



The higher-order structure of common DSM mental disorders: internalization, externalization, and their connections to personality

Robert F. Krueger *, Matt McGue, William G. Iacono

Department of Psychology, University of Minnesota, 75 E. River Road, Minneapolis, MN 55455, USA

Received 9 February 2000; received in revised form 20 April 2000

Abstract

Comorbidity among mental disorders is commonly observed in clinical and epidemiological samples. Can comorbidity be understood as meaningful covariance, and is this covariance structure linked with personality? We addressed this question in a sample of 634 female and 549 male, middle-aged participants in the Minnesota Twin-Family Study (MTFS). Mental disorders were assessed using the Structured Clinical Interview for DSM-III-R, the Substance Abuse Module from the Composite International Diagnostic Interview, and a specially-designed interview for the assessment of antisocial personality disorder. Personality was assessed using the Multidimensional Personality Questionnaire. Relations among symptom scales for eight common DSM disorders were compatible with hypothesized underlying bivariate normal distributions. Polychoric correlations among these scales were well-fit by a two-factor model positing *internalizing* and *externalizing* factors, which, in turn, were correlated with broad personality dimensions. Internalizing was positively correlated with *negative emotionality* (and negatively with *positive emotionality* in women) and externalizing was negatively correlated with *constraint*.

These findings suggest that internalization, externalization, and their links to personality may provide a useful framework for understanding covariance among common adult mental disorders. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Personality traits; Mental disorders; DSM; Internalizing disorders; Externalizing disorders; Comorbidity

* Corresponding author. Tel.: +1-612-624-8204; fax: +1-612-626-2079.

E-mail address: krueg038@tc.umn.edu (R.F. Krueger).

1. Introduction

In recent history, personality and mental disorders have been studied in distinct and non-overlapping literatures. This historical bifurcation, however, is clearly waning, as the field bears witness to a rapprochement between personality and psychopathology research. At one time, these fields were essentially united as topics of study under the broad rubric of human individual differences. Now that the age of doubt regarding the reality and utility of personality is essentially over, personality and psychopathology research are again showing signs of cross-fertilization (Watson & Clark, 1994).

In the research reported herein, we pursued the rapprochement between personality and psychopathology. We examined the higher-order structure of common mental disorders in a community-based sample of men and women, and we mapped this higher-order structure onto the higher-order structure of personality. Thus, our analyses addressed two basic issues. First, could “comorbidity” among mental disorders be understood as meaningful covariance, and hence, modeled as the result of basic underlying factors of psychopathological variation? Second, how did these basic underlying factors of psychopathological variation map onto higher-order personality traits? To set the stage for our analysis, we briefly outline the historical context of our work.

One key theme underlying the renewed interest of psychopathologists in personality has been the strong resurgence of personality psychology as a vital academic enterprise. This resurgence was based on a number of key research findings and closely-related conceptual developments. First, personality traits were found to be heritable. Evidence from twin studies, for example, is consistent in documenting broad heritabilities of 30–50% for higher-order dimensions of personality (Loehlin, 1992; Bouchard, 1994). In addition, molecular genetic work is beginning to bear fruit, as specific markers (e.g., DRD4) are being linked to personality variation (Plomin & Caspi, 1998). If personality traits are fictional, they are fictions with a notable genetic basis.

Second, personality traits were found to be stable in adulthood. Summarizing studies that varied markedly in instrumentation, sample size, gender composition, initial age of participants, and retest interval (two to thirty years), Costa and McCrae (1997) found that the median stability of different personality traits ranges from 0.34 to 0.77. As these authors point out, disattenuated for unreliability, these correlations would be even higher.

Finally, personality traits were found to be consequential. One could list a compendium of important social outcomes that are meaningfully correlated with personality traits, but such a “laundry list” does not do justice to the *systematic* nature of the relevant relations. Personality trait models provide the “glue” for a comprehensive psychological theory linking individual differences in domains such as affect and impulse control to consequential social outcomes (Watson, Clark & Harkness, 1994). For example, why do substance use disorders and antisocial behavior tend to occur in the same persons (Kessler et al., 1997; Morgenstern, Langenbucher, Labouvie & Miller, 1997; Rounsaville et al., 1998)? Perhaps because low levels of constraint — a tendency toward impulsivity, sensation seeking, and non-traditional values — lead to behaviorally distinctive outcomes that, nevertheless, are psychologically coherent when reframed as indicative of impulsivity (Sher & Trull, 1994; Krueger, 1999b). In this way, personality trait models can predict and explain specific social behaviors, and the relations among such behaviors.

While these developments in personality psychology were underway, psychopathology research witnessed the rise of the modern Diagnostic and Statistical Manuals, i.e., DSM-III (American

Psychiatric Association, 1980), DSM-III-R (American Psychiatric Association, 1987), and DSM-IV (American Psychiatric Association, 1994), and their operationalization in various structured interviews. One of the most interesting findings from research using structured interviews is the pervasiveness of “comorbidity” among DSM-defined mental disorders — the tendency for mental disorders to co-occur at greater than chance rates (Clark, Watson & Reynolds, 1995). This phenomenon is usually viewed through the lens of the categorical, “neo-Kraepelinian” measurement model underlying the modern DSMs. From this perspective, comorbidity is typically viewed as either a crisis or as a nuisance. Conceptually, comorbidity is a crisis in that it undermines the neo-Kraepelinian idea that there are various, separate, and discrete mental disorders. Operationally, comorbidity is a nuisance because various gymnastics are necessary to obtain samples for research, such as identifying rare “pure” individuals in highly comorbid treatment-seeking samples.

Rather than viewing comorbidity as a crisis or a nuisance, it may be better viewed as an opportunity. Comorbidity suggests that something is awry with the conceptual measurement model underlying recent DSMs, and provides the impetus for improving that model. It is not, however, a threat to the key, seminal contribution of the neo-Kraepelinians: The notion that mental disorders are real, and can therefore be reliably measured. Indeed, if mental disorders were not reliably measurable, comorbidity would not be possible (i.e., unreliability of measurement sets a ceiling on covariance among disorders).

Comorbidity is, however, a problem that is difficult to cope with under the categorical, neo-Kraepelinian model. Nevertheless, no specific measurement model has made a strong bid to take the place of the categorical model. Psychometric theory may provide such a model, for at least some psychopathological conditions. For instance, in recent years, psychometric methods have begun to influence conceptualizations of the Axis II (personality) disorders (Cloninger, 1999; Costa & Widiger, 1994; Strack & Lorr, 1994). Although agreement in this literature is not universal, researchers studying personality disorders have begun to realize the advantages of the methods utilized by personality researchers studying normal-range personality variation. Specifically, personality psychologists have historically coped with phenotypic diversity and covariance among indicators by using tools developed by psychometricians, such as factor analysis. These tools have served personality research well. Although researchers may still disagree about the precise number and nature of the basic dimensions of personality (Eysenck, 1992; Livesley, Jang & Vernon, 1998), the notion that a small number of basic higher-order dimensions are jointly necessary and sufficient to summarize personality variation is fairly incontrovertible (Watson et al., 1994).

Does this perspective also have utility in understanding Axis I (syndrome) mental disorders? Recently, Krueger applied the psychometric perspective to the comorbidity problem in research on common Axis I mental disorders in the Dunedin Multidisciplinary Health and Development Study (DMHDS; Krueger, Caspi, Moffitt & Silva, 1998) and in the National Comorbidity Survey (NCS; Krueger, 1999a). Epidemiological research shows that certain DSM-defined mental disorders occur with greater frequency in community samples. Specifically, prevalence rates for unipolar affective disorders, anxiety disorders, substance use disorders, and antisocial behavior disorders exceed prevalence rates for “severe” psychopathologies, such as mania and the non-affective psychoses (Kessler et al., 1994). Moreover, the common DSM disorders are systematically covariant. Krueger showed how two higher-order, psychologically-coherent dimensions

of variation, *internalization* and *externalization*, were able to account for this systematic covariance. The factor connecting the unipolar mood and anxiety disorders was labeled *internalization*, to describe the propensity to express distress inwards that unites these disorders. The factor connecting the substance use and antisocial behavior disorders was labeled *externalization*, to describe the propensity to express distress outwards that unites these disorders. Importantly, internalization and externalization are conceived as separate dimensions, as opposed to opposite ends of the same dimension.

In the work reported herein, we attempted to replicate and extend this previous research on the structure of common mental disorders, and to connect this line of research to research on the structure of personality. Specifically, we employed data from middle-aged parent participants in the Minnesota Twin-Family Study (MTFS) to build upon Krueger's previous work in a number of specific ways. First, participants in the DMHDS were interviewed using the Diagnostic Interview Schedule (DIS; Robins, Helzer, Cottler & Goldring, 1989), and participants in the NCS were interviewed using the University of Michigan version of the Composite International Diagnostic Interview (UM-CIDI; Wittchen, Kessler, Zhao & Abelson, 1995). In the MTFS, a different set of diagnostic instruments — the Structured Clinical Interview for DSM-III-R (SCID; Spitzer, Williams & Gibbon, 1987), the Substance Abuse Module from the Composite International Diagnostic Interview (SAM; Robins, Babor & Cottler, 1987) and a specially-designed interview for the assessment of antisocial personality disorder (Holdcraft, Iacono & McGue, 1998) were employed. In addition, interview data collected in the MTFS were submitted to a consensus process, whereas DMHDS and NCS diagnoses were made directly from interview data. Thus, we were able to determine the robustness of Krueger's model to variation in diagnostic instrumentation and method.

Second, Krueger et al. (1998) and Krueger (1999a) employed the tetrachoric correlation as an index of association between DSM diagnoses. The tetrachoric correlation operationalizes a *liability-threshold* model of disorder risk. This model states that a normally distributed continuum of risk underlies observed dichotomous mental disorder diagnoses. When only dichotomous diagnoses are used, there is no direct way to test the appropriateness of the liability-threshold model. Therefore, in the current research, we employed multi-point ordinal scales representing the number of criteria participants met toward a given DSM disorder. When more than two categories are available for at least one of the two variables being correlated, it is possible to test the fit of the underlying bivariate normal distribution to the frequency table formed by the cross-classification of participants on the two ordinal variables. Thus, the appropriateness of viewing bivariate associations among DSM variables (a.k.a. comorbidity) as jointly normal and continuous was submitted to statistical scrutiny.

Finally, Krueger et al. (1998) and Krueger (1999a) speculated about relations between internalization, externalization, and broad traits identified in the personality literature, but presented no empirical data with regard to this issue. Specifically, Krueger et al. (1998) speculated that internalization might be linked with *negative emotionality* (a tendency to experience anxiety, anger, and alienation), and externalization with a lack of *constraint* (a tendency to engage in risky behavior, to act on impulse, and to endorse non-traditional values). Negative emotionality and constraint are constructs measured by Tellegen's Multidimensional Personality Questionnaire (Tellegen, 1985, 2000; Tellegen & Waller, 2000), which was completed by the MTFS parents. Thus, in the current research, we were able to directly examine relations between internalization, externalization, and the higher-order personality traits measured by the MPQ.

2. Method

2.1. Participants

Participants were parents of adolescent Minnesota-born twins enrolled in the Minnesota Twin-Family Study (MTFS), an epidemiological investigation of the development of substance abuse and related mental disorders. Although parents of both male and female twins are enrolled in the MTFS, for this report, data from the parents of female twins were used because, at intake, these parents received a more extensive mental disorder assessment than parents of male twins. Twin families were ascertained from State of Minnesota birth records and located using various public databases including telephone directories and driver license registrations. For any given birth year, more than 90% of twin births have been located. Families were excluded from participation if they lived further than a day's drive from our Minneapolis laboratories, or if either of the twins in the family had a physical or cognitive handicap that would preclude their completing our day-long, in-person intake assessment. Among eligible families, 17% refused participation, with 78% of the non-refusing families ultimately completing the intake assessment. Through telephone interviews, mail surveys, and access to information recorded on the birth records, we were able to determine the socioeconomic status and complete brief mental health assessments on 82.6% of the non-participating families. Participating families differed minimally from non-participating families, with the largest differences occurring on measures of socioeconomic status (parents in participating families averaged 0.25 more years of education than non-participating parents). Consistent with the demographics of the state of Minnesota, over 98% of participating parents were Caucasian. For this report, participants were required to have no missing data for both the mental disorder and personality assessments, resulting in list-wise sample sizes of 634 mothers and 549 fathers. Further details on the MTFS design, sample, and assessment may be found in Iacono, Lykken, and McGue (1996).

2.2. Measures

2.2.1. Assessment of mental disorders

Participants were assessed in-person by trained bachelors and masters-degree-level interviewers for lifetime DSM-III-R mental disorders upon intake into the study (DSM-III-R was current at the time the study began). Specific affective and anxiety disorders were assessed using the Structured Clinical Interview for DSM-III-R (SCID; Spitzer et al., 1987). Drug and alcohol disorders were assessed using the Substance Abuse Module from the Composite International Diagnostic Interview (SAM; Robins et al., 1987). Antisocial personality disorder symptoms were assessed using an interview designed by the MTFS staff (Holdcraft et al., 1998). Prior to the assignment of mental disorder symptoms, a clinical case conference was completed in which the evidence for every symptom was reviewed (including, if necessary, the replaying of audio tapes from the interviews) by at least two advanced graduate students in clinical psychology. The reliability of the consensus process was excellent, with kappas for each of the DSM diagnoses studied herein exceeding 0.77.

After consensus was reached, computer algorithms were used to create symptom counts corresponding to criteria for DSM-III-R disorders. Criteria were included in the symptom counts if

they referred to symptoms of disorders, rather than symptom duration or hierarchical exclusionary rules. Symptoms were not considered present if they were deemed due to an organic etiology or bereavement (in the case of major depression). Specifically, eight symptom count variables were employed in the current research. These variables corresponded to: (1) the 10 criterion C (adult antisocial behavior) symptoms of antisocial personality disorder (AAB); (2) the 9 criterion A symptoms of major depressive episode (MDE); (3) the 13 criterion C and 1 criterion D symptoms of panic disorder (PD); (4) criteria A, C, D, E, and F of social phobia (SOP); (5) criteria A-E of simple phobia (SIP); and (6, 7, 8) the nine criterion A symptoms of alcohol dependence (AD), cannabis dependence (CD), and drug dependence (DD). (The drug dependence variable represented the average of symptom counts across amphetamine, cocaine, hallucinogens, opioids, and sedatives.)

2.2.2. *Assessment of personality*

Personality was assessed using the 198-item version of the Multidimensional Personality Questionnaire (MPQ; Tellegen, 2000). The MPQ measures 11 primary personality traits; these 11 traits, in turn, form a three-factor higher-order structure. The first higher-order factor of the MPQ has been called *positive emotionality (PEM)*, and refers to a propensity to experience positive emotions resulting from active engagement in work and social environments. Primary scales labeled wellbeing, social potency, achievement, and social closeness load highly on this factor. The second MPQ factor has been called *negative emotionality (NEM)*, and refers to a propensity to experience negative mood states such as anxiety, alienation from others, and anger. Primary scales labeled stress reaction, alienation, and aggression load highly on this factor. The third MPQ factor has been called *constraint (CON)* and refers to a propensity for cautious and restrained behavior and endorsement of traditional values. Primary scales labeled control, harmavoidance, and traditionalism load highly on this factor. The final scale, absorption, does not load primarily on a specific higher-order factor.

MPQs were mailed to parents prior to their scheduled in-person visit. Parents were asked to return the MPQ by mail or to bring the MPQ along when they visited in-person. If we failed to obtain a completed MPQ by the time of the in-person assessment, the parent was asked to complete it at home and return it by mail. One additional telephone reminder was made if we still did not receive a completed MPQ. We were more likely to receive a completed MPQ from women than from men (91% vs 85%). In addition, we compared MPQ-returners vs non-returners on the eight DSM scales, separately for men and women. Only one of 16 t-tests was significant: male non-returners reported slightly higher levels of major depression than returners [$t(144.68, \text{unequal variances}) = 2.23, P = 0.03$]. As this was about the number of significant tests expected by chance, we concluded that MPQ-returners were representative of the overall sample on the DSM scales.

3. Results

3.1. *Appropriateness of the liability-threshold model*

Does the liability-threshold model hold for pairwise relations among the DSM symptom-count variables? The computer program PRELIS 2.2 (Jöreskog & Sörbom, 1996b) was used to create

contingency tables for all possible pairings of the eight DSM variables. That is, for each pairing of variables (e.g., *A* with *x* scale points, and *B* with *y* scale points), PRELIS creates a table with dimensions *x* by *y* showing the number of persons who fell in each cell of the table, due to their combination of scores on the two variables (*A* and *B*). For each such contingency table, PRELIS computes a chi-square goodness of fit statistic. The null hypothesis for these chi-square tests is that a bivariate normal distribution underlies the observed pattern of cell counts. If the chi-square is not significant (in a large sample), it is reasonable to conclude that the observed contingency table is compatible with the hypothesis of an underlying bivariate normal distribution.

There were 28 pairings among the eight DSM variables. For women, none of the 28 chi-square tests were significant (all *P* values > 0.07). For men, only one of these 28 tests was significant (*P* = 0.02). Thus, across men and women, the number of significant chi-square tests was less than the number expected by chance. We concluded from this that the liability-threshold model was reasonable for the DSM variables in our study.

3.2. Structure of DSM symptom counts

Because we had no evidence to reject the liability-threshold model for the bivariate associations among DSM variables, we computed polychoric correlations among these variables (Table 1). The polychoric correlation is an estimate of the correlation between the two hypothesized, normally-distributed liability variables underlying the two observed ordinal variables. We used the method of Weighted Least Squares (WLS) in LISREL 8.2 (Jöreskog & Sörbom, 1996a) to fit factor models to these polychoric correlations and their asymptotic variances and covariances. (Asymptotic covariance matrices are necessary for analysis of polychoric correlation matrices in LISREL. Because these are very large matrices, they are not reprinted here but may be obtained from the authors upon request.)

Specifically, we fit three different factor models to the polychoric correlations among the eight DSM variables, separately for men and women. The first model was a one factor model, in which all DSM variables loaded on a single “general psychopathology” factor. The second model was a two-factor (internalization and externalization) model where MDE, PD, SIP, and SOP were

Table 1
Polychoric correlations for eight DSM symptom scales^a

	MDE	PD	SIP	SOP	AD	CD	DD	AAB
MDE	1.00	0.37	0.12	0.17	0.14	0.16	0.19	0.19
PD	0.34	1.00	0.09	0.21	0.10	0.27	0.35	0.14
SIP	0.11	0.15	1.00	0.29	0.06	0.02	0.06	0.10
SOP	0.09	0.17	0.21	1.00	0.17	0.03	0.05	0.11
AD	0.06	0.21	0.09	0.12	1.00	0.56	0.63	0.64
CD	0.20	0.10	-0.07	-0.02	0.50	1.00	0.71	0.63
DD	0.01	0.21	0.05	0.00	0.59	0.75	1.00	0.65
AAB	0.10	0.16	0.05	0.10	0.55	0.60	0.72	1.00

^a Correlations for women (*n* = 634) are above the diagonal, correlations for men (*n* = 549) are below the diagonal. MDE = major depressive episode; PD = panic disorder; SIP = simple phobia; SOP = social phobia; AD = alcohol dependence; CD = cannabis dependence; DD = drug dependence; AAB = adult antisocial behavior.

indicators of internalization, and AAB, AD, CD, and DD were indicators of externalization. The third model was a four-factor, DSM-derived model where MDE indicated an affective disorder factor, PD, SIP, and SOP indicated an anxiety disorder factor, AD, CD, and DD indicated a substance dependence factor, and AAB indicated an antisocial behavior factor.

We compared the fit of our models using information-theoretic measures of fit: Akaike's Information Criterion ($AIC = \chi^2 - 2 \text{ df}$; Akaike, 1987) and the Bayesian Information Criterion ($BIC = \chi^2 - \text{df} \ln n$; Raftery, 1995). Information-theoretic fit indices are useful in comparative model fitting because they balance fit (in the sense of a small chi-square) with parsimony (in the sense of a *large* number of degrees of freedom) in determining the "best" model. Larger, negative values of BIC and AIC are associated with better-fitting models.

For women, the two factor model provided the best fit to the data [χ^2 ($\text{df} = 19, n = 634$) = 33.11, $P = 0.02$, $AIC = -4.89$, $BIC = -89.44$]. In comparison, the one-factor model [χ^2 ($\text{df} = 20, n = 634$) = 81.18, $P < 0.01$, $AIC = 41.18$, $BIC = -47.82$] and the four-factor model [χ^2 ($\text{df} = 16, n = 634$) = 31.52, $P = 0.01$, $AIC = -0.48$, $BIC = -71.68$] fit less well. Similarly, for men, the two factor model provided the best fit to the data [χ^2 ($\text{df} = 19, n = 549$) = 29.24, $P = 0.06$, $AIC = -8.76$, $BIC = -90.65$]. In comparison, the one-factor model [χ^2 ($\text{df} = 20, n = 549$) = 55.86, $P < 0.01$, $AIC = 15.86$, $BIC = -70.34$] and the four-factor model [χ^2 ($\text{df} = 16, n = 549$) = 29.03, $P = 0.02$, $AIC = -2.97$, $BIC = -71.93$] fit less well.

3.3. Linking internalization and externalization to broad personality dimensions

How do internalization and externalization map onto the three higher-order dimensions of the MPQ? To address this issue, we used principal components analysis (PCA) to compute component scores (separately for men and women) on internalization (I) and externalization (E).¹ First, the DSM polychoric correlation matrices were submitted to PCA. The results from these analyses dovetailed well with the results from the confirmatory factor analyses. Specifically, examination of the eigenvalues of the matrix (and the scree plot) suggested two components for both men and women. For men, the first two eigenvalues were 2.97 and 1.46 (the third was marginal with a value of 1.04). When extracted and rotated via varimax, the first two components accounted for 55.4% of the variance and were clearly identifiable as E and I, respectively. Component I (E) was defined by strong loadings on AAB (0.84), AD (0.75), CD (0.85) and DD (0.91), a low secondary

¹ An alternative approach to this issue would involve simultaneous structural modeling of the DSM variables, the MPQ variables, and the interrelations among the latent factors underlying the DSM and MPQ variables. We chose not to take this approach because analysis of personality trait data via confirmatory structural modeling is controversial, due to the fact that such models tend not to fit very well (unless multiple, complex, and hard-to-interpret parameters — such as cross-loadings and correlated uniquenesses — are added to the model; Church & Burke, 1994; McCrae, Zonderman, Costa, Bond, & Paunonen, 1996). This complexity regarding the structure of personality is tangential to our purpose. Thus, we turned to PCA as a robust approach, familiar to personality psychologists, that essentially accomplishes the same goal as structural modeling with one important exception: Unlike paths linking latent variables in structural models, estimates of correlations between component scores and other variables will be attenuated by random error variance. This limitation should be borne in mind when considering the magnitude of the correlations we report between component scores and NEM, PEM and CON. We also note that we used PCA instead of exploratory factor analysis (EFA) because in PCA (as opposed to EFA) communalities are not estimated, thereby circumventing the factor score indeterminacy problem.

loading on PD (0.19) and zero loadings on the remaining DSM variables. Component II (I) was defined by strong loadings on MDE (0.60), PD (0.68), SIP (0.59), and SOP (0.59), a low secondary loading on AD (0.17) and zero loadings on the remaining DSM variables.

For women, the results were very similar. The first two eigenvalues for the women's matrix were 3.13 and 1.44 (the third was marginal with a value of 1.06). When extracted and rotated via varimax, the first two components accounted for 57.1% of the variance and were clearly identifiable as E and I, respectively. Component I (E) was defined by strong loadings on AAB (0.84), AD (0.81), CD (0.86) and DD (0.87), low secondary loadings on PD (0.24) and MDE (0.16), and zero loadings on the remaining DSM variables. Component II (I) was defined by strong loadings on MDE (0.62), PD (0.61), SIP (0.59), and SOP (0.70), a low secondary loading on DD (0.14) and zero loadings on the remaining DSM variables. We used the component score coefficient matrices from these analyses (computed via the regression method) to compute component scores for men and women on E and I.²

Pearson product-moment correlations were then computed between component scores on I and E, and PEM, NEM, and CON, separately for men and women. These correlations can be seen in Table 2, which shows a clean convergent–discriminant pattern of correlations, with significant links between NEM and I and CON and E, respectively. The one exception is the significant (albeit small) negative correlation between PEM and I for women. To explore this finding further, significant correlations (PEM with I, NEM with I and CON with E) were compared across men and women via Fisher's $r \rightarrow z$ transformation. The PEM-I correlations were marginally different for men vs women ($z = 1.91$, $P = 0.06$), and the NEM-I ($z = 0.84$) and CON-E ($z = -1.35$) correlations were not significantly different for men vs women.

Table 2

Pearson product-moment correlations between internalization, externalization, positive emotionality, negative emotionality, and constraint^a

	Positive emotionality	Negative emotionality	Constraint
Men ($n = 549$)			
Internalization	0.00	0.27*	0.03
Externalization	-0.04	0.04	-0.28*
Women ($n = 634$)			
Internalization	-0.12*	0.22*	0.04
Externalization	0.02	0.00	-0.20*

^a Note: * $P < 0.01$.

² We also performed these principal component analyses on Pearson product-moment correlations among the DSM variables. The results produced by this approach were very similar to the results utilizing polychoric correlations. The intercorrelations between the component scores derived using both approaches were 0.99, for both the internalizing and externalizing components, for both men and women.

4. Discussion

We investigated the latent structure of variables representing the criteria for eight common DSM mental disorders in a large, representative sample of middle-aged men and women. We found no reason to reject a liability-threshold model to account for the bivariate relations among these variables. The hypothesis that normally-distributed liabilities underlie observed patterns of co-risk for these eight DSM variables could not be rejected. This was a key finding because it allowed us to compare the fit of specific continuous factor models to explain correlations among the eight variables. Our modeling efforts (using both confirmatory factor modeling and exploratory principal components analysis) provided solid evidence for the existence of two broad underlying dimensions, internalization (I) and externalization (E), in both men and women. Internalization (I) was linked with higher negative emotionality (NEM) — and lower positive emotionality (PEM) in women — and E was linked with lower constraint (CON). These findings thus replicate Krueger's earlier work (Krueger et al., 1998; Krueger, 1999a) across instruments, diagnostic procedures, and level of measurement (i.e., dichotomous vs ordinal), and extend these findings by linking dimensions of mental disorder with dimensions of personality.

In considering our results, some important limitations should be borne in mind. First, we studied only a limited range of DSM psychopathology. The disorders we studied, however, represent the “common” mental disorders (i.e., unipolar affective disorder, anxiety disorder, substance use disorder, and antisocial behavior disorder) — the broad classes of disorder that have higher prevalence rates in epidemiological samples, such as the sample studied herein (Kessler et al., 1994). Future research, however, could examine a broader range of disorders by studying clinical samples, where other syndromes are likely to be observed.

Second, we modeled only eight DSM variables. Would more factors be identified if more primary variables (e.g., other anxiety disorders, other forms of substance dependence) were modeled? This is possible, but we would be surprised if the first two factors emerging from such analyses could not be identified as I and E. We hypothesize that, like personality traits, common psychopathological syndromes are organized hierarchically. I and E are super-ordinate organizing axes of common psychopathological variation, as has been found consistently in factor-analytic studies of child and adolescent behavioral and emotional problems (Achenbach & Edelbrock, 1978, 1984). However, finer grained *facets* may be identified at a level between symptoms and the I/E super-factor level. As noted by Watson et al. (1994), models that “split” variation into multiple facets, and models that “lump” variation into broad, super-ordinate dimensions are hierarchically nested. Identifying the facets of broad dimensions in a hierarchical model is akin to extracting more factors from the matrix of correlations among primary variables. In this way, “comorbidity” (covariance among mid-level entities) is the logical result of the hierarchical structure of common mental disorders.

Along these lines, consider factor-analytic work on the clinical and validity scales of the MMPI (Graham, 2000), which might be conceived of as “facets of psychopathology”. The first broad factor emerging from these analyses is often interpreted as “general psychological maladjustment”, involving pessimism, anxiety, and apathy. To our reading, this factor resembles I and NEM. The second factor is often interpreted as conventionality and formality. To our reading, this factor resembles E (reflected) and CON.

Waller (1999), however, noted that previous work in this area has been hampered by the problem of multiple overlapping items on the MMPI clinical and validity scales (thus introducing spurious components of variance into the correlations among these scales). Thus, Waller began with an item-level factor analysis of the MMPI (using tetrachoric correlations), and created 16 “pure” (without overlapping items) scales to measure the factors derived from this analysis. He then extracted three meaningful super-factors from the intercorrelations of these 16 scales. These three factors (by our reading) bear a notable resemblance to NEM and I, PEM, and CON (reflected) and E, respectively. Thus, the DSM, the MMPI, and the MPQ seem to converge on similar higher-order factors, even though all three systems measure diverse “primary traits”. From this, we conclude that the influence of broad I and E-like factors on common psychopathological variation is inescapable, even when the primary indicators stem from distinctive measurement traditions.

Although PEM emerges from the MMPI and the MPQ, it can not be seen clearly in our DSM analyses (i.e., we find only two replicable factors, I and E, linked with NEM and CON, respectively). In addition, PEM had an inconsistent relationship with the DSM factors, being weakly linked to I for women but not for men — a borderline-significant gender difference. One possible interpretation of this finding might involve gender differences in I-levels, i.e., a higher prevalence of anxiety-depression in women vs men (see Weich, Sloggett & Lewis, 1998, for recent evidence). The relation between PEM and I may not be seen in men because of a restricted range of I-levels in men, relative to women (note also that male MPQ-returners in our sample were slightly less depressed than MPQ non-returners). This speculation should be revisited if the weak PEM-I link and gender difference found herein are strengthened via replication in future studies.

Finally, a third weakness stems from the correlations we found linking personality and mental disorder, which were small. There may be a number of reasons for this finding. First, the correlations are not disattenuated for random error variance. Second, personality and mental disorder were measured using rather different approaches, a self-report questionnaire filled out on the participant’s own time, and an intensive interview completed in our laboratories. These method differences may contribute to diminished relations among constructs. Thus, one strategy for evaluating the closeness of the link between personality and mental disorder would involve multi-method assessment of both constructs. For example, personality could be measured by non-traditional means (e.g., via clinical interview or other-report), as could the symptoms of mental disorder (e.g., via self-report or other-report). A multi-trait, multi-method approach may prove enlightening in determining the “true” magnitude of the relation between personality and mental disorder. Finally, there may be range restriction on our measurement of I and E. We conceive of these dimensions as spanning a wide range of functioning, but we suspect that the DSM is measuring only the higher, maladaptive ends of both dimensions. This range restriction may place a limit on correlations between I and E and the “normal range” personality variables measured by the MPQ.

Bearing these concerns in mind, our findings offer insight into reasons why comorbidity is notable throughout the DSM system. Specifically, we found that “comorbidity” could be modeled by hypothesizing the existence of broad, continuous variables underlying observed patterns of correlation among DSM constructs. These broad variables, in turn, were linked to broad variables from the personality literature. Thus, one could hypothesize that comorbidity occurs because basic dimensions of temperamental variation confer risk for a broad range of maladaptive

outcomes. This model has been referred to as a *vulnerability model* of the relationship between personality and mental disorder — a model in which personality contributes to the risk of experiencing mental disorder (Akiskal, Hirschfeld & Yerevanian, 1983; Eysenck, 1994; Klein, Wonderlich & Shea, 1993; Clark, Watson & Mineka, 1994). In this respect, each DSM disorder contains both a specific-non-temperamental component of variance, and a general-temperamental component of variance. Consider substance dependence as an example. The likelihood of an individual developing dependence on a specific substance is a function of both temperament and non-temperamental, specific circumstances. For example: We hypothesize that a lack of constraint is a persistent risk factor for substance dependence across time and circumstance, but secular and sub-cultural changes in the availability and popularity of specific substances (or for that matter, other reinforcers) will modify the *apparent* phenotype. The synthesis and street availability of a new designer drug (“Drug Z”) could result in a rash of cases of “Z-dependence”, but we would anticipate that the persons willing to get involved with Z in the first place are those lacking in constraint. The fundamental problem (a.k.a. distal etiology) in this scenario is not Drug Z, but low constraint. In this sense, attending to the sequela of Z-dependence (perhaps, e.g., severe physiological withdrawal) is an important initial component of the treatment plan, but it would be a mistake to not follow-up by addressing the way in which low constraint renders the person vulnerable to the next powerful reinforcer they encounter (gambling, Drug Y, and so on).

How, then, do our findings and perspective relate to previous work on the link between personality and mental disorder? Most previous work on personality and common mental disorders has focused on single disorders in single studies (for reviews of research on anxiety and depression, see Clark et al., 1994; for substance and antisocial disorders, see Sher & Trull, 1994). Single-disorder studies are problematic in the face of “comorbidity”, and little work has examined multiple disorders and multiple traits in the same sample. The “multi-trait/multi-diagnosis” work that has been conducted, using both the “Big five” (Trull & Sher, 1994) and Tellegen (Krueger, Caspi, Moffitt, Silva & McGee, 1996) personality models, has dealt with comorbidity in less-than-satisfactory ways, controlling for its influence or presenting findings separately for “comorbid vs pure” groups. Our work herein suggests that these approaches are not entirely satisfactory because a primary *cause* of comorbidity might be broad personality traits. That is, the typical design examines the link between personality and a specific mental disorder or multiple mental disorders, not the link between personality and the covariance among disorders (i.e., comorbidity), treated as meaningful variance (i.e., I and E factors). Evaluating further the possibility that personality is a major contributor to comorbidity will require multivariate modeling of multi-method personality and mental disorder indicators in large, representative samples.

If comorbidity is treated as meaningful covariance, then the common mental disorder domain can be viewed as hierarchical — like the personality trait domain. From this perspective, the DSM focuses on an extensive number of mid-level dichotomized entities. Is this the best focus for our research efforts, or would a higher-order, more continuous focus be more useful? We feel a reasonable argument can be made for focusing on a higher level of the mental disorder hierarchy — on broad I and E factors, and their links to personality. This change of focus resolves various conundrums that result from the comorbidity concept. For example, recent clinical guidelines emphasize the role of the serotonin reuptake inhibitors (SSRIs) in treating a host of “separate and often comorbid” internalizing conditions (Dunner, 1998). Given evidence that SSRI administration reduces negative affect levels in persons without a history of mental

disorder (Knutson et al., 1998), SSRIs may be better viewed as treatments for the core, shared temperamental component of the internalizing disorders. Further research on core I and E dimensions and their links to personality may assist in parsimoniously accounting for these and other phenomena that do not readily fit the “comorbidity” concept. Such research, in turn, could form the basis for a psychometrically-sophisticated, empirically-based approach to assessment and taxonomy of mental disorder in adults (Achenbach, 1995).

Acknowledgements

The Minnesota Twin Family Study is supported in part by USPHS grants AA00175, AA09367, and DA05147. We thank Auke Tellegen, Kristen Kling, and the anonymous reviewers for their helpful comments.

References

- Achenbach, T. M. (1995). Empirically based assessment and taxonomy: applications to clinical research. *Psychological Assessment*, 7, 261–274.
- Achenbach, T. M., & Edelbrock, C. S. (1978). The classification of child psychopathology: a review and analysis of empirical efforts. *Psychological Bulletin*, 85, 1275–1301.
- Achenbach, T. M., & Edelbrock, C. S. (1984). Psychopathology of childhood. *Annual Review of Psychology*, 35, 227–256.
- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52, 317–332.
- Akiskal, H. S., Hirschfeld, R. M. A., & Yerevanian, B. I. (1983). The relationship of personality to affective disorders. *Archives of General Psychiatry*, 40, 801–810.
- American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: APA.
- American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev. ed.). Washington, DC: APA.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: APA.
- Bouchard, T. J. (1994). Genes, environment and personality. *Science*, 264, 1700–1701.
- Church, A. T., & Burke, P. J. (1994). Exploratory and confirmatory tests of the big five and Tellegen’s three- and four-dimensional models. *Journal of Personality and Social Psychology*, 66, 93–114.
- Clark, L. A., Watson, D., & Mineka, S. (1994). Temperament, personality, and the mood and anxiety disorders. *Journal of Abnormal Psychology*, 103, 103–116.
- Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: challenges to the current system and future directions. *Annual Review of Psychology*, 46, 121–153.
- Cloninger, C. R. (1999). *Personality and psychopathology*. Washington, DC: American Psychiatric Press.
- Costa, P. T. Jr, & McCrae, R. R. (1997). Longitudinal stability of adult personality. In R. Hogan, J. Johnson, & S. Briggs, *Handbook of personality psychology* (pp. 269–290). San Diego, CA: Academic Press.
- Costa, P. T. Jr, & Widiger, T. A. (1994). *Personality disorders and the five-factor model of personality*. Washington, DC: American Psychological Association.
- Dunner, D. L. (1998). The issue of comorbidity in the treatment of panic. *International Clinical Psychopharmacology*, 13(supplement 4), S19–S24.
- Eysenck, H. J. (1992). Four ways five factors are not basic. *Personality and Individual Differences*, 13, 667–673.
- Eysenck, H. J. (1994). Normality-abnormality and the three-factor model of personality. In S. Strack, & M. Lorr, *Differentiating normal and abnormal personality* (pp. 3–25). New York: Springer.

- Graham, J. R. (2000). *MMPI-2: assessing personality and psychopathology* (3rd ed.). New York: Oxford University Press.
- Holdcraft, L. C., Iacono, W. G., & McGue, M. K. (1998). Antisocial Personality Disorder and depression in relation to alcoholism: a community-based sample. *Journal of Studies on Alcohol*, 59, 222–226.
- Iacono, W. G., Lykken, D. T., & McGue, M. (1996). Psychophysiological prediction of substance abuse. In H. W. Gordon, & M. D. Glantz, *Individual differences in the biobehavioral etiology of drug abuse* (pp. 129–160). Rockville, MD: National Institute on Drug Abuse.
- Jöreskog, K., & Sörbom, D. (1996a). *LISREL 8: user's reference guide*. Chicago, IL: Scientific Software International.
- Jöreskog, K., & Sörbom, D. (1996b). *PRELIS 2: user's reference guide*. Chicago, IL: Scientific Software International.
- Kessler, R. C., Crum, R. M., Warner, L. A., Nelson, C. B., Schulenberg, J., & Anthony, J. C. (1997). Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. *Archives of General Psychiatry*, 54, 313–321.
- Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H. U., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Survey. *Archives of General Psychiatry*, 51, 8–19.
- Klein, M. H., Wonderlich, S., & Shea, M. T. (1993). Models of relationships between personality and depression: toward a framework for theory and research. In M. H. Klein, S. Wonderlich, & M. T. Shea, *Personality and depression: a current view* (pp. 1–54). New York: Guilford Press.
- Knutson, B., Wolkowitz, O. M., Cole, S. W., Chan, T., Moore, E. A., Johnson, R. C., Terpstra, J., Turner, R. A., & Reus, V. I. (1998). Selective alteration of personality and social behavior by serotonergic intervention. *American Journal of Psychiatry*, 155, 373–379.
- Krueger, R. F. (1999a). The structure of common mental disorders. *Archives of General Psychiatry*, 56, 921–926.
- Krueger, R. F. (1999b). Personality traits in late adolescence predict mental disorders in early adulthood: a prospective-epidemiological study. *Journal of Personality*, 67, 39–65.
- Krueger, R. F., Caspi, A., Moffitt, T. E., & Silva, P. A. (1998). The structure and stability of common mental disorders (DSM-III-R): a longitudinal-epidemiological study. *Journal of Abnormal Psychology*, 107, 216–227.
- Krueger, R. F., Caspi, A., Moffitt, T. E., Silva, P. A., & McGee, R. (1996). Personality traits are differentially linked to mental disorders: a multitrait-multidiagnosis study of an adolescent birth cohort. *Journal of Abnormal Psychology*, 105, 299–312.
- Livesley, W. J., Jang, K. L., & Vernon, P. A. (1998). Phenotypic and genetic structure of traits delineating personality disorder. *Archives of General Psychiatry*, 55, 941–948.
- Loehlin, J. C. (1992). *Genes and environment in personality development*. Newbury Park, CA: Sage.
- McCrae, R. R., Zonderman, A. B., Costa, P. T. Jr, Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the revised NEO personality inventory: confirmatory factor analysis versus procrustes rotation. *Journal of Personality and Social Psychology*, 70, 552–566.
- Morgenstern, J., Langenbucher, J., Labouvie, E., & Miller, K. J. (1997). The comorbidity of alcoholism and personality disorders in a clinical population: prevalence rates and relation to alcohol typology variables. *Journal of Abnormal Psychology*, 106, 74–84.
- Plomin, R., & Caspi, A. (1998). DNA and personality. *European Journal of Personality*, 12, 387–407.
- Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25, 111–163.
- Robins, L. M., Babor, T. F., & Cottler, L. B. (1987). *Composite international diagnostic interview: expanded substance abuse module*. St. Louis, MO: Washington University Department of Psychiatry.
- Robins, L. N., Helzer, J. E., Cottler, L., & Goldring, E. (1989). *Diagnostic interview schedule, version III-R*. St. Louis, MO: Washington University Department of Psychiatry.
- Rounsaville, B. J., Kranzler, H. R., Ball, S., Tennen, H., Poling, J., & Triffleman, E. (1998). Personality disorders in substance abusers: relation to substance use. *Journal of Nervous and Mental Disease*, 186, 87–95.
- Sher, K. J., & Trull, T. J. (1994). Personality and disinhibitory psychopathology: alcoholism and antisocial personality disorder. *Journal of Abnormal Psychology*, 103, 92–102.
- Spitzer, R. L., Williams, J. B. W., & Gibbon, M. (1987). *Structured clinical interview for DSM-III-R*. New York: Biometrics Research Department, New York State Psychiatric Institute.
- Strack, S., & Lorr, M. (1994). *Differentiating normal and abnormal personality*. New York: Springer.

- Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma, & J. D. Maser, *Anxiety and the anxiety disorders* (pp. 681–706). Hillsdale, NJ: Erlbaum.
- Tellegen, A. (2000). *Manual for the Multidimensional Personality Questionnaire*. University of Minnesota Press: Minneapolis, MN, in press.
- Tellegen, A., & Waller, N. G. (2000). Exploring personality through test construction: development of the Multidimensional Personality Questionnaire. In S. R. Briggs and J. M. Cheek (Eds.), *Personality measures: development and evaluation* (Vol. 1). JAI Press: Greenwich, CT, in press.
- Trull, T. J., & Sher, K. J. (1994). Relationship between the five-factor model of personality and Axis I disorders in a nonclinical sample. *Journal of Abnormal Psychology*, 103, 350–360.
- Waller, N. G. (1999). Searching for structure in the MMPI. In S. E. Embretson, & S. L. Hershberger, *The new rules of measurement: what every psychologist and educator should know* (pp. 185–217). Mahwah, NJ: Erlbaum.
- Watson, D., & Clark, L. A. (1994). Introduction to the special issue on personality and psychopathology. *Journal of Abnormal Psychology*, 103, 3–5.
- Watson, D., Clark, L. A., & Harkness, A. R. (1994). Structures of personality and their relevance to psychopathology. *Journal of Abnormal Psychology*, 103, 18–31.
- Weich, S., Sloggett, A., & Lewis, G. (1998). Social roles and gender difference in the prevalence of common mental disorders. *British Journal of Psychiatry*, 173, 489–493.
- Wittchen, H. U., Kessler, R. C., Zhao, S., & Abelson, J. (1995). Reliability and clinical validity of UM-CIDI DSM-III-R generalized anxiety disorder. *Journal of Psychiatric Research*, 29, 95–110.