxiety, Angry hostility, Depression, Self-consciousness, and Vulnerability facets had strong loadings on the general Negative Affectivity factor. By contrast, only three items from the original Impulsiveness facet loaded on the general factor. For five of the six specific factors, at least four items from the original facet had salient loadings on the relevant specific factor. The exception was the Vulnerability specific factor, on which only two items from the original facet loaded. The remaining items had weak loadings on the Vulnerability specific factor, with most <0.20.

A primary question in the current study concerned whether the specific factors from the best-fitting bifactor model were independent of current symptom severity. Table 3 displays the correlations between the original unit-weighted neuroticism scores (total and facet-level scores) and symptoms of depression and anxiety. Significant, positive correlations were observed between depression symptoms and standard, unit-weighted neuroticism total and facet scores. The same pattern was observed for correlations with anxiety symptoms, with the exception of the Impulsiveness facet, which was not significantly correlated with anxiety. Table 3 also displays the corresponding correlations between the elements of the bifactor model and symptom scores. For both symptom types, the correlation between total neuroticism score and symptoms was largely captured in the bifactor model structure by the general Negative Affectivity factor, as hypothesized. That is, the correlation between the general Negative Affectivity factor and each symptom type was similar to the corresponding correlation between the standard, united-weighted neuroticism total score and each symptom type. By contrast, the correlations between the specific factors and symptoms were substantially reduced. No significant positive correlations were observed between the specific factors and depression symptoms, and the only significant associations observed with depression symptoms were small and negative. Together, these patterns suggest that when the common variance across the items of the NEO-PI-R is represented in a general factor, the unique elements that remain capture elements of individual differences that are distinct from acute symptoms of depression. This pattern is particularly notable for the Depression specific factor. The lack of residual associations with depressive symptoms suggests that the majority of variance in depressivity is captured by the general factor.

Regarding associations with anxiety symptoms, we observed a small, significant positive association between the Anxiety specific factor and SCID-rated anxiety symptom counts in the larger of the two subsamples of the depression sample (Table 3). This suggests that after parceling out the common variance, the Anxiety specific factor captures some information

about concurrent anxiety disorder symptoms, whereas the remaining specific factors are relatively independent from these symptoms. However, we did not observe significant associations with anxiety symptoms in the smaller subsample from the Pittsburgh site (see the Supplement). Ultimately, all specific factor associations with both depression and anxiety symptoms were small (i.e., < |0.20|).

Variance explained and factor determinacy.—A high percentage of variance (ωH=0.87) in standard, unit-weighted neuroticism total scores can be attributed to individual differences on the general Negative Affectivity factor; however, the general factor accounted for only 60% of the total reliable variance among the items. This indicates that a substantial portion of reliable variance remained for the formation of the residual specific factors. Factor score determinacies were: general Negative Affectivity factor (0.96), Anxiety specific factor (0.76), Angry Hostility specific factor (0.89), Depression specific factor (0.74), Self-Consciousness specific factor (0.77), Impulsiveness specific factor (0.88), and Vulnerability specific factor (0.91).

Community Sample

In the community sample, average neuroticism levels were closer to population norms (M=74.32, SD=23.22), and a wide range of scores was observed (*range*=10-151). Average BDI depression scores (*M*=4.03, *SD*=4.76) were low, as expected given the low percentage of participants in the sample who met criteria for a diagnosable depressive illness. See Supplemental Table 1 for individual facet scores and additional details.

Model Fit.—As reported in Table 1, the pattern of model fit statistics in the community sample was quite similar to that in the depression sample. As was the case for the depression sample, the fit indices for the first two bifactor models did not meet standard cut-off criteria, and the majority of the loadings on the specific factors in those models were < |0.30|. The best fitting bifactor model, and the only model in which all fit statistics were in acceptable ranges, was Model 5 – the bifactor structure with one general negative affectivity factor and six specific factors for each of the six original neuroticism facets. Table 2 displays the loadings of the best fitting model (Model 5) in the community sample. The pattern of loadings was similar between the depression and community samples for the majority of the specific factors. As was the case in the depression sample, the impulsiveness specific factor appeared to be the most distinct from the general factor, with only four of the eight items contributing salient loadings to the general factor. The largest discrepancy between the two samples occurred with the depression specific factor, on which only two items had loadings > 0.30 and the majority had loadings < 0.20 in the community sample. The similarity of the bifactor solutions in the two samples was reflected in strong factor congruency estimates, which were ≥ 0.95 for all the specific factors with the exception of the depression specific factor (congruency = 0.75). Congruence values in this range (≥ 0.95) suggest that the two factors being compared are essentially equivalent (Lorenzo-Seva & ten Berge, 2006).[1]

[1]The factor solutions for the best fitting bifactor models in both samples contained a small number of items with negative loadings. We re-estimated the models in both samples, setting these negative loading items to zero, and we observed that the resulting solutions were nearly identical to those presented in Table 2. Congruencies between the original solutions and the modified solutions were 1.00 for all elements of the structure across both samples, with the exception of the depression specific factor in the community sample

In the community sample, the general Negative Affectivity factor accounted for 66% of the total common reliable variance among the items. As was the case in the depression sample, this suggests that sufficient reliable variance remained for the formation of the residual specific factors. Factor determinacies were more varied in the community sample and ranged from 0.74-0.97 (Table 2).

Separation from Symptomatic Distress.—As in the depression sample, the bifactor modeling approach in the community sample yielded a solution in which the specific factors were relatively independent of concurrent depressive symptomatic distress. As displayed in Table 3, the correlation between neuroticism total scores and depression symptoms in the community sample was high (r=0.50), as were the correlations between the original neuroticism facet scales and depression symptoms (rs=0.30-0.51). As in the depression sample, these relationships were largely captured in the bifactor model by the general factor (r=0.52), whereas the correlations between the specific factors and symptoms were small and substantially lower than the correlations observed for the original facet scores. Given the larger sample size compared to the original depression sample, we were able to detect a significant, small relationship (r=0.11) between the depression specific factor and symptoms of depression, and small negative relationships (rs≤−0.11) between the anxiety, self-consciousness, and vulnerability specific factors and depression symptoms. As with the clinical sample, all specific factor associations with symptoms were small (i.e., < |0.20|).

Criterion Validity

Table 4 displays the associations in the community sample between each of the elements of the best fitting bifactor model and the scales of the IIP, controlling for symptoms of depression and all other parts of the structure. Given the large size of the sample, we had substantial power to detect even subtle effects. Focusing on larger effects, |β|s>0.30, we observed that the general Negative Affectivity factor was associated with the Interpersonal Sensitivity, Need for Approval, and Lack of Sociability IIP sub-scales. In addition, two specific factors (Angry Hostility and Self-consciousness) demonstrated significant, incremental associations with IIP subscale scores, controlling for symptoms of depression, general negative affectivity, and scores on each of the other specific factors. The Self-Consciousness specific factor explained additional variance in the Interpersonal Sensitivity, Need for Approval, and Lack of Sociability IIP factors, whereas the Angry Hostility specific factor explained additional variance in Aggression (Table 4). Recognizing that these patterns might simply represent residual associations between the specific factors and the remaining four domains of the Five Factor Model of Personality (Openness, Conscientiousness, Extraversion, and Agreeableness), we re-estimated these models controlling also for total scores from these domains. The associations between the Self-Consciousness and Angry-Hostility specific factors and interpersonal functioning remained strong and significant (Self-Consciousness and Interpersonal Sensitivity, B=0.29, p<0.001, Need for Approval, B=0.25, p<0.001, Lack of Sociability, B=0.27, p<0.001; Angry-Hostility and Aggression, B=0.37, p<0.001, see Supplement for complete model results). Results of the validity

(congruency=0.97), which contained three items that were modified. Given the similarity between the original and modified solutions, we elected to retain the original solution.

analysis in a small subset (n=59) of the depression sample largely replicated these patterns of associations (see the Supplement for the results in this sample as well as for relationships between elements of the bifactor structure and the octants of the IIP circumplex scales within the community sample).

Discussion

We used bifactor measurement models to examine the structure of neuroticism to identify unique elements of the structure that are independent of depression and anxiety symptoms. Using data from the NEO PI-R, we observed that a bifactor model with one general factor and six specific factors, representing the six facets proposed by the developers of the NEO PI-R, fit well and better than alternative bifactor models both in a sample of depressed individuals and in a separate sample of community dwelling adults. Not only were the congruencies between the solutions in the two samples strong, but we also observed that in both samples, the general factor captured much of the association between neuroticism and measures of acute symptoms, whereas the six residual specific factors were relatively independent of symptoms. Finally, and most importantly, we observed replicated evidence that certain of the specific factors accounted for additional variance in particular domains of interpersonal functioning over-and-above general negative affectivity and acute psychiatric symptoms.

The best fitting model suggested that although a strong general factor was observed, sufficient reliable residual variance remained to form meaningful specific factors. This result has important implications for the structure of neuroticism, as it suggests that there are nonignorable subcomponents of the domain that cannot be captured adequately by the common element of negative affectivity. The general Negative Affectivity factor was composed of loadings from the majority of items from five of the six original neuroticism facets. The notable exception was the Impulsiveness facet where the loadings on the general factor from items representing the original Impulsiveness facet were small, suggesting that the general factor is capturing something distinct from impulsivity, at least as measured by the NEO PI-R. The general Negative Affectivity factor was associated with symptoms of depression and anxiety, and the strength of these relationships were similar to that observed for standard unit-weighted neuroticism total scores. This suggests that the general Negative Affectivity factor captured much of the association between neuroticism and acute measures of depression and anxiety, as hypothesized. In addition, we observed that the general factor was broadly associated with interpersonal functioning deficits. Taken together, we interpret the general Negative Affectivity factor to represent a global marker of negative emotionality that is composed of a mixture of state and trait influences. We believe it to be a broad indicator of emotional distress, and as such, we do not expect it to be particularly helpful in discriminating among different possible mechanisms of illness (or treatment response) operating within different individuals. What is lost with respect to the granularity of the general factor, however, is preserved in the specific factors which represent constellations of reliable variance distinct both from current symptoms and from the broad general factor.

In the best fitting bifactor solution, we observed that two of the specific factors (Angry hostility and Self-consciousness) were particularly helpful in identifying important and

distinct individual differences in interpersonal functioning. Across both samples, six of the eight original Angry Hostility items had salient loadings on the Angry Hostility Specific factor. This specific factor was independent of depression or anxiety symptoms, and it uniquely predicted interpersonal problems associated with aggression in both samples, over and above psychiatric symptoms, general negative affectivity, each of the other specific factors, and (in the community sample) the remaining four domains of the Five Factor Model. The Self-Consciousness specific factor also comprised a consistent pattern of salient loadings in both the depression and community samples. Items that loaded on the specific factor represented the tendency to feel frank self-consciousness, embarrassment, and concern over making mistakes in social settings. We observed a small, negative relationship between it and depression symptoms in both samples, and we observed that the Self-Consciousness specific factor was associated with interpersonal problems associated with heightened sensitivity to the criticism of others, lack of sociability, and need for others' approval. Again, these associations were present even when controlling for general negative affectivity, depressive symptoms, each of the other specific factors, and (in the community sample) the remaining four domains of the Five Factor Model. The differentiation observed between the interpersonal problems associated with the Angry Hostility and Self-Consciousness specific factors is consistent with the findings of Pilkonis and colleagues (Pilkonis et al., 1996; Stern, Kim, Trull, Scarpa, & Pilkonis, 2000), who observed that certain of the IIP subscales, including aggression, were more strongly associated with Cluster B personality pathology, whereas others, including lack of sociability and need for approval, were strongly associated with Cluster C pathology. It may be that the Angry Hostility and Self-Consciousness specific factors are tapping into elements of the interpersonal problems associated with these forms of personality pathology.

The fact that these two components of neuroticism, Angry Hostility and Self-Consciousness, are not substantially correlated with symptoms and that each incrementally predicts deficits in interpersonal functioning over and above symptom scores suggests that these two dimensions are capturing meaningful individual differences that cannot be attributed to the individuals' current symptomatic state. The fact that they can be separated from the general Negative Affectivity factor and predict incremental variability in interpersonal functioning over and above the general factor and the remaining domains of the Five Factor Model suggests that they represent important parts of the structure of personality that cannot be adequately captured by the other elements of the model. Rather, they represent unique and nonignorable facets of neuroticism. Angry Hostility captures the proclivity to lose one's temper and react in an aggressive and hostile manner. Self-Consciousness marks heightened sensitivity to others and the tendency to react with an increased self-focus coupled with feelings of embarrassment. These dimensions, and their relationships with symptom and functioning measures, were quite similar across both the community and depression samples. Thus, these two dimensions represent unique components of neuroticism with distinct interpersonal functioning profiles that can be assessed during an acute episode and that vary independently of each other and of acute symptoms.

On the one hand, it is not surprising that there are elements of neuroticism that are separable from acute distress. As reviewed above, measures of neuroticism predict differential treatment response and change during treatments even when acute symptomatic distress is

well controlled, suggesting that the variance in neuroticism that is nonoverlapping with acute symptoms is important and meaningful. At the same time, this is the first study of which we are aware to characterize the nature of those unique components of neuroticism, to replicate their structure across diverse samples, and to identify their incremental relationships with functioning. With the identification of these dimensions, future work will be able to examine whether change in these components of neuroticism mediates symptom change during treatments for depression. In addition, future research will be able to examine whether the finer granularity of these dimensions will aid in identifying better and more robust neurobiological markers than has been possible to date either for symptom levels or frank diagnosis of depression (see Murray & Johnson, 2013 for a similar suggestion with respect to the structure of cognitive abilities).

Two additional specific factors (Anxiety and Impulsiveness) may show promise in describing meaningful and distinct dimensions on which individuals differ, although additional work will be needed to confirm these effects. The majority of items (6 of 8) from the original Anxiety facet loaded on the Anxiety specific factor in both samples. Although we did not observe an association between the Anxiety specific factor and Hamilton anxiety scale scores in the smaller of the two depression subsamples (reported in the Supplement), we did observe a significant association between the Anxiety specific factor and the number of anxiety disorder criteria endorsed during a diagnostic interview in the larger of the two depression subsamples. We interpret this finding as preliminary evidence suggesting that the Anxiety specific factor may be capturing features of anxiety disorder that are distinct from general negative affectivity.

The Impulsiveness specific factor demonstrated the largest congruency between the two samples, but relatively few items from the original Impulsiveness facet loaded on the general Negative Affectivity factor. This pattern informs a long-running debate about the proper placement of impulsivity within structural models of personality. Although many theorists identify trait impulsiveness as an important part of personality, there is little agreement about whether dimensions of impulsivity are best captured by domains representing negative affectivity, extroversion, constraint, or conscientiousness (see Whiteside & Lynam, 2001 for a review). Part of the challenge no doubt stems from the multidimensional nature of impulsivity, such that different components may well fit best with different domains of personality (Dvorak, Pearson, & Kuvaas, 2013; Miller, Zeichner, & Wilson, 2012; Whiteside & Lynam, 2001). A further complication stems from the general tendency for questionnaire items that tap negative elements of affect and cognition to be captured by factors that represent neuroticism in structural models (Clark & Watson, 1995). In the current study, we observed that most of the NEO-PI-R impulsivity items were not strongly associated with the same construct of general negative affectivity as were the remainder of the neuroticism items, suggesting that the component of impulsivity measured by the NEO PI-R is relatively distinct from the rest of the measure. Given the scope of the present project, we did not examine whether the impulsivity items would fit better on a different broad domain of personality. We did not observe substantial associations between the Impulsiveness specific factor and the interpersonal functioning measure under consideration. Others have observed interactions between neuroticism and components of impulsivity in the prediction of interpersonal problems (Dvorak et al., 2013; Miller et al., 2012). As such, more work, with

additional measures that more fully capture the multidimensionality of impulsivity, may be needed to fully determine whether elements of impulsivity should be considered as parts of neuroticism, as well as how they relate to functioning.

The remaining two specific factors (Depression and Vulnerability) do not appear to be particularly useful in describing meaningful individual differences over and above the general Negative Affectivity factor. The Depression specific factor was associated with a disparate pattern of loadings across the two samples and was not associated with a strong or replicable pattern of interpersonal functioning deficits. Moreover, it was not substantially associated with depressive symptoms once the general Negative Affectivity factor was covaried. We interpret these results to indicate that there simply may not be sufficient reliable variance in the NEO-PI-R assessment of depressionality that is distinct and separable from the common dimension of negative emotionality that runs across the measure. A similar finding was recently obtained by Arrindell and colleagues (Arrindell et al., 2017) who examined the structure of symptoms across the clinical psychiatric disorders and likewise observed that depressive symptoms were largely captured by a general psychopathology factor, with little meaningful specific variance left over to form a robust depression specific factor. Similarly, for the Vulnerability specific factor, only two items from the original Vulnerability facet loaded on the residual Vulnerability specific factor, and we observed no significant associations between this specific factor and difficulties in interpersonal functioning. As such, individual differences in vulnerability likewise appear to be well accounted for by the broad dimension of negative affectivity that captures the common variance in the measure. That is, neither the Depression nor the Vulnerability specific factors may be useful on their own. Rather than capturing meaningful individual differences, they appear to be capturing relatively uninformative residual correlations between a small number of items. It is not uncommon in bifactor models to identify specific factors like these that do not on their own have substantial explanatory power (Reise et al., 2010; Reise, Morizot, & Hays, 2007; Rodriguez et al., 2015). We decided to retain these specific factors in the final model in order to help resolve the other elements of the structure, and we would caution against over interpretation of their content.

Comparisons with Prior Literature and Limitations

This is the largest study to date to examine the components of neuroticism in a sample of depressed adults. However, one limitation of the study is the heterogeneity of the samples with respect to the nature of the study of origin, diagnosis, treatment status, and clinical course. Data regarding these possible differences were not uniformly available, and as such, their impact could not be assessed. Although future work will be needed to examine the impact of clinical features, we view the high degree of heterogeneity within and between the two samples as a substantial strength, supporting the potential generalizability of the present findings. Related to this, given the small sample size used for the validity analyses in the depression sample, additional replication in larger clinical samples is warranted.

Contrary to prior work, we did not observe clear support of a two-factor model, either in the depression sample or in the community sample. Several differences between the current study and previous reports may account for the difference in findings: First, we used bifactor

models, whereas prior reports used exploratory factor analytic techniques to examine the structure of neuroticism. Controversies in the literature remain about how best to compare traditional factor solutions and bifactor solutions (see, e.g., Chen et al., 2006; Gignac, 2016; Murray & Johnson, 2013; Reise, 2012). Indeed, some have observed bias in model fit statistics that favor bifactor models over standard confirmatory or higher order models (Bonifay, Lane, & Reise, 2016; Gignac, 2016; Molenaar, 2016; Morgan, Hodge, Wells, & Watkins, 2015; Murray & Johnson, 2013). At the same time, recent work suggests that the superior fit of bifactor models in these studies may stem not from bias per se but from the failure to meet a statistical assumption (the proportionality constraint) inherent in standard higher-order factor models that is not relevant to bifactor models (Gignac, 2016). Although this debate is ongoing, those on both sides remind researchers that no modelling strategy or test statistic can definitively uncover the truth of how a construct is organized in nature (Gignac, 2016; Morgan et al., 2015; Murray & Johnson, 2013; Rodriguez et al., 2015). Rather, the most appropriate modeling strategy depends on the aims of the project (Murray & Johnson, 2013), and the ultimate utility of a model will depend on whether it can provide reliable and valid information that is not well captured by other approaches. In line with the recommendations of others (Benson et al., 2016; Chen et al., 2006; Murray & Johnson, 2013), in this study we used the bifactor measurement model primarily to determine whether separable subcomponents of neuroticism exist that predict incremental variability in functioning over and above general negative affectivity and symptom scores. Although we did observe that meaningful specific factors could be formed, we did not observe clear support for the two-factor models identified by others using standard confirmatory approaches. Rather, we observed that the general factor accounted for much of the variance in the items that might otherwise have contributed to one of those two factors.

A second difference with previous work is that we examined data from one self-report personality measure, an instrument that was designed a-priori to capture six components of neuroticism. This design feature may well have contributed to our finding that the six-specific-factor model fit the data better than the alternative bifactor models. Prior reports identifying two-factor solutions have combined data from multiple measures. It is possible that two of the specific factors that we identified in the current work, Angry hostility and Self-Consciousness, could represent seeds for two strong subfactors of neuroticism that are well aligned with the constructs captured by those two factor solutions. Future work should incorporate multiple measures and examine whether elements that were part of prior two-factor solutions and that were not well represented in the current study (e.g., measures of calmness, stability, moderation, toughness, stress reactivity, and even-temperedness) might contribute to two such factors. Incorporating additional measures may also improve the observed factor determinacies, which were low for some of the specific factors. Determinacies that are lower than optimal can signal concerns about the validity of factor estimates. These concerns are mitigated for some of the specific factors with low determinacies in this study, e.g., Self-Consciousness, by the observed replication of patterns of association with external interpersonal functioning variables across the two samples.

Additionally, self-report instruments like the one used in this study, require individuals to have intact insight into their cognitive, affective, behavioral, and motivational tendencies. Future work could combine data from multiple sources, including self-report, informant, and

interview-based measures in order to help ensure that more of the relevant information is captured. Examining these components in longitudinal and lab-based observation and challenge studies would also provide valuable information not just about the long-term stability of these markers, but also about the conditions in which they are or are not associated with changes in affect and behavior. Finally, we validated the components of the structure with respect to one domain of functioning. Future work will be needed to examine associations with other important areas of functioning (e.g., education, employment, quality of life).

Conclusions

The findings from the current study suggest that neuroticism measured either in the context of depression or among community dwelling adults can be decomposed into a combination of general negative affectivity, which is associated with symptom severity, and nonignorable subcomponents, which are relatively uncorrelated with depression and anxiety. At least two of these subcomponents, Angry-hostility and Self-consciousness, are distinguishable from acute symptoms, are independently associated with particular deficits in interpersonal functioning, and are not adequately captured by a common element of negative affectivity that explains a large portion of the variance in neuroticism. Using elements of the structure identified in this study, future work could examine which components of neuroticism are most predictive of treatment success or failure and which are themselves affected by treatments. Findings from such work would advance our ability to target treatments more efficiently to specific individuals.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The data from the University of Toronto/CAMH site of the depression sample were collected with the support of an operating grant and a Senior Investigator Award from the Ontario Mental Health Foundation conferred to RMB. Data from the Pittsburgh site of the depression sample were collected with the support of the National Institutes of Health (PAP, MH056888). Data from the community sample were collected with the support of the National Institutes of Health (PO1 HL040962, RO1 HL065137). JCF is supported by MH097889 and AGCW by L30 MH101760 from the National Institutes of Health.

References

Arrindell WA, Urbán R, Carrozzino D, BECH P, Demetrovics Z, & Roozen HG (2017). SCL-90-R emotional distress ratings in substance use and impulse control disorders_ One-factor, oblique first-order, higher-order, and bi-factor models compared. Psychiatry Research, 255, 173–185. doi: 10.1016/j.psychres.2017.05.019 [PubMed: 28558358]

Bagby RM, Psych C, Quilty LC, & Ryder AC (2008). Personality and depression. Canadian Journal of Psychiatry, 53(1), 14–25. doi:10.1177/070674370805300104 [PubMed: 18286868]

Bagby RM, Quilty LC, Segal ZV, McBride CC, Kennedy SH, & Costa PT Jr. (2008). Personality and differential treatment response in major depression: A randomized controlled trial comparing cognitive-behavioural therapy and pharmacotherapy. Canadian Journal of Psychiatry, 53(6), 361–370. [PubMed: 18616856]

Barlow DH, Ellard KK, Sauer-Zavala S, Bullis JR, & Carl JR (2014). The origins of neuroticism. Perspectives on Psychological Science, 9(5), 481–496. doi:10.1177/1745691614544528 [PubMed: 26186755]

Barrett MS, & Barber JP (2007). Interpersonal profiles in major depressive disorder. Journal of Clinical Psychology, 63(3), 247–266. doi:10.1002/jclp.20346 [PubMed: 17211875]

Beaujean AA, Parkin J, & Parker S (2014). Comparing Cattell–Horn–Carroll factor models: Differences between bifactor and higher order factor models in predicting language achievement. Psychological Assessment, 26(3), 789–805. doi:10.1037/a0036745 [PubMed: 24840178]

Beck AT, Epstein N, & Harrison R (1983). Cognitions, attitudes and personality dimensions in depression. British Journal of Cognitive Psychotherapy, 1(1), 1–16.

Beck AT, Steer RA, & Brown GK (1996). Manual for the BDI-II. San Antonio, TX: Psychological Corporation.

Benson NF, Kranzler JH, & Floyd RG (2016). Examining the integrity of measurement of cognitive abilities in the prediction of achievement: Comparisons and contrasts across variables from higher-order and bifactor models. Journal of School Psychology, 58(C), 1–19. doi:10.1016/j.jsp.2016.06.001 [PubMed: 27586067]

Bentler PM (1990). Comparative fit indexes in structural models. Psychological Bulletin, 107(2), 238–246. doi:10.1037/0033-2909.107.2.238 [PubMed: 2320703]

Blatt SJ, Shahar G, & Zuroff DC (2001). Anaclitic (sociotropic) and introjective (autonomous) dimensions. Psychotherapy: Theory, Research, Practice, Training, 38, 449–454.

Bonifay W, Lane SP, & Reise SP (2016). Three concerns with applying a bifactor model as a structure of psychopathology. Clinical Psychological Science, 5(1), 184–186. doi:10.1177/2167702616657069

Cain NM, Ansell EB, Wright AGC, Hopwood CJ, Thomas KM, Pinto A, . . . Grilo CM (2012). Interpersonal pathoplasticity in the course of major depression. Journal of Consulting and Clinical Psychology, 80(1), 78–86. doi:10.1037/a0026433 [PubMed: 22103955]

Chen FF, West S, & Sousa K (2006). A comparison of bifactor and second-order models of quality of life. Multivariate Behavioral Research, 41(2), 189–225. doi:10.1207/s15327906mbr4102_5 [PubMed: 26782910]

Clark LA, & Watson D (1995). Constructing validity: Basic issues in objective scale development. Psychological Assessment, 7(3), 309.

Costa PT, & McCrae RR (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Professional Manual. Odessa, FL: Psychological Assessment Resources.

Dawood S, Thomas KM, Wright AGC, & Hopwood CJ (2013). Heterogeneity of interpersonal problems among depressed young adults: Associations with substance abuse and pathological personality traits. Journal of Personality Assessment, 95(5), 513–522. doi:10.1080/00223891.2013.781031 [PubMed: 23560433]

DeYoung CG, Quilty LC, & Peterson JB (2007). Between facets and domains: 10 aspects of the Big Five. Journal of Personality and Social Psychology, 93(5), 880–896. doi:10.1037/0022-3514.93.5.880 [PubMed: 17983306]

Dvorak RD, Pearson MR, & Kuvaas NJ (2013). The Five-Factor Model of impulsivity-like traits and emotional lability in aggressive behavior. Aggressive Behavior, 39(3), 222–228. doi:10.1002/ab.21474 [PubMed: 23471690]

Endler NS, Rutherford A, & Denisoff E (1997). Neuroticism: How does one slice the PI(e)? European Journal of Personality, 11(2), 133–145. doi:10.1002/(sici)1099-0984(199706)11:2<133::aid-per260>3.0.co;2-f

Erickson KI, Banducci SE, Weinstein AM, MacDonald III AW, Ferrell RE, Halder I, . . . Manuck SB (2013). The brain-derived neurotrophic factor Val66Met polymorphism moderates an effect of physical activity on working memory performance. Psychological Science, 24(9), 1770–1779. doi:10.1177/0956797613480367 [PubMed: 23907543]

Eysenck HJ (1967). The Biological Basis of Personality. Springfield, IL: Thomas.

First M, Spitzer R, Gibbon M, Willians J, & Benjamin L (1995). Structured Clinical Interview for DSM-IV Axis I Disorders (SCID) New York: Biometric Research Department, New York State Psychiatric Institute.

Fournier JC, DeRubeis RJ, Shelton RC, Gallop R, Amsterdam JD, & Hollon SD (2008). Antidepressant medications v. cognitive therapy in people with depression with or without personality disorder. British Journal of Psychiatry, 192, 124–129. doi:10.1192/bjp.bp.107.037234 [PubMed: 18245030]

Gignac GE (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. Intelligence, 55(C), 57–68. doi:10.1016/j.intell.2016.01.006

Gorsuch RL (1983). Factor Analysis (2nd ed.). Hillsdale, NJ: Erlbaum.

Gotlib IH, & Joormann J (2010). Cognition and depression: Current status and future directions. Annual Review of Clinical Psychology, 6(1), 285–312. doi:10.1146/annurev.clinpsy.121208.131305

Hames JL, Hagan CR, & Joiner TE (2013). Interpersonal processes in depression. Annual Review of Clinical Psychology, 9(1), 355–377. doi:10.1146/annurev-clinpsy-050212-185553

Hamilton MA (1959). The assessment of anxiety states by rating. British Journal of Medical Psychology, 32, 50–55. [PubMed: 13638508]

Hamilton MA (1960). A rating scale for depression. Journal of Neurology, Neurosurgery, and Psychiatry, 23, 56–62. doi:10.1136/jnnp.23.1.56

Hirschfeld RM, Klerman GL, Clayton PJ, Keller MB, McDonald-Scott P, & Larkin BH (1983). Assessing personality: Effects of the depressive state on trait measurement. The American Journal of Psychiatry, 140(6), 695–699. [PubMed: 6846626]

Holgado-Tello FP, Chacón-Moscoso S, Barbero-García I, & Vila-Abad E (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. Quality & Quantity, 44, 153–166.

Horowitz LM, Rosenberg SE, Baer BA, Ureño G, & Villaseñor VS (1988). Inventory of Interpersonal Problems: Psychometric properties and clinical applications. Journal of Consulting and Clinical Psychology, 56(6), 885–892. doi:10.1037/0022-006x.56.6.885 [PubMed: 3204198]

Hu L.-t., & Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6(1), 1–55. doi:10.1080/10705519909540118

Jang KL, Livesley WJ, Angleitner A, Riemann R, & Vernon PA (2002). Genetic and environmental influences on the covariance of facets defining the domains of the five-factor model of personality. Personality and Individual Differences, 33(1), 83–101. doi:10.1016/S0191-8869(01)00137-4

Klerman GL, Weissman MM, Rounsaville BJ, & Chevron ES (1984). Interpersonal psychotherapy of depression. New York: NY:: Basic Books.

Kotov R, Gamez W, Schmidt F, & Watson D (2010). Linking “big” personality traits to anxiety, depressive, and substance use disorders: A meta-analysis. Psychological Bulletin, 136(5), 768–821. doi:10.1037/a0020327 [PubMed: 20804236]

Lorenzo-Seva U, & ten Berge JMF (2006). Tucker’s congruence coefficient as a meaningful index of factor similarity. Methodology, 2(2), 57–64. doi:10.1027/1614-2241.2.2.57

Manuck SB, Phillips JE, Gianaros PJ, Flory JD, & Muldoon MF (2010). Subjective socioeconomic status and presence of the metabolic syndrome in midlife community volunteers. Psychosomatic Medicine, 72(1), 35–45. doi:10.1097/PSY.0b013e3181c484dc [PubMed: 19933505]

Marsh HW, Hau K-T, & Wen Z (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. Structural Equation Modeling, 11(3), 320–341. doi:10.1207/s15328007sem1103_2

Miller JD, Zeichner A, & Wilson LF (2012). Personality correlates of aggression. Journal of Interpersonal Violence, 27(14), 2903–2919. doi:10.1177/0886260512438279 [PubMed: 22610830]

Molenaar D (2016). On the distortion of model fit in comparing the bifactor model and the higher-order factor model. Intelligence, 57(C), 60–63. doi:10.1016/j.intell.2016.03.007

Morgan G, Hodge K, Wells K, & Watkins M (2015). Are fit indices biased in favor of bi-factor models in cognitive ability research?: A comparison of fit in correlated factors, higher-order, and bi-factor

models via monte carlo simulations. Journal of Intelligence, 3(1), 2–20. doi:10.3390/jintelligence3010002

Murray AL, & Johnson W (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. Intelligence, 41(5), 407–422. doi:10.1016/j.intell.2013.06.004

Muthen LK, & Muthen BO (2012). Mplus: Statistical Analysis with Latent Variables User's Guide. Los Angeles: Muthen & Muthen.

Naragon-Gainey K, & Watson D (2014). Consensually defined facets of personality as prospective predictors of change in depression symptoms. Assessment, 21(4), 387–403. doi:10.1177/1073191114528030 [PubMed: 24671734]

Ormel J, Jeronimus BF, Kotov R, Riese H, Bos EH, Hankin B, . . . Oldehinkel AJ (2013). Neuroticism and common mental disorders: Meaning and utility of a complex relationship. Clinical Psychology Review, 33(5), 686–697. doi:10.1016/j.cpr.2013.04.003 [PubMed: 23702592]

Pashler H, & Wagenmakers EJ (2012). Editors' introduction to the special section on replicability in Psychological Science. Perspectives on Psychological Science, 7(6), 528–530. doi:10.1177/1745691612465253 [PubMed: 26168108]

Pilkonis PA, Kim Y, Proietti JM, & Barkham M (1996). Scales for personality disorders developed from the Inventory of Interpersonal Problems. Journal of Personality Disorders, 10(4), 355–369. doi:10.1521/pedi.1996.10.4.355

Quilty LC, Meusel LA, & Bagby RM (2008). Neuroticism as a mediator of treatment response to SSRIs in major depressive disorder. Journal of Affective Disorders, 111(1), 67–73. doi:10.1016/j.jad.2008.02.006 [PubMed: 18384882]

Reise SP (2012). The rediscovery of bifactor measurement models. Multivariate Behavioral Research, 47(5), 667–696. doi:10.1080/00273171.2012.715555 [PubMed: 24049214]

Reise SP, Moore TM, & Haviland MG (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. Journal of Personality Assessment, 92(6), 544–559. doi:10.1080/00223891.2010.496477 [PubMed: 20954056]

Reise SP, Morizot J, & Hays RD (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. Quality of Life Research, 16(S1), 19–31. doi:10.1007/s11136-007-9183-7 [PubMed: 17479357]

Reise SP, Scheines R, Widaman KF, & Haviland MG (2012). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. Educational and Psychological Measurement, 73(1), 5–26. doi:10.1177/0013164412449831

Rodriguez A, Reise SP, & Haviland MG (2015). Evaluating bifactor models: Calculating and interpreting statistical indices. Psychological Methods, 21(2), 137–150. doi:10.1037/met0000045 [PubMed: 26523435]

Rosellini AJ, & Brown TA (2011). The NEO Five-Factor Inventory: Latent structure and relationships with dimensions of anxiety and depressive disorders in a large clinical sample. Assessment, 18(1), 27–38. doi:10.1177/1073191110382848 [PubMed: 20881102]

Santor DA, Bagby RM, & Joffe RT (1997). Evaluating stability and change in personality and depression. Journal of Personality and Social Psychology, 73(6), 1354–1362. doi:10.1037/0022-3514.73.6.1354 [PubMed: 9418283]

Scott LN, Kim Y, Nolf KA, Hallquist MN, Wright AGC, Stepp SD, . . . Pilkonis PA (2013). Preoccupied attachment and emotional dysregulation: Specific aspects of borderline personality disorder or general dimensions of personality pathology? Journal of Personality Disorders, 27(4), 473–495. doi:10.1521/pedi_2013_27_099 [PubMed: 23586934]

Soldz S, Budman S, Demby A, & Merry J (1995). A short form of the inventory of interpersonal problems circumples scales. Assessment, 2(1), 53–63.

Steiger JH, & Lind JM (1980). Statistically based tests for the number of common factors. Paper presented at the Annual Meeting of the Psychometric Society, Iowa City, IA.

Stern BL, Kim Y, Trull TJ, Scarpa A, & Pilkonis P (2000). Inventory of Interpersonal Problems Personality Disorder scales: Operating characteristics and confirmatory factor analysis in nonclinical samples. Journal of Personality Assessment, 74(3), 459–471. doi:10.1207/S15327752JPA7403_9 [PubMed: 10900572]

Tackett JL, Lilienfeld S, Patrick C, Johnson S, Krueger R, Miller J, . . . Shrout PE (2017). It's time to broaden the replicability conversation: Thoughts for and from clinical psychological science. Perspectives on Psychological Science, 12(5), 742–756. [PubMed: 28972844]

Tackett JL, Quilty LC, Sellbom M, Rector NA, & Bagby RM (2008). Additional evidence for a quantitative hierarchical model of mood and anxiety disorders for DSM-V: The context of personality structure. Journal of Abnormal Psychology, 117(4), 812–825. doi:10.1037/a0013795 [PubMed: 19025228]

Tang TZ, DeRubeis RJ, Hollon SD, Amsterdam J, Shelton R, & Schalet B (2009). Personality change during depression treatment: A placebo-controlled trial. Archives of General Psychiatry, 66(12), 1322–1330. doi:10.1001/archgenpsychiatry.2009.166 [PubMed: 19996037]

Tucker LR, & Lewis C (1973). A reliability coefficient for maximum likelihood factor analysis. Psychometrika, 38(1), 1–10. doi:10.1007/BF02291170

Uliaszek AA, Hauner KKY, Zinbarg RE, Craske MG, Mineka S, Griffith JW, & Rose RD (2009). An examination of content overlap and disorder-specific predictions in the associations of neuroticism with anxiety and depression. Journal of Research in Personality, 43(5), 785–794. doi:10.1016/j.jrp.2009.05.009 [PubMed: 20161016]

Watson D, & Clark LA (1995). Depression and the melancholic temperament. European Journal of Personality, 9(5), 351–366. doi:10.1002/per.2410090505

Whiteside SP, & Lynam DR (2001). The Five Factor Model and impulsivity: using a structural model of personality to understand impulsivity. Personality and Individual Differences, 30(4), 669–689. doi:10.1016/s0191-8869(00)00064-7

Widiger TA, & Trull TJ (1992). Personality and psychopathology: An application of the five-factor model. Journal of Personality, 60(2), 363–393. doi:10.1111/j.1467-6494.1992.tb00977.x [PubMed: 1635047]

Wright AGC, Scott LN, Stepp SD, Hallquist MN, & Pilkonis PA (2015). Personality pathology and interpersonal problem stability. Journal of Personality Disorders, 29(5), 684–706. doi:10.1521/pedi_2014_28_171 [PubMed: 25562539]

Table 1.

Fit of Structural Models of the Neuroticism Scale of the NEO PI-R.

Model	RMSEA [90% CI]	CFI	TLI
Depression Sample			
1. CFA: One Factor	0.080 [0.078 - 0.082]	0.765	0.755
2. CFA: Six Factors	0.061 [0.059 - 0.063]	0.866	0.859
3. Bifactor: One General Factor, Two Specific Factors	0.060 [0.058 - 0.062]	0.874	0.862
4. Bifactor: One General Factor, Three Specific Factors[a]	0.059 [0.057 - 0.061]	0.876	0.865
5. Bifactor: One General Factor, Six Specific Factors	0.050 [0.048 - 0.053]	0.911	0.902
Community Sample			
1. CFA: One Factor	0.079 [0.078 - 0.081]	0.798	0.789
2. CFA: Six Factors	0.064 [0.062 - 0.065]	0.871	0.864
3. Bifactor: One General Factor, Two Specific Factors	0.062 [0.061 - 0.064]	0.881	0.870
4. Bifactor: One General Factor, Three Specific Factors[a]	0.061 [0.060 - 0.063]	0.885	0.874
5. Bifactor: One General Factor, Six Specific Factors	0.053 [0.051 - 0.054]	0.915	0.907

Note. CFA = Confirmatory Factor Analysis.

[a]This model tested a structure similar to the two specific factor structure represented in Model 3, with the addition of a third, depression-related specific factor.

Table 2.

Model 5: Bifactor Solution - 1 General Factor 6 Specific Factors.

		Depression Sample							Community Sample						
NEO PI-R		General	Specific						General	Specific					
Item#	Facet	N.A.	A.	A.H.	D.	S.C.	I.	V.	N.A.	A.	A.H.	D.	S.C.	I.	V.
001	A.	**0.37**	**0.30**						**0.33**	**0.34**					
031	A.	**0.48**	**0.44**						**0.44**	**0.35**					
061	A.	**0.63**	**0.44**						**0.64**	**0.43**					
091	A.	**0.59**	**0.34**						**0.72**	0.19					
121	A.	**0.48**	0.09						**0.53**	**0.35**					
151	A.	**0.61**	**0.36**						**0.63**	**0.36**					
181	A.	**0.49**	**0.43**						**0.44**	**0.46**					
211	A.	**0.48**	0.26						**0.57**	0.24					
006	A.H.	**0.51**		**0.41**					**0.54**		**0.30**				
036	A.H.	**0.43**		**0.64**					**0.47**		**0.65**				
066	A.H.	0.23		**0.77**					**0.38**		**0.69**				
096	A.H.	**0.35**		**0.47**					**0.47**		**0.41**				
126	A.H.	**0.36**		**0.32**					**0.47**		0.26				
156	A.H.	**0.35**		**0.64**					**0.40**		**0.65**				
186	A.H.	**0.49**		0.20					**0.54**		0.13				
216	A.H.	**0.57**		0.25					**0.54**		**0.32**				
011	D.	**0.49**			**0.42**				**0.62**			**0.45**			
041	D.	**0.67**			**0.44**				**0.74**			0.18			
071	D.	**0.54**			**0.42**				**0.68**			**0.52**			
101	D.	**0.55**			0.09				**0.50**			−0.12			
131	D.	**0.54**			0.17				**0.53**			−0.12			
161	D.	**0.71**			**0.32**				**0.75**			−0.03			
191	D.	**0.67**			**0.34**				**0.76**			0.18			
221	D.	**0.73**			0.03				**0.78**			0.00			
016	S.C.	**0.51**				**0.57**			**0.41**				**0.46**		
046	S.C.	**0.50**				**0.47**			**0.45**				**0.30**		

		Depression Sample							Community Sample						
NEO PI-R		General	Specific						General	Specific					
Item#	Facet	N.A.	A.	A.H.	D.	S.C.	I.	V.	N.A.	A.	A.H.	D.	S.C.	I.	V.
076	S.C.	**0.58**				0.09			**0.64**				0.13		
106	S.C.	**0.37**				**0.42**			**0.32**				**0.44**		
136	S.C.	**0.67**				0.21			**0.73**				0.23		
166	S.C.	**0.53**				0.17			**0.50**				0.15		
196	S.C.	**0.46**				**0.33**			**0.43**				**0.41**		
226	S.C.	0.27				0.24			0.15				**0.35**		
021	I.	0.26					**0.65**		0.27					**0.52**	
051	I.	**0.33**					**0.68**		**0.42**					**0.65**	
081	I.	0.13					**0.33**		0.17					0.25	
111	I.	0.18					**0.60**		0.25					**0.59**	
141	I.	0.22					**0.50**		0.22					**0.48**	
171	I.	0.17					**0.53**		**0.40**					**0.45**	
201	I.	**0.38**					**0.34**		**0.51**					**0.32**	
231	I.	**0.48**					0.08		**0.46**					0.13	
026	V.	**0.61**						0.12	**0.64**						0.15
056	V.	**0.67**						0.16	**0.60**						0.25
086	V.	**0.62**						−0.01	**0.67**						0.09
116	V.	0.28						**0.65**	**0.41**						**0.61**
146	V.	**0.54**						0.08	**0.55**						0.17
176	V.	**0.36**						**0.85**	**0.50**						**0.82**
206	V.	**0.50**						0.29	**0.58**						**0.31**
236	V.	**0.72**						0.11	**0.67**						0.15
F.D.		0.96	0.76	0.89	0.74	0.77	0.88	0.91	0.97	0.77	0.88	0.76	0.74	0.86	0.94
Factor Congruencies:									0.99	0.95	0.99	0.75	0.97	0.995	0.99

Note. N.A.=Negative Affectivity, A.=Anxiety, A.H.=Angry Hostility, D.=Depression, S.C.=Self-Consciousness, I.=Impulsiveness, V.=Vulnerability, F.D. = Factor Determinacy. Values represent loadings from an orthogonal bifactor model. Loadings > |0.30| are bolded.

Table 3.

Correlations between Components of Neuroticism and Depression and Anxiety Measures.

	Depression Sample		Community Sample
Neuroticism Component	**Depression**[a]	**Anxiety**[b]	**Depression**[c]
Unit-Weighted Raw Scores			
Total Neuroticism	**0.31 [0.25 – 0.38]**	**0.21 [0.13 – 0.28]**	**0.50 [0.46 – 0.54]**
Neuroticism Facet:			
Anxiety	**0.30 [0.23 – 0.37]**	**0.24 [0.17 – 0.31]**	**0.37 [0.32 – 0.41]**
Angry hostility	**0.17 [0.10 – 0.24]**	**0.13 [0.05 – 0.20]**	**0.39 [0.34 – 0.43]**
Depression	**0.30 [0.23 – 0.37]**	**0.17 [0.09 – 0.24]**	**0.51 [0.47 – 0.55]**
Self-consciousness	**0.19 [0.12 – 0.26]**	**0.14 [0.07 – 0.22]**	**0.33 [0.28 – 0.38]**
Impulsiveness	**0.09 [0.01 – 0.16]**	0.02 [−0.06 – 0.10]	**0.30 [0.25 – 0.35]**
Vulnerability	**0.34 [0.27 – 0.40]**	**0.17 [0.10 – 0.25]**	**0.42 [0.38 – 0.47]**
Bifactor Components			
General Factor: Negative Affectivity	**0.36 [0.28 – 0.44]**	**0.22 [0.15 – 0.29]**	**0.52 [0.49 – 0.56]**
Specific Factor:			
Anxiety	0.06 [−0.05 – 0.16]	**0.18 [0.11 – 0.25]**	**−0.08 [−0.13 – −0.02]**
Angry hostility	−0.04 [−0.12 – 0.04]	0.02 [−0.05 – 0.10]	0.03 [−0.02 – 0.07]
Depression	0.04 [−0.05 – 0.14]	−0.01 [−0.09 – 0.06]	**0.11 [0.07 – 0.16]**
Self-consciousness	**−0.18 [−0.27 – −0.08]**	0.01 [−0.07 – 0.08]	**−0.11 [−0.17 – −0.05]**
Impulsiveness	**−0.11 [−0.19 – −0.03]**	−0.07 [−0.14 – 0.01]	0.04 [−0.01 – 0.08]
Vulnerability	0.11 [0.03 – 0.19]	0.01 [−0.06 – 0.09]	**−0.10 [−0.16 – −0.03]**

Note. Values represent correlation coefficients and 95% confidence intervals.

[a]For the depression sample, *n*=732 provided data on the depression severity measure (the Hamilton Rating Scale for Depression).

[b]Participants from the Toronto site, *n* = 658, provided anxiety symptom data on the SCID (Structural Clinical Interview for DSM Disorders). Anxiety symptoms were quantified as the total count of symptoms endorsed in Section F: Anxiety Disorders. Because the distribution of the data was non-normal, values represent Spearman's rank order correlations.

[c]For the community sample, *n*=1276 provided data on the depression severity measure (the Beck Depression Inventory). Bolded values are statistically significant at p<0.05.

Table 4.

Criterion Validity: Associations with Interpersonal Problems in the Community Sample.

	Symptoms	General			Specific			
IIP Subscales[a]	**Dep.**	**N.A.**	**A.**	**A.H.**	**D.**	**S.C.**	**I.**	**V.**
Interpersonal Sensitivity	**0.33*****	**0.40*****	0.11**	−0.01	−0.08**	**0.33*****	0.07**	−0.01
Interpersonal Ambivalence	0.22***	0.16***	−0.13**	0.10**	0.05	−0.06	0.05†	0.01
Aggression	0.24***	0.29***	−0.13***	**0.51*****	0.00	−0.19***	0.13***	−0.01
Need for Approval	0.25***	**0.40*****	0.12**	−0.18***	−0.14***	**0.32*****	0.04	−0.05
Lack of Sociability	**0.33*****	**0.36*****	0.00	−0.11***	−0.04	**0.37*****	−0.07*	0.03

Note. IIP = Inventory of Interpersonal Problems, Dep. = Depression symptoms, N.A.=Negative Affectivity, A.=Anxiety, A.H.=Angry Hostility, D.=Depression, S.C.=Self-Consciousness, I.=Impulsiveness, V.=Vulnerability. Values represent standardized beta estimates from structural equation models in which all of the independent variables, displayed in the columns, were examined simultaneously. Separate models were estimated for each IIP subscale (row). N=1282 for each IIP subscale except for Interpersonal Ambivalence and Lack of Sociability, for which N=1281. Values > 0.30 are bolded.

†=p<0.10,

*=p<0.05,

**=p<.01,

***p<0.001