Psychological Medicine (2008), **38**, 51–61. © 2007 Cambridge University Press
doi:10.1017/S0033291707001572 Printed in the United Kingdom

ORIGINAL ARTICLE

The role of internalizing and externalizing liability factors in accounting for gender differences in the prevalence of common psychopathological syndromes

Mark D. Kramer*, Robert F. Krueger* and Brian M. Hicks

Department of Psychology, University of Minnesota, Twin Cities Campus, Minneapolis, MN, USA

Background. We hypothesized that gender differences in average levels on the internalizing and externalizing factors that account for co-morbidity among common psychopathological syndromes in both men and women account for gender differences in the prevalence of specific syndromes.

Method. The latent structure of 11 syndromes was examined in a middle-aged (mean age = 52.66 years, S.D. = 5.82) sample of 2992 (37% men) members of the community-based Minnesota Twin Registry (MTR) assessed using 10 scales of the Psychiatric Diagnostic Screening Questionnaire (PDSQ) and an adult antisocial behavior scale. Confirmatory factorial invariance models were applied to a best-fitting, internalizing–externalizing model.

Results. A 'strong gender invariance model' fit best, indicating that gender differences in the means of individual syndromes were well accounted for by gender differences in mean levels of internalizing and externalizing. Women exhibited higher mean levels of internalizing ($d = 0.23$) and lower mean levels of externalizing ($d = -0.52$) than men.

Conclusions. These findings suggest that risk factors for common mental disorders exhibiting gender differences may influence prevalence at the latent factor level. Future research may benefit from focusing on both the latent factor and individual syndrome levels in explaining gender differences in psychopathology.

Received 19 July 2006; Revised 10 July 2007; Accepted 17 July 2007; First published online 25 September 2007

Key words: Antisocial, anxiety, co-morbidity, depression, substance.

Introduction

Gender differences in the prevalence of common mental disorders have long been of interest in the mental health field because, in part, identification of the causes of such discrepancies may help to elucidate the etiology of various psychopathological conditions. Epidemiological and community-based surveys consistently report higher rates of mood and anxiety disorders in women than in men, and higher rates of substance use and antisocial behavior disorders in men than in women (Robins & Regier, 1991; Kessler *et al*. 1994; Bijl *et al*. 1998; Jenkins *et al*. 2003; Grant *et al*. 2004*a*; Klose & Jacobi, 2004; Compton *et al*. 2005). Structural models of co-morbidity among common mental disorders have revealed that the observed co-morbidity structure (e.g. Krueger *et al*. 1998; Krueger, 1999; Kendler *et al*. 2003), as well as the genetic and environmental architecture of co-morbidity (Krueger *et al*. 2002; Kendler *et al*. 2003), is consistent across gender. For both genders, mood and anxiety disorders can be modeled as elements within an etiologically coherent internalizing spectrum, and substance use and antisocial behavior disorders can be modeled as elements within an etiologically coherent externalizing spectrum.

A reasonable hypothesis that can be derived by integrating these lines of research is that gender differences in average levels on the internalizing and externalizing factors that account for the co-morbidity among common mental disorders in *both* genders might account in part for gender *differences* in the prevalence of specific disorders. Evidence in support of this hypothesis would provide a parsimonious model to frame future research on gender differences in common forms of psychopathology. In addition to research focusing on specific syndromes, such a model would suggest investigation of the sources of gender differences in both the average levels of the general internalizing and externalizing liabilities and the unique variances of individual syndromes within the spectra. We applied formal factorial invariance procedures (Meredith, 1993) within the

* Address for correspondence: M. D. Kramer or R. F. Krueger, Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455, USA.
(Email: krame016@umn.edu *or* krueg038@umn.edu)

internalizing–externalizing framework to tie together the prevalence and co-morbidity literatures on gender differences.

Method

Participants

Participants were members of the Minnesota Twin Registry (MTR), a registry of twin pairs born in Minnesota identified through birth records provided by the Minnesota State Health Department. For a comprehensive description of the establishment of the MTR, see Lykken *et al*. (1990). Twin pairs who were prospective participants in this study were selected from the MTR on the basis of the following criteria: both were living in the USA, zygosity information had been obtained by the method described in Lykken *et al*. (1990), and one or both of the twins of each pair had previously participated in postal data collection. The current study focused on the 3025 participants from whom we received completed mental health questionnaires by mail; the overall response rate to this mailing was 50%. Participants missing one page (31 items) or more of the questionnaire were not included in the analyses ($n = 33$). The sample included in the analyses consisted of 2992 participants, 1112 men and 1880 women, aged 38 to 76 years (mean = 52.66, S.D. = 5.82), and most were Caucasian (94%).

To examine non-response bias, we compared participants to non-participants on the 11 primary trait scales of the Multidimensional Personality Questionnaire (MPQ; Tellegen, in press), a self-report personality inventory with scale scores that evince a systematic pattern of correlations with internalizing and externalizing syndromes (Krueger *et al*. 1996). Most of the twins targeted for the current project had completed the MPQ at some point prior to the mailing, including the Psychiatric Diagnostic Screening Questionnaire (PDSQ; Zimmerman & Mattia, 2001*a*) (91%). Differences on the 11 personality scales were modest ($\text{mean}_{|d|} = 0.09$, S.D. = 0.06, range −0.20 to 0.15), indicating that participating and non-participating members of the MTR are likely to be comparable in levels of self-reported psychopathological symptomatology.

Measures

The PDSQ is a self-report questionnaire of dichotomously scored (yes–no) items designed to assess the symptoms of common Axis I disorders of the DSM-IV-TR (APA, 2000). Ten scales of the PDSQ were administered to the twins: generalized anxiety disorder, major depressive disorder, panic disorder, social phobia, hypochondriasis, obsessive–compulsive disorder (OCD), bulimia/binge-eating disorder, agoraphobia, alcohol abuse/dependence, and drug abuse/dependence (Zimmerman & Mattia, 2001*a*). In the original PDSQ, two time-frames are assessed using these scales, with the major depressive disorder, panic disorder, OCD, and bulimia/binge-eating disorder referring to the past 2 weeks, and the generalized anxiety disorder, social phobia, hypochondriasis, agoraphobia, alcohol abuse/dependence and drug abuse/dependence scales referring to the past 6 months (Zimmerman & Mattia, 2001*b*). We modified the instructions to assess symptoms experienced during the past month for each scale in order to standardize the duration assessed across scales and to better ensure that co-occurrence of syndromes was assessed rather than fluctuation of symptomatology from one domain to another over a 6-month period.

The 10 scales used in the current study have strong psychometric properties, demonstrating good to excellent internal consistency, test–retest reliability, convergent and discriminant validity (Zimmerman & Mattia, 2001*a*). In reference to diagnoses made using the Structured Clinical Interview for DSM-IV (SCID; First *et al*. 1997), the scale sensitivities (mean = 89.10, S.D. = 2.51) at recommended cut-off scores for outpatients resulted in mean scale specificities of 70.50 (S.D. = 12.28, range 50–89) (Zimmerman & Mattia, 2001*b*). In addition, 16 dichotomous items were included to assess DSM-IV criteria for antisocial personality disorder, with the exception of the conduct disorder criterion. These adult antisocial behavior items were based on an interview assessment of antisocial personality disorder used by the Minnesota Twin Family Study (Holdcraft *et al*. 1998).

Analyses

Confirmatory factor analysis (CFA) using MPlus version 3.0 (Muthén & Muthén, 2004) to estimate maximum likelihood model parameters from the PDSQ score data was conducted in the full sample. Each scale score was treated as a continuous variable and, to take into account both skewness of the symptom count scores and dependence of twin observations, maximum likelihood parameter estimates with standard errors and χ^2 test statistic robust to non-normality and non-independent observations were used (MLR; Muthén & Muthén, 2004). A one-factor model in which all of the disorders were indicators of a single, underlying propensity to experience common mental disorders was evaluated. Tested also was a two-factor model which posited that generalized anxiety disorder, major depressive disorder, panic disorder, social phobia, hypochondriasis, OCD, bulimia/binge-eating disorder, and agoraphobia reflect internalizing

https://doi.org/10.1017/S0033291707001572 Published online by Cambridge University Press

problems, and alcohol abuse/dependence, adult antisocial behavior, and drug abuse/dependence reflect externalizing problems.

In addition to the χ^2 test statistic, which indexes overall fit of the model, fit statistics for CFA included the root mean square error of approximation (RMSEA), the standardized root mean residual (SRMR), and the Bayesian information criterion (BIC). RMSEA indexes the degree of discrepancy between the estimated population covariance matrix of observed variables and the covariance matrix reproduced as a function of parameters of the confirmatory model, per degree of freedom. SRMR estimates the discrepancy between statistics predicted by the model and those derived from the sample. Values of both indexes range from 0 to 1, with lower values indicating better model fit. BIC values cover a wider range than RMSEA and SRMR. BIC balances both fit and parsimony, and is a function of the model χ^2 and its degrees of freedom [$\chi^2 - \mathrm{df}(\ln N)$]. Because increasing the number of parameters in a model tends to reduce the χ^2 fit value, models with excessive numbers of parameters can appear to fit better than simpler, more parsimonious models. BIC produces lower values for models superior in both fit and parsimony. In evaluating the fit of competing models, we judged the best-fitting model to be the one that produced the largest negative BIC values while retaining acceptable absolute fit (i.e. low RMSEA and SRMR values).

Modeling the data from each gender individually does not take into account all of the data from both genders simultaneously. To model the prevalence and co-morbidity among the 11 psychopathological syndromes in both men and women, factorial invariance models were fit to the means, variances and covariances of the observed variables in each gender. Confirmatory factorial invariance models involve the estimation of factor loadings, factor covariances, factor variances, residual variances, intercepts and factor means. The last two parameters – not estimated in CFA using single groups – can be estimated because factorial invariance models take into account the additional information provided by the means of the observed variables in each group being modeled.

Increasingly stringent levels of factorial invariance can be distinguished through imposition of constraints on parameter estimates across genders (Meredith, 1993). Configural invariance exists if the same factor structure can be imposed across genders. Metric invariance assumes configural invariance and includes the additional constraint that the loadings must be equal across groups. Strong invariance assumes configural and metric invariance and adds the constraint that the intercepts (terms added to the structural equations linking observed and latent variables that allow latent variable means to account for observed means) are constant across genders. Strict invariance assumes configural, metric and strong invariance, in addition to the constraint that residual variances of the observed variables must be equal for men and women. In the strong and strict invariance models, differences in means of the latent factors account for differences in the mean levels of each of the respective syndromes, but the strict invariance model also specifies that variances in the syndromes unaccounted for by the latent factors are the same in men and women.

Results

Gender differences in mean levels of individual syndromes

Descriptive statistics for each scale for men and women and internal consistency for each of the scales are provided in Table 1†. PDSQ scales with means of 0.30 in the whole sample or less were those with recommended cut-off scores of 1 in out-patient samples (with the exception of agoraphobia, which had a cut-off score of 4; Zimmerman & Mattia, 2001*b*). This suggests that symptoms tapped by items on these scales were rare in our community-dwelling sample, with endorsement of one item indicating a high likelihood of the presence of the disorder as defined in DSM-IV. In addition, PDSQ scales with α values less than 0.70 also had recommended cut-off scores of 1. Given the apparent severity of these items, the low variance in past-month symptomatology in a community-dwelling sample of middle-aged twins is not surprising, and as a result internal consistency in this sample could be considered lower bound estimates. Independent-sample t tests indicated that the women's mean scale score for each internalizing syndrome, with the exceptions of OCD and hypochondriasis, was significantly higher than that of the men's ($p < 0.05$). For externalizing syndromes, men had significantly higher mean scores on the alcohol abuse/dependence and adult antisocial behavior scales ($p < 0.001$) but did not differ significantly from women on the drug abuse/dependence scale.

Confirmatory factor analysis (*CFA*)

The two-factor model in the CFA fit better than the one-factor model. The two-factor model exhibited a

† One item from the AAB scale was dropped prior to computing scale scores used in the analyses because it was essentially uncorrelated with the sum of the other items ($r = -0.035$). Internal consistency reliability for the scales was generally acceptable, with the exception of AAB (see Table 1). We nevertheless retained AAB in the analyses because it shared non-trivial variance with other externalizing scales despite its lower α.

https://doi.org/10.1017/S0033291707001572 Published online by Cambridge University Press

Table 1. *Scale descriptive statistics by gender (1112 men, 1880 women), internal consistency, and independent-sample t tests*

Scale	Items	Men Mean[a] (S.D.)	Women Mean[a] (S.D.)	α	t	df[b]	p
Generalized anxiety disorder	10	0.88 (1.92)	1.36 (2.31)	0.87	−6.06	2672.30	<0.001
Major depressive disorder	21	1.58 (2.43)	2.01 (2.65)	0.82	−4.53	2495.35	<0.001
Panic disorder	8	0.24 (0.88)	0.39 (1.14)	0.81	−3.84	2780.99	<0.001
Social phobia	15	1.00 (2.22)	1.18 (2.53)	0.89	−2.02	2575.72	0.044
Hypochondriasis	5	0.16 (0.60)	0.18 (0.62)	0.69	−0.95	2990.00	0.344
Obsessive–compulsive disorder	7	0.11 (0.48)	0.11 (0.50)	0.64	−0.07	2990.00	0.946
Bulimia/binge-eating disorder	10	0.42 (1.17)	0.85 (1.81)	0.85	−7.95	2965.14	<0.001
Agoraphobia	11	0.20 (0.87)	0.30 (1.09)	0.83	−2.77	2738.40	0.006
Alcohol abuse/dependence	6	0.50 (1.11)	0.19 (0.69)	0.77	8.44	1620.70	<0.001
Adult antisocial behavior	15	0.28 (0.65)	0.13 (0.49)	0.46	6.44	1867.65	<0.001
Drug abuse/dependence	6	0.04 (0.28)	0.02 (0.23)	0.67	1.17	1946.05	0.243

α, Cronbach's α; df, degrees of freedom.

[a] Mean for dichotomous item scales equals the average number of items endorsed.

[b] Degrees of freedom are fractional when homogeneity of variance was not assumed according to Levene's test.

lower χ^2 [$\chi^2 = 255.51$ (43)], RMSEA (0.041), SRMR (0.035) and BIC (−88.65) than the one-factor model [$\chi^2 = 300.85$ (44), RMSEA = 0.044, SRMR = 0.042, BIC = −51.31]. More negative values of BIC indicate better fit while taking into account parsimony of the model. In addition to a lower BIC value than the one-factor model, the two-factor model demonstrated excellent absolute fit, as indicated by the lower RMSEA and SRMR values.

Factorial invariance models

Factorial variance and invariance models differ from conventional CFA models in that all data from both groups are taken into account simultaneously in the estimation of model parameters and evaluation of model fit. The configural, metric, strong, and strict invariance models imposed increasingly stringent constraints on the two-factor model in order to assess measurement invariance across sex.

The configural invariance model specified the factor structure to be invariant across gender, with internalizing disorders loading on one factor and externalizing disorders loading on another, for both men and women. If this model fit best, it would indicate that the structure of co-morbidity among these disorders was the same for men and women, with the same underlying factors (internalizing and externalizing) accounting for the shared variance among the same groups of disorders.

The metric invariance model assumed configural invariance (equivalent factor structure across gender) and thus allowed for the estimation of model fit when factor loadings were held constant across gender. The factor loadings are the coefficients in the structural equation models linking the observed and latent variables. If the metric invariance model fit better than the configural invariance model, it would imply that the structure of co-morbidity was the same across gender, and that the magnitude of relations linking the syndromes with the latent factors of internalizing and externalizing was equivalent for men and women.

The strong invariance model assumed that the factor structure was invariant across gender (configural invariance), that the magnitude of relationships between the syndromes and latent factors were the same across gender (metric invariance), and added the constraint that intercepts of the structural equation models linking latent factors and the observed syndromes were constant across gender. If the strong invariance model fit better than the metric invariance model, it would also indicate that gender differences in average levels of each disorder could be accounted for by differences between men and women on the mean levels of the latent factors, internalizing and externalizing.

The strict invariance model assumed that the factor structure was invariant across gender (configural invariance), that the magnitude of relationships between the syndromes and latent factors were the same across gender (metric invariance), that the intercepts of the structural equation models linking latent factors and the observed syndromes were the same for men and women (strong invariance), and added the constraint that the residual variances in the syndromes were constant across gender. If the strict invariance model fit better than the strong invariance model, it would also imply that there were no gender differences in variance of the syndromes that were unaccounted for by the latent factors.

https://doi.org/10.1017/S0033291707001572 Published online by Cambridge University Press

Table 2. *Fit statistics for factorial invariance models across gender*

	χ^2	df	RMSEA	SRMR	BIC
Configural invariance	305.00	86	0.041	0.040	−383.32
Metric invariance	307.20	95	0.039	0.047	−453.15
Strong invariance	352.05	104	0.040	0.050	−480.34
Strict invariance	434.73	115	0.043	0.093	−485.70

χ^2, adjusted χ^2 fit statistic with robust standard errors; df, degrees of freedom; RMSEA, root mean square error of approximation; SRMR, standardized root mean residual; BIC, Bayesian information criterion.

Best-fitting model

Taking into account the pattern across fit statistics, the two-factor strong gender invariance model fit best (Table 2). The strong gender invariance model had a lower, more negative BIC value than the configural and metric invariance models while retaining acceptable fit according to RMSEA and SRMR. The strict gender invariance model exhibited a slightly lower BIC value than the strong invariance model, but RMSEA and SRMR statistics indicated a decrement in absolute fit. Thus, the two-factor strong gender invariance model was selected as the best-fitting model.

The best-fitting strong gender invariance model specifies that the factor structure, factor loadings and intercepts were equivalent for men and women, indicating that the differences in means of the observed disorders could be explained by between-gender differences in the means of the factors. The structure of co-morbidity among common mental disorders was the same for men and women (see Fig. 1). The women's mean score on the latent internalizing factor was greater than that of the men ($d = 0.23$), and the women's internalizing scores were 1.43 times more variable than the men's. The women's mean score on the latent externalizing factor was less than that of the men ($d = -0.52$), and the women's externalizing scores showed approximately one-third (0.35) of the variance of the men's externalizing scores.

Estimates presented in Fig. 1 are unstandardized and identified in terms of the loadings to illustrate their equivalence across gender. The unstandardized loadings for each syndrome differed in magnitude as a function of both the observed variance accounted for by the latent factors and the differences in observed variance in each of the syndromes, and are therefore not directly comparable. To further evaluate how well the strong invariance model accounted for observed gender differences in the syndromes, we correlated the observed syndrome means and variances with model-estimated means (i.e. factor mean + intercept) and variances [i.e. residual variance + (loading × factor variance × loading)]. The rank order of strong invariance model estimates closely approximated the observed means in men ($r_{\text{spearman rank order}} = 0.98$) and women ($r_s = 0.95$), and perfectly approximated the observed variances in men ($r_s = 1.00$) and women ($r_s = 1.00$). These findings provide further support for the parsimony of the strong invariance model relative to the metric invariance model, and for the latent factors' ability to account for the observed mean differences in this sample. The degradation in fit of the strict relative to the strong invariance model indicates that there are important gender differences in variances of the individual disorders that are not accounted for by the latent factors.

Discussion

We examined the prevalence and co-morbidity of 11 common psychopathological syndromes in middle-aged male and female twins by testing models of factorial invariance across gender. The best-fitting two-factor strong invariance model differentiated internalizing (i.e. generalized anxiety disorder, major depressive disorder, panic disorder, social phobia, hypochondriasis, OCD, bulimia/binge-eating disorder, and agoraphobia) from externalizing (i.e. alcohol abuse/dependence, adult antisocial behavior, and drug abuse/dependence) syndromes in men and women indicated that the magnitude of the relationships between the latent factors and the observed syndromes (i.e. loadings) were equivalent across gender, and revealed that the differences in average levels of the syndromes between men and women were well accounted for by differences in means of the latent factors of internalizing and externalizing. Compared to men, women exhibited greater mean levels and variability on the internalizing factor, and lower mean levels and less variability on the externalizing factor. These findings simultaneously and parsimoniously account for both the similarities in patterns of co-morbidity and differences in the mean levels of a broad range of internalizing and externalizing syndromes among men and women.

Certain limitations of the current study should be taken into account when interpreting the results. The psychopathological symptom data were acquired using a self-report inventory, the PDSQ. While one of the goals of this research was to replicate the internalizing–externalizing model using a method of assessment other than structured interview, the structured interview remains the typical approach to assessing psychopathological constructs in community-dwelling samples. Nevertheless, given the acceptable to strong sensitivity and specificity of these

https://doi.org/10.1017/S0033291707001572 Published online by Cambridge University Press

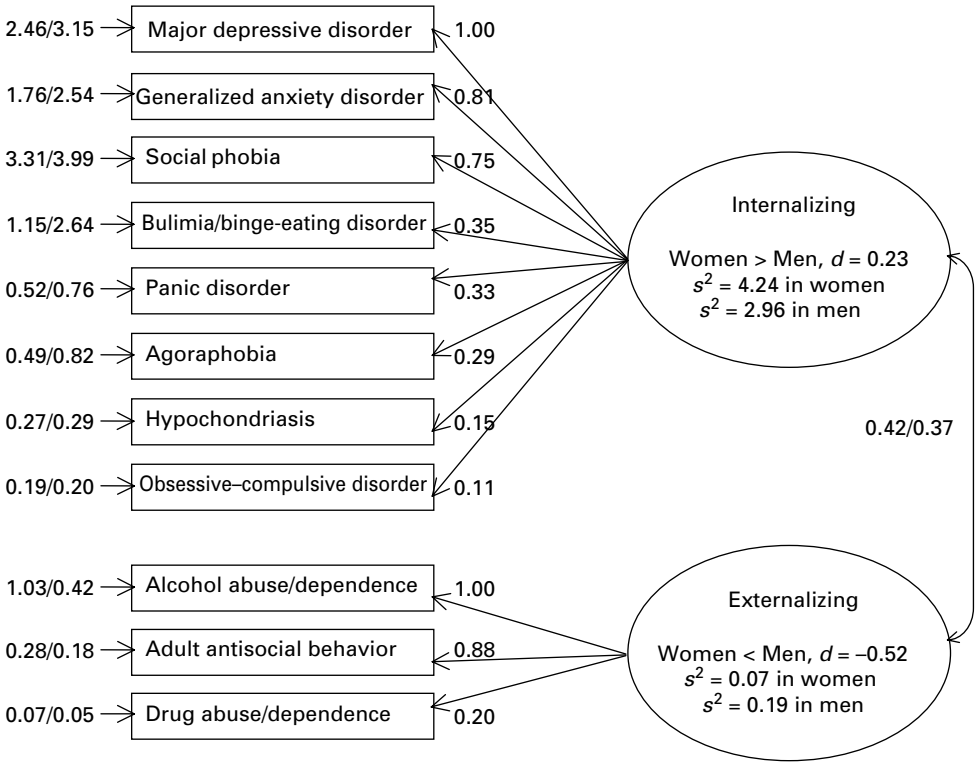

Fig. 1. Best-fitting strong gender invariance model with factor structure, factor loadings, and intercepts constrained to be equal in men and women. Loadings are unstandardized estimates. Residual variances and factor covariances are presented as men/women.

scales with clinical out-patients, the PDSQ has been demonstrated to be a valid screen for DSM-defined mental disorders assessed using structured interviews (Zimmerman & Mattia, 2001*a*, *b*). We have also extended the finding that mean-level gender differences in externalizing account for gender differences in prevalence of a broader array of externalizing syndromes using symptom-count data assessed by structured interview (Hicks *et al*. 2007). Extension of the present findings using interview-derived DSM diagnoses will nevertheless be important. For example, future work could investigate whether gender differences in mean levels of internalizing and externalizing can account for gender differences in the prevalence of dichotomous diagnoses. This approach would complement evidence presented herein regarding mean numbers of symptoms within syndromes.

In addition, the limited variance in a number of syndromes contributed to the relative sizes of the loadings of the syndromes in the best-fitting model. At the same time, the different proportions of variance in each syndrome accounted for by the internalizing and externalizing factors emphasizes the higher-order structure of the framework. That is, each syndrome consists of both common and unique variance and the syndromes differ in terms of the extent to which they reflect common vs. unique variance. For example, in men, 55% and 16% of the variance in major depressive disorder and OCD, respectively, was accounted for by internalizing. That is, etiological mechanisms not associated with internalizing may account for more of the variance in OCD than major depressive disorder. In addition, mechanisms that account for the gender difference in internalizing will be likely to explain more of the variance in major depressive disorder than OCD.

In conducting gender differences research on specific syndromes, both common and unique variance in the syndrome should be taken into account. Certain risk factors exhibiting gender differences may account for a proportion of the common variance in the syndrome (e.g. the gender difference in internalizing), the unique variance associated with the syndrome, or differential proportions of the unique variance in the syndrome in men and women. The caution of taking into account both general and specific risk factors in gender differences research, and in researching risk factors other than gender, is not a new idea; however, the high rates of co-morbidity among disorders suggest that rarely will risk factors be

https://doi.org/10.1017/S0033291707001572 Published online by Cambridge University Press

identified that can be dichotomously defined as either general or specific. The internalizing–externalizing model provides an empirical framework in which to estimate a risk factor's general and specific contribution to variance in a particular syndrome. Nevertheless, while taking into account the fact that multiple outcomes is a benefit of using the internalizing–externalizing model, the current findings are limited in the use of a single predictor (i.e. gender). The association between gender and internalizing and externalizing syndromes may differ when additional predictors are taken into account. Investigation of the ultimate plausibility of the strong gender invariance model would be bolstered by replication and demonstration of the utility of the approach to gender differences research.

Finally, the current findings are based on modeling the point prevalence (i.e. past-month symptomatology) of the syndromes in a middle-aged sample and thus may not be applicable to gender differences in onset of psychopathology. The associations of the latent factors with the course of these syndromes in men and women may differ from the current findings as well. For example, the cross-sectional assessment did not allow for modeling the possibility that exposure to specific syndromes may differentially predispose men and women to other syndromes within or across spectra. Despite these limitations, our findings provide preliminary evidence that suggests that investigations of the causes of gender differences in the prevalence of specific disorders may benefit from elucidating the causes of gender differences in average levels of the latent factors that account for the patterns of co-morbidity among internalizing and externalizing syndromes, respectively.

Using the internalizing–externalizing model in gender differences research

One goal of investigations of sex differences in the prevalence rates of psychopathological syndromes is to reveal etiological risk factors. The internalizing–externalizing model accounts for two key observations in the epidemiology of common mental disorders: the patterns of co-morbidity and the patterns of gender differences in mean levels of internalizing and externalizing syndromes. The degradation in fit of the strict, relative to the strong, invariance model indicates that there are important gender differences in variances of the individual syndromes that are unaccounted for by the latent factors as well. Use of the internalizing–externalizing model in gender differences research would benefit from inclusion of a broad array of risk factors to investigate their association with psychopathology in general, the latent internalizing and externalizing factors, and residual variance in the individual syndromes unaccounted for by internalizing and externalizing. Our results indicate that, of the many risk factors posited to explain sex differences in diverse disorders, risk factors that contribute to gender differences in internalizing and externalizing may best account for gender differences in the mean levels of specific disorders within these spectra. While a review of general and specific risk factors accounting for gender differences in variance (but not to idiosyncratic mean-level differences) in the syndromes is beyond the scope of this discussion, we turn to those risk factors likely to be associated with gender differences in internalizing and externalizing.

As there is currently little evidence that risk factors for internalizing and externalizing disorders are specific to men or women, the sex difference in internalizing and externalizing may arise from a sex difference in the differential *exposure* to risk factors. Risk factors that are likely to account for gender differences at the factor level will probably: (1) exhibit a reliable mean-level sex difference (i.e. there is differential exposure to the risk factor) and (2) be associated with multiple disorders within the internalizing or externalizing spectrum. In turn, identification of such risk factors will provide greater understanding of the etiology of internalizing and externalizing, and the internalizing–externalizing model may provide a useful framework to organize risk factors for common mental disorders. We now discuss putative risk factors that may contribute to gender differences in mean levels of externalizing and internalizing, respectively.

Gender differences in risk factors for externalizing

Men exhibit a higher mean level of a number of risk factors that are associated with multiple disorders within the externalizing spectrum. Neuro-cognitive deficits, early-emerging under-controlled temperament, hyperactivity, deviant peer relationships, and personality traits associated with heightened negative affectivity (especially aggression) and weak behavioral constraint are associated with sex differences in adolescent antisocial behavior (Moffitt *et al*. 2001). Similar risk factors have been implicated in men's greater rates of alcohol use disorders, including traits associated with the masculine gender role (particularly aggressiveness), impulsivity, behavioral undercontrol, sensation seeking, and antisociality (Nolen-Hoeksema, 2004). While sex differences in drug use disorders have received less attention in the literature, a number of these risk factors have also been associated with drug use disorders, including

https://doi.org/10.1017/S0033291707001572 Published online by Cambridge University Press

peer influence (Hawkins *et al*. 1992; Guo *et al*. 2002), traits related to behavioral under-control, sensation seeking, and impulsivity (Wills *et al*. 1994; Iacono *et al*. 1999; Sher *et al*. 2000; Agrawal *et al*. 2004; de Wit & Richards, 2004; Lynam & Miller, 2004), neuro-cognitive deficits (Giancola & Tarter, 1999; Tarter *et al*. 2003), antisocial personality disorder (Grant *et al*. 2004*b*) and aggression (Brook *et al*. 1996; Reinherz *et al*. 2000; Giancola & Parker, 2001).

Gender differences in risk factors for internalizing

Other risk factors that exhibit a gender difference have also been posited to account for women's greater rates of multiple disorders within the internalizing spectrum. The personality trait of neuroticism or stress reaction exhibits a reliable gender difference with greater mean levels observed for women (Costa *et al*. 2001) and has been consistently associated with depression and anxiety at both the phenotypic (Watson *et al*. 2005) and genetic level (Roberts & Kendler, 1999; Fanous *et al*. 2002). Neuroticism also partially accounts for the gender difference in depression (Goodwin & Gotlib, 2004), and accounts for much of the genetic overlap between depression and anxiety (Jardine *et al*. 1984; Kendler *et al*. 1993). Rumination is another risk factor that exhibits a reliable gender difference (Nolen-Hoeksema & Jackson, 2001) and is associated with anxiety as well as depression symptoms (Nolen-Hoeksema, 2000) and partially accounts for women's greater depressive symptoms (Nolen-Hoeksema, 1991; Nolen-Hoeksema *et al*. 1999). Ruminative response style may be conceptualized as a facet of a higher-order neuroticism/negative temperament construct (Watson *et al*. 2006).

The greater hypothalamic–pituitary–adrenal (HPA) axis dysregulation in response to stress exhibited by women has also been implicated in their greater rates of depression relative to men (Young & Korszun, 1999). Abnormalities in the HPA axis activity have also been proposed as playing a causal role in the pathogenesis of both anxiety and depressive disorders (Butler & Nemeroff, 1990), and are associated with co-morbid cases of depression and anxiety (Young *et al*. 2004). Reduced rates of serotonin synthesis (Nishizawa *et al*. 1997) have also been implicated in the greater rates of depression in women. The use of selective serotonin reuptake inhibitors and other antidepressants for the treatment of a variety of internalizing syndromes (Hudson & Pope, 1990; Hudson *et al*. 2003) suggests that serotonin abnormalities are not specific to depression, and that the action of these medications may be at the level of the general internalizing vulnerability (e.g. Knutson *et al*. 1998).

Summary and future directions

Based upon the available evidence, we reviewed a number of risk factors that exhibit reliable gender differences and are associated with multiple disorders within the internalizing and externalizing spectra. It appears that neuro-cognitive deficits, early-emerging under-controlled temperament, hyperactivity, deviant peer relationships and personality traits related to behavioral disinhibition and aggression are associated with and may account for much of the sex difference in externalizing. Neuroticism, rumination, HPA axis dysregulation and serotonin synthesis are associated with and may account for the sex difference in internalizing. Notably, these risk factors seem to form a coherent cluster of risk factors *within* spectra, but are distinctive *across* spectra. Psychosocial and psychobiological risk factors reflecting distress and stress reactivity may be more strongly associated with internalizing, while those reflecting relatively poorer prefrontal cortex development and executive functioning may be associated with externalizing.

Using our results that gender differences on internalizing and externalizing factors account for gender differences in specific syndromes, risk factors for internalizing and externalizing syndromes may be organized into distinct clusters, and the structure of the risk factors themselves may reflect similar organization to the internalizing–externalizing model. Future research may benefit from assessing multiple disorders and risk factors and examining the structure of each along with their inter-relationships. Empirical investigation of the organization of risk factors revealed in the literature may elucidate latent 'risk' dimensions representing the etiological processes underlying the internalizing and externalizing spectra. The internalizing and externalizing dimensions may not only account for the co-morbidity among the disorders but may also serve as 'linkage factors' between these risk dimensions and the syndromes. The risk dimensions may account for mean-level gender differences in internalizing and externalizing, which in turn account for the gender differences in specific syndromes. Busseri *et al*. (2007) provide an empirical framework in which to examine the relationship between risk predictors and latent linkage factors accounting for covariation among adolescent problem behaviors. Such an approach may be a focus of future research using the internalizing–externalizing model.

In addition, investigation of the gender invariance of the model should be assessed longitudinally during periods in which prevalence of syndromes in men and women increase (e.g. puberty) to assess whether mean-level changes in the latent factors account for the

gender differences in prevalence. The gender invariance of specific criteria of internalizing and externalizing syndromes may also be addressed in future work. Through systematic investigation of the risk factors contributing to gender differences in internalizing and externalizing, we may increase our understanding of the etiology of a broad array of common psychopathological syndromes.

Declaration of Interest

None.

References

Agrawal A, Jacobson KC, Prescott CA, Kendler KS (2004). A twin study of personality and illicit drug use and abuse/dependence. *Twin Research* **7**, 72–81.

APA (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, text revision American Psychiatric Association: Washington, DC.

Bijl RV, Ravelli A, van Zessen G (1998). Prevalence of psychiatric disorder in the general population: results of the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology* **33**, 587–595.

Brook JS, Whiteman M, Finch SJ, Cohen P (1996). Young adult drug use and delinquency: childhood antecedents and adolescent mediators. *Journal of the American Academy of Child and Adolescent Psychiatry* **35**, 1584–1592.

Busseri MA, Willoughby T, Chalmers H (2007). A rationale and method for examining reasons for linkages among adolescent risk behaviors. *Journal of Youth and Adolescence* **36**, 279–289.

Butler PD, Nemeroff CB (1990). Corticotropin-releasing factor as a possible cause of comorbidity in anxiety and depressive disorders. In *Comorbidity of Mood and Anxiety Disorders* (ed. J. D Maser and C. Cloninger), pp. 413–435. American Psychiatric Association: Washington, DC.

Compton WM, Conway KP, Stinson FS, Colliver JD, Grant BF (2005). Prevalence, correlates, and comorbidity of DSM-IV antisocial personality syndromes and alcohol and specific drug use disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry* **66**, 677–685.

Costa Jr. P, Terracciano A, McCrae RR (2001). Gender differences in personality traits across cultures: robust and surprising findings. *Journal of Personality and Social Psychology* **81**, 322–331.

de Wit H, Richards JB (2004). Dual determinants of drug use in humans: reward and impulsivity. In *Nebraska Symposium on Motivation. Vol. 50: Motivational Factors in the Etiology of Drug Abuse* (ed. R. A. Bevins and M. T. Bardo), pp. 19–55. University of Nebraska Press: Lincoln.

Fanous A, Gardner CO, Prescott CA, Cancro R, Kendler KS (2002). Neuroticism, major depression and gender: a population-based twin study. *Psychological Medicine* **32**, 719–728.

First MB, Spitzer RL, Williams JBW, Gibbon M (1997). *Structured Clinical Interview for DSM-IV* (*SCID*). American Psychiatric Association: Washington, DC.

Giancola PR, Parker AM (2001). A six-year prospective study of pathways toward drug use in adolescent boys with and without a family history of a substance use disorder. *Journal of Studies on Alcohol* **62**, 166–178.

Giancola PR, Tarter RE (1999). Executive cognitive functioning and risk for substance abuse. *Psychological Science* **10**, 203–205.

Goodwin RD, Gotlib IH (2004). Gender differences in depression: the role of personality factors. *Psychiatry Research* **126**, 135–142.

Grant BF, Hasin DS, Chou SP, Stinson FS, Dawson DA (2004*a*). Nicotine dependence and psychiatric disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry* **61**, 1107–1115.

Grant BF, Stinson FS, Dawson DA, Chou SP, Ruan WJ, Pickering RP (2004*b*). Co-occurrence of 12-month alcohol and drug use disorders and personality disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry* **61**, 361–368.

Guo J, Hill KG, Hawkins JD, Catalano RF, Abbott RD (2002). A developmental analysis of sociodemographic, family, and peer effects on adolescent illicit drug initiation. *Journal of the American Academy of Child and Adolescent Psychiatry* **41**, 838–845.

Hawkins J, Catalano RF, Miller JY (1992). Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: implications for substance abuse prevention. *Psychological Bulletin* **112**, 64–105.

Hicks BM, Blonigen DM, Kramer MD, Krueger RF, Patrick CJ, Iacono WG, McGue M (2007). Gender differences and developmental change in externalizing disorders from late adolescence to early adulthood: a longitudinal-twin study. *Journal of Abnormal Psychology* **116**, 433–447.

Holdcraft LC, Iacono WG, McGue MK (1998). Antisocial personality disorder and depression in relation to alcoholism: a community-based sample. *Journal of Studies on Alcohol* **59**, 222–226.

Hudson JI, Mangweth B, Pope Jr. HG, De Col C, Hausmann A, Gutweniger S, Laird NM, Biebl W, Tsuang MT (2003). Family study of affective spectrum disorder. *Archives of General Psychiatry* **60**, 170–177.

Hudson JI, Pope Jr. HG (1990). Affective spectrum disorder: does antidepressant response identify a family of disorders with a common pathophysiology? *American Journal of Psychiatry* **147**, 552–564.

Iacono WG, Carlson SR, Taylor J, Elkins IJ, McGue M (1999). Behavioral disinhibition and the development of substance-use disorders: findings from the Minnesota Twin Family Study. *Development and Psychopathology* **11**, 869–900.

Jardine R, Martin NG, Henderson AS (1984). Genetic covariation between neuroticism and the symptoms of anxiety and depression. *Genetic Epidemiology* **1**, 89–107.

Jenkins R, Lewis G, Bebbington P, Brugha T, Farrell M, Gill B, Meltzer H (2003). The national psychiatric morbidity surveys of Great Britain: initial findings from the Household Survey. *International Review of Psychiatry* **15**, 29–42.

Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ (1993). A longitudinal twin study of personality and major depression in women. *Archives of General Psychiatry* **50**, 853–862.

Kendler KS, Prescott CA, Myers J, Neale MC (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry* **60**, 929–937.

Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, Wittchen HU, Kendler KS (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.

Klose M, Jacobi F (2004). Can gender differences in the prevalence of mental disorders be explained by sociodemographic factors? *Archives of Women's Mental Health* **7**, 133–148.

Knutson B, Wolkowitz OM, Cole SW, Chan T, Moore EA, Johnson RC, Terpstra J, Turner RA, Reus VI (1998). Selective alteration of personality and social behavior by serotonergic intervention. *American Journal of Psychiatry* **155**, 373–379.

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

Krueger RF, Caspi A, Moffitt TE, Silva PA (1998). The structure and stability of common mental disorders (DSM-III-R): a longitudinal–epidemiological study. *Journal of Abnormal Psychology* **107**, 216–227.

Krueger RF, Caspi A, Moffitt TE, Silva PA, 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.

Krueger RF, Hicks BM, Patrick CJ, Carlson SR, Iacono WG, McGue M (2002). Etiologic connections among substance dependence, antisocial behavior and personality: modeling the externalizing spectrum. *Journal of Abnormal Psychology* **111**, 411–424.

Lykken DT, Bouchard Jr. TJ, McGue M, Tellegen A (1990). The Minnesota Twin Family Registry: some initial findings. *Acta Geneticae Medicae et Gemellologiae* **39**, 35–70.

Lynam DR, Miller JD (2004). Personality pathways to impulsive behavior and their relations to deviance: results from three samples. *Journal of Quantitative Criminology* **20**, 319–341.

Meredith W (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika* **58**, 525–543.

Moffitt TE, Caspi A, Rutter M, Silva PA (2001). *Sex Differences in Antisocial Behaviour: Conduct disorder, Delinquency, and Violence in the Dunedin Longitudinal Study*. Cambridge University Press: New York, NY.

Muthén LK, Muthén BO (2004). *MPlus User's Guide*. Muthén & Muthén: Los Angeles, CA.

Nishizawa S, Benkelfat C, Young SN, Leyton M, Mzengeza S, de Montigny C, Blier P, Diksic M (1997). Differences between males and females in rates of serotonin synthesis in human brain. *Proceedings of the National Academy of Sciences USA* **94**, 5308–5313.

Nolen-Hoeksema S (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology* **100**, 569–582.

Nolen-Hoeksema S (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology* **109**, 504–511.

Nolen-Hoeksema S (2004). Gender differences in risk factors and consequences for alcohol use and problems. *Clinical Psychology Review* **24**, 981–1010.

Nolen-Hoeksema S, Jackson B (2001). Mediators of the gender difference in rumination. *Psychology of Women Quarterly* **25**, 37–47.

Nolen-Hoeksema S, Larson J, Grayson C (1999). Explaining the gender difference in depressive symptoms. *Journal of Personality and Social Psychology* **77**, 1061–1072.

Reinherz HZ, Giaconia RM, Hauf AM, Wasserman MS, Paradis AD (2000). General and specific childhood risk factors for depression and drug disorders by early adulthood. *Journal of the American Academy of Child and Adolescent Psychiatry* **39**, 223–231.

Roberts SB, Kendler KS (1999). Neuroticism and self-esteem as indices of the vulnerability to major depression in women. *Psychological Medicine* **29**, 1101–1109.

Robins LN, Regier DA (1991). *Psychiatric Disorders in America: The Epidemiologic Catchment Area Study*. Free Press: New York.

Sher KJ, Bartholow BD, Wood MD (2000). Personality and substance use disorders: a prospective study. *Journal of Consulting and Clinical Psychology* **68**, 818–829.

Tarter RE, Kirisci L, Mezzich A, Cornelius JR, Pajer K, Vanyukov M, Gardner M, Blackson T, Clark D (2003). Neurobehavior disinhibition in childhood predicts early age at onset of substance use disorder. *American Journal of Psychiatry* **160**, 1078–1085.

Tellegen A (in press). *Manual for the Multidimensional Personality Questionnaire*. University of Minnesota Press: Minneapolis.

Watson D, Gamez W, Simms LJ (2005). Basic dimensions of temperament and their relation to anxiety and depression: a symptom-based perspective. *Journal of Research in Personality* **39**, 46–66.

Watson D, Kotov R, Gamez W (2006). Basic dimensions of temperament in relation to personality and psychopathology. In *Personality and Psychopathology* (ed. R. F. Krueger and J. L. Tackett). Guilford Press: New York, NY.

Wills TA, Vaccaro D, McNamara G (1994). Novelty seeking, risk taking, and related constructs as predictors of adolescent substance use: an application of Cloninger's theory. *Journal of Substance Abuse* **6**, 1–20.

Young E, Korszun A (1999). Women, stress, and depression: sex differences in hypothalamic–pituitary–adrenal axis regulation. In *Review of Psychiatry Series. Vol. 18: Gender Differences in Mood and Anxiety Disorders: From Bench to*

Bedside (ed. E. Leibenluft), pp. 31–52. American Psychiatric Association: Washington, DC.

Young EA, Abelson JL, Cameron OG (2004). Effect of comorbid anxiety disorders on the hypothalamic–pituitary–adrenal axis response to a social stressor in major depression. *Biological Psychiatry* **56**, 113–120.

Zimmerman M, Mattia JI (2001*a*). The Psychiatric Diagnostic Screening Questionnaire: development, reliability and validity. *Comprehensive Psychiatry* **42**, 175–189.

Zimmerman M, Mattia JI (2001*b*). A self-report scale to help make psychiatric diagnoses: the Psychiatric Diagnostic Screening Questionnaire. *Archives of General Psychiatry* **58**, 787–794.