Journal of Abnormal Psychology
1998, Vol. 107, No. 2, 216–227

0021-843X/98/$3.00

The Structure and Stability of Common Mental Disorders (*DSM–III–R*): A Longitudinal–Epidemiological Study

Robert F. Krueger
University of Wisconsin—Madison

Avshalom Caspi and Terrie E. Moffitt
Institute of Psychiatry, University of London and
University of Wisconsin—Madison

Phil A. Silva
University of Otago Medical School

The latent structure and stability of 10 common mental disorders were examined in a birth cohort at ages 18 and 21. A 2-factor model, in which some disorders were presumed to reflect internalizing problems and others were presumed to reflect externalizing problems, provided a more optimal fit to the data than either a 1- or a 4-factor model. To a significant extent, persons in the sample retained their relative positions on the latent factors across the 3-year period from age 18 to age 21. Results offer potential clarification of the meaning of comorbidity in psychopathology research by suggesting that comorbidity may result from common mental disorders being reliable, covariant indicators of stable, underlying "core psychopathological processes."

Classification of mental disorders remains a controversial topic in psychopathology research; numerous and varied approaches to establishing a scientific taxonomy of mental disorders have been proposed (for reviews, see Blashfield, 1984; Millon, 1991; Skinner, 1981). However, one classification scheme has gained relative ascendancy in recent years. As stated in a recent *Journal of Abnormal Psychology* editorial, "the perspective of our field is substantially influenced by the *Diagnostic and Statistical Manual of Mental Disorders*" (*DSM;* Strauss, 1995, p. 555). Accordingly, much recent research on mental disorders has been conducted using a classification strategy implicitly endorsed by the *DSM*—a strategy in which disordered participants are carefully screened for the presence of a "target" disorder, to the exclusion of other disorders.

Nevertheless, the results of research pursued using this strategy may be difficult to interpret: High rates of comorbidity among mental disorders have been observed in numerous clinical and epidemiological samples (Clark, Watson, & Reynolds, 1995; Kessler et al., 1994; Maser & Cloninger, 1990). That is, mental disorders co-occur at greater than chance rates, calling into question the meaning of research performed on "pure" cases of disordered participants (e.g., research involving persons suffering from a major depressive episode, who do not meet criteria for other *DSM* disorders). Because of high rates of comorbidity, such pure cases are not only rare, they may also be unrepresentative of the entire spectrum of persons suffering from the target disorder. Consider the high rates of documented comorbidity among major depression and the *DSM* anxiety disorders (Kendall & Watson, 1989; Maser & Cloninger, 1990). If most persons who suffer from major depressive episode symptoms also suffer from the symptoms of an anxiety disorder, a study that involves only persons who have major depressive episode symptoms may inadvertently involve an unusual and rare group of persons who underrepresent the entire range of persons in the population who suffer from major depressive episode symptoms. The generalizability and relevance of the findings from such a study would be limited because its participants have been systematically chosen to not represent the entire range of persons in the population who suffer from major depressive episode symptoms.

Observations such as these led a recent review of research on the diagnosis and classification of psychopathology to call for "a systematic examination of comorbidity patterns to elucidate the broad, higher-order structure of phenotypic psychopathology" (Clark et al., 1995, p. 131). In response to this call,

Robert F. Krueger, Department of Psychology, University of Wisconsin—Madison; Avshalom Caspi and Terrie E. Moffitt, Institute of Psychiatry, University of London, and Department of Psychology, University of Wisconsin—Madison; Phil A. Silva, Dunedin Multidisciplinary Health and Development Research Unit, University of Otago Medical School, Dunedin, New Zealand. Robert F. Krueger is now at the Department of Psychology, University of Minnesota.

Robert F. Krueger was supported by a Javits Fellowship from the U.S. Department of Education. This research was also supported by United States Public Health Service (USPHS) Grant MH-45070 to Terrie E. Moffitt from the Violence and Traumatic Stress Branch of the National Institute of Mental Health (NIMH), by USPHS Grant MH-49414 to Avshalom Caspi from the Personality and Social Processes Branch of the NIMH, by the William T. Grant Foundation, and by the William Freeman Vilas Trust and H. I. Romnes Fellowship at the University of Wisconsin. The Dunedin Multidisciplinary Health and Development Research Unit is supported by the New Zealand Health Research Council. We are grateful to Roger Blashfield, Kristen Kling, Scott Lilienfeld, Thomas Piasecki, Timothy Strauman, Brad Wright, and two anonymous reviewers for their technical and editorial assistance. We are also grateful to the Dunedin Unit investigators and staff and to the study members and their families.

Correspondence concerning this article should be addressed to Terrie E. Moffitt, Social, Genetic, and Developmental Psychiatry Centre, Institute of Psychiatry, III Denmark Hill, London SE5 8AF, England.

the present article investigates the latent structure underlying 10 common *DSM–III–R* (*DSM*, 3rd ed. rev.; American Psychiatric Association, 1987) mental disorders using a well-known clinical interview. We assessed these disorders on two occasions (ages 18 and 21) in an epidemiological sample. The 10 disorders we studied (major depressive episode, dysthymia, generalized anxiety disorder, agoraphobia, social phobia, simple phobia, obsessive–compulsive disorder, marijuana dependence, alcohol dependence, and conduct disorder–antisocial personality disorder) are important not only because many people suffer from them (Kessler et al., 1994), but also because they entail high social costs (e.g., diminished productivity at work and at school; Newman et al., 1996). In addition, we chose to model common mental disorders because their base rates of occurrence (in our epidemiological sample as well as in others; cf. Kessler et al., 1994; Robins & Regier, 1991) were high enough to allow for variance in psychopathological status among our participants. For rarer disorders, there are simply too few opportunities to determine whether the disorder co-occurs with other disorders at greater than chance rates (i.e., the correlation between the rare disorder and other disorders cannot be accurately assessed because there is too little variance in the sample in terms of the rare disorder). For example, we did not model panic disorder because the variance in this disorder was too restricted to produce meaningful correlations with other disorders (0.9% of the sample members who completed our diagnostic interview at age 18 met the criteria for this diagnosis; at age 21, the corresponding figure was 0.6%).

Our study has two aims. The first aim was to use confirmatory factor analysis (CFA) to evaluate alternative hypotheses about the latent structure underlying our 10 disorders. The research strategy implicit in the *DSM* has led researchers to study common mental disorders in relative isolation from one another; for example, separate journals are devoted to the study of affective disorders (e.g., *Journal of Affective Disorders*), anxiety disorders (e.g., *Journal of Anxiety Disorders*), substance abuse (e.g., *Journal of Studies on Alcohol*) and antisocial behavior (e.g., *Criminology*). However, the high rates of comorbidity observed among all these disorders suggest that a more parsimonious structure than the one implied by the *DSM* (and reflected in current trends in research specialization) may underlie common mental disorders. If this is true, then research on any one of these disorders should not proceed in isolation from research on other disorders; each disorder may be an alternate expression of one of a number of common, underlying core psychopathological deficits. These hypothetical core psychopathological deficits themselves (regardless of their specific expressions) may be the more profitable topic of study. In the current research, we used CFA to evaluate different models of the structure of phenotypic psychopathology. We compared and cross-validated these different models in an epidemiological sample that has been assessed with the same clinical interview in late adolescence and again in early adulthood.

In addition to using CFA to evaluate the latent structure underlying our 10 disorders, a second aim of this study was to use structural equation modeling to examine the stability of this structure across a 3-year period, tracking our research participants from age 18 to age 21. In addition to its influence on research strategies and specializations, the *DSM*'s categorical conception of mental disorder also influences conceptions of continuity and change in psychopathological status. For example, the *DSM–IV* describes many mental disorders as episodic (e.g., major depressive disorder; American Psychiatric Association, 1994, pp. 344–345). There are at least two key issues that must be addressed in evaluating the *DSM*'s claim that mental disorders such as major depression are episodic. The first issue involves the psychometric distinction between actual and artifactual change: An individual's score on a given measure (e.g., the presence vs. absence of a psychiatric diagnosis) is influenced by at least two components, true score variance and random error variance. Apparent change in psychiatric status (e.g., the observation that a person who met the *DSM* criteria for major depression no longer meets the criteria) may sometimes be due to errors of measurement rather than to actual change (Fergusson, Horwood, & Lynskey, 1995). Because structural equation modeling is capable of separating true score from error variance, we were able to determine the stability of mental disorder in our sample, corrected for random measurement error.

The second key issue involves the conceptual distinction between absolute stability and differential (or rank-order) stability: Even if an individual no longer meets the criteria for a given diagnosis, he or she may still be experiencing a similar overall level of psychiatric symptomatology relative to his or her peers. Consider the example of an individual (A) who once met the criteria for major depression and now no longer meets those criteria. Even if this observed change is nonartifactual (i.e., not due to errors of measurement), A may have remained at the same relative level of symptom expression (e.g., at the 75th percentile) in the population. This is an important distinction because, although A may not be currently diagnosable according to *DSM* criteria, A has not changed in relative psychopathological level. But this distinction is lost under the categorical *DSM* model, which regards A as psychopathologically identical to another individual (B) at the 25th percentile in the population, virtually symptom-free. Now suppose, for example, that the same stressful life event occurs in the lives of both A and B (e.g., a severe loss; see Oatley, 1988). It seems likely that their differences in rank order will have implications for their continued functioning (i.e., A would be more likely to relapse to a diagnosable state). Potentially valuable information about rank-order stability is lost when A and B are equated subsequent to A's relapse. Hence, under the categorical *DSM* model, the observation of a large number of individuals such as A can lead to the inference that major depression is an episodic disorder. This inference, however, seems unwarranted if these individuals, although in remission, continue to be the most disordered persons in the population.

In the current research, we used structural equation modeling (performed on latent factors derived from CFA) to assess the rank-order stability of psychiatric disorder in our sample over time. We turn now to further explore the conceptual and methodological advantages of confirmatory factor-analytic–structural equation models in the classification of mental disorder and to consider candidate dimensions that may underlie common mental disorders.

Conceptual Advantages of the CFA Model

The CFA approach to mental disorders may provide certain conceptual advantages in modeling a number of vexing empiri-

cal phenomena related to the classification of mental disorder. As noted earlier, one vexing problem is comorbidity (Feinstein, 1970). The existence of high levels of comorbidity among *DSM–III* (American Psychiatric Association, 1980) and *DSM–III–R*[1] disorders calls into question the neo-Kraepelinian assumptions embodied in these recent *DSM* systems; that is, psychopathological signs and symptoms do not divide easily into discrete and mutually exclusive categories (Maser & Cloninger, 1990). In this article, we attempt to determine if the substantial levels of comorbidity that have been observed among *DSM–III* and *DSM–III–R* disorders (Kessler et al., 1994; Robins & Regier, 1991) might be understood as meaningful covariation among diagnoses. Evidence that comorbidity can be modeled as meaningful covariation would support a conceptualization of *DSM* disorders as manifest indicators of latent factors that underlie varied forms of psychiatric distress.

A second vexing problem concerns the severity of mental illness. Severity predicts not only the intensity and longevity of the disorder under consideration but also the likelihood that the individual will meet the criteria for another disorder. That is, severity and comorbidity are positively correlated: Persons who are severely disordered (i.e., highly and persistently symptomatic) are more likely to meet criteria for other disorders than are persons who are less severely disordered (Clark et al., 1995). This phenomenon is difficult to explain if mental disorders are presumed to be relatively independent, discrete entities. However, if psychiatric diagnoses are indicators of latent factors underlying multiple mental disorders, we would, in fact, expect there to be a positive correlation between severity and comorbidity. Conceptualized in these terms, persons with more positive indicators (i.e., more disorders) are better representatives of the high poles of the latent dimensions underlying manifest disorders. Consider the general factor theory of psychometric intelligence as an analogy (e.g., Jensen, 1980). The higher an individual is on the latent general intelligence dimension (g), the greater the likelihood that he or she will pass any given test item and the greater the likelihood that he or she will pass multiple test items. Similarly, if a latent dimension underlies multiple mental disorders, then the higher an individual is on the latent general psychopathology dimension (p), the greater the likelihood that he or she will meet the criteria for any given diagnosis (pass any given test item) and the greater the likelihood that he or she will meet the criteria for multiple diagnoses (pass multiple test items).

Methodological Advantages of the CFA Model

The CFA model conceives of psychiatric diagnoses as indicators of underlying dimensions: In this model, the presence or absence of a psychiatric disorder provides information about where an individual lies on a latent continuum underlying multiple disorders. Although enthusiasm about the utility of dimensional models in the classification of psychopathology has been expressed in the past (e.g., Cattell, 1965; Eysenck, 1960; Menninger, 1963), such enthusiasm has waned in recent years.[2] Critics of dimensional models have asserted that there is little agreement regarding both the number and the nature of the dimensions that are jointly necessary and sufficient to characterize various psychopathologies (see, e.g., Millon, 1991). Much of this controversy has centered around problems with exploratory factor analysis (EFA), the principal statistical tool used by early proponents of dimensional systems. Two distinct problems with EFA have been noted (Watson, Clark, & Harkness, 1994). Critics have noted that there are no infallible guidelines to aid the investigator in determining the appropriate number of factors to extract in EFA. In addition, critics have noted that there are no infallible guidelines to aid the investigator in determining the correct orientation of the resultant factors in multidimensional space. Although these problems with EFA are not insuperable (Watson et al., 1994), the CFA model offers certain advantages over EFA in evaluating dimensional models for the classification of psychopathology. In contrast to EFA, CFA requires the investigator to specify a theory about the number and orientation of the latent factors that may explain observed variations in psychiatric signs and symptoms. Moreover, in CFA, such theories are expressed as mathematical formalisms that can then be tested for their verisimilitude, thereby providing a riskier test of the investigator's theory than that provided by EFA (cf. Meehl, 1978).

Candidate Dimensions Underlying Multiple Mental Disorders

What latent dimensions might underlie multiple common mental disorders? Three possibilities were evaluated in our research. First, we considered the possibility that one single latent dimension underlies all 10 disorders we studied. This possibility was suggested to us by factor analyses of self-report personality and psychopathology inventories, such as the Minnesota Multiphasic Personality Inventory (MMPI). For example, factor analyses of the MMPI Clinical and Validity scales have consistently revealed the presence of one dominating factor that appears to reflect general maladjustment (Graham, 1993). Hence, we evaluated the possibility that this single general maladjustment factor might be sufficient to explain the patterns of covariance we observed among the 10 disorders we studied.

Second, we turned to literature on the classification of childhood psychopathology. In contrast to the prominence of categorical models in the classification of adult psychopathologies, dimensional models have enjoyed much success in the classification of childhood psychopathologies. This is most likely due to the fact that child psychopathology researchers have traditionally relied on symptom checklists completed by parents and teachers to assess psychopathology (as opposed to diagnostic interviews). Numerous factor analyses of such checklists have been performed, and despite the use of diverse item sets and methodologies, these studies have converged in identifying two primary dimensions of childhood psychopathology: *internalizing* and *externalizing* (Achenbach & Edelbrock, 1978; Quay,

[1] Although *DSM–IV* is the current official classification system, the large-scale epidemiological studies that have convincingly established the magnitude of the comorbidity phenomenon used the *DSM–III* and *DSM–III–R* systems (Kessler et al., 1994; Robins & Regier, 1991).

[2] One important exception to this general trend is in the area of personality disorder, in which dimensional models are gaining increasing acceptance (see, e.g., Clark, Livesley, Schroeder, & Irish, 1996; Costa & Widiger, 1994).

1986). The internalizing dimension summarizes anxious, depressed, somatic, obsessive, and compulsive symptoms, whereas the externalizing dimension summarizes attention deficit, aggressive, and delinquent symptoms (Achenbach & Edelbrock, 1984). Both dimensions show moderate stability across time (Ollendick & King, 1994), suggesting the possibility that this two-dimensional structure may persist into adulthood. Nevertheless, from childhood to adulthood, psychopathology changes its phenotypic expression, particularly with the emergence of substance abuse problems in adolescence. The extant literature, however, suggests that substance abuse disorders belong with other externalizing disorders. For example, in the sample used in the current research (which has been studied longitudinally since childhood), many persons meeting adult criteria for substance dependence at age 21 met childhood criteria for mental disorder at a prior assessment. Conduct disorder (an externalizing disorder) was the most common childhood diagnosis to antedate a diagnosis of substance dependence at age 21 (Newman et al., 1996).

Third, and finally, we evaluated a model inspired by current trends in research specialization and by the organization of recent *DSM* manuals. As noted earlier, separate journals are devoted to reporting research specifically on affective, anxiety, substance dependence, and antisocial disorders. Moreover, the *DSM–IV* (American Psychiatric Association, 1994) has separate sections for substance-related disorders, mood disorders, and anxiety disorders, and places disorders involving antisocial behavior in other, separate sections (conduct disorder appears in a section entitled Disorders Usually First Diagnosed in Infancy, Childhood, or Adolescence, and antisocial personality disorder appears in a section entitled Personality Disorders). Hence, we also evaluated a four-factor model in which separate factors were presumed to underlie affective, anxiety, substance dependence, and antisocial disorders.

A CFA Model of *DSM–III–R* Mental Disorders in a Birth Cohort

In the current research, we compared three accounts (one-, two-, and four-factor models) of potential dimensional structures that may underlie 10 common *DSM–III–R* diagnoses assessed in an entire birth cohort. We conducted the present research in an epidemiological sample in order to obtain accurate estimates of the correlations among the diagnoses; such estimates can be biased in nonrepresentative samples (Mednick, 1978). In addition, we attempted to assess the replicability and stability of the latent dimensions by conducting independent assessments of disorder at two points in time: ages 18 and 21. First, in the age-18 data, we tested a one-factor general maladjustment model in which all disorders were specified as indicators of a single latent dimension. We compared the fit of this model with the fit of a two-factor model in which major depressive episode, dysthymia, generalized anxiety disorder, agoraphobia, social phobia, simple phobia, and obsessive–compulsive disorder were conceptualized as indicators of a latent internalizing dimension; and conduct disorder, marijuana dependence, and alcohol dependence were conceptualized as indicators of a latent externalizing dimension. We then compared the fit of the one- and two-factor models with the fit of a four-factor model in which major depressive episode and dysthymia were indicators of a latent affective disorder factor; generalized anxiety disorder, agoraphobia, social phobia, simple phobia, and obsessive–compulsive disorder were indicators of a latent anxiety disorder factor; marijuana dependence and alcohol dependence were indicators of a latent substance dependence factor; and conduct disorder was an indicator of a latent antisocial behavior factor.

Next, we attempted to replicate our findings at age 18 by examining and comparing the fit of the one-, two-, and four-factor models in the age-21 data, with the exception that conduct disorder was replaced with the more age-appropriate diagnosis of antisocial personality disorder. Although conduct disorder is a required antecedent of antisocial personality disorder, such conduct disorder must occur before age 15 (American Psychiatric Association, 1987); hence, the age-18 conduct disorder and age-21 antisocial personality disorder indicators do not overlap for spurious reasons. Contingent on replicating a specific factor structure at ages 18 and 21, we assessed the rank-order stability of disorder in our sample by estimating the correlations linking the latent factors across the 3-year interval from age 18 to age 21.

Method

Sample Members

Sample members belonged to an unselected birth cohort that has been studied extensively since birth, as part of the Dunedin Multidisciplinary Health and Development Study. The sample and the history of the study have been described in detail by Silva (1990). Briefly, the study is a longitudinal investigation of the health, development, and behavior of a complete cohort of births between April 1, 1972, and March 31, 1973, in Dunedin, New Zealand, a city of 120,000 people. Perinatal data were obtained at delivery, and when the children were later traced for follow-up at age 3, 1,037 (52% boys and 48% girls, 91% of the eligible births) participated in the assessment, forming the base sample for the longitudinal study. Since age 3, 17 sample members have died. With regard to social stratification, the children's families were representative of the social class and ethnic distribution of the general population on New Zealand's South Island. The sample members are of predominantly European ancestry; fewer than 7% identify themselves as Maori or Polynesian.

Data collection procedure. The Dunedin sample has been reassessed with a diverse battery of psychological, medical, and sociological measures at ages 3, 5, 7, 9, 11, 13, 15, 18, and most recently at age 21. The basic research procedure involves bringing each sample member into the research unit within 60 days of his or her birthday for a full day of individual data collection. The various research topics are presented in different private interview rooms as standardized modules by different trained examiners in counterbalanced order throughout the day. Because there has never been a violation of confidentiality, this sample is by now unusually willing to provide frank reports. Printed brochures about how to get help for mental disorders were made available in the waiting area, and referral was available for sample members reporting suicidal intent.

Attrition. In 1990–1991, 98.8% of the living members of the cohort agreed to participate in the age-18 follow-up assessment. Mental health interviews were completed for 930 of the 18-year-old study members. Study members who completed the mental health interview at age 18 did not differ significantly from nonparticipants ($n = 107$) in socioeconomic status (SES), $t(939) = 1.85$, *ns*, or sex, $\chi^2(1, N = 1{,}037) = .92$, *ns*.

In 1993–1994, 97.3% of the living members of the cohort agreed to

participate in the age-21 follow-up assessment. Mental health interviews were completed for 961 of the 21-year-old study members. Study members who completed the mental health interview at age 21 did not differ significantly from nonparticipants ($n = 76$) in SES, $t(939) = 1.35$, *ns*, or sex, $\chi^2(1, N = 1{,}037) = 1.26$, *ns*.

Measurement of Psychopathology

Disorders at age 18. At age 18, the Diagnostic Interview Schedule (DIS; Version III–R; Robins, Helzer, Cottler, & Goldring, 1989) was used to obtain diagnoses of mental disorder in the previous 12 months. The DIS was developed by the National Institute of Mental Health for the Epidemiologic Catchment Area (ECA) program (Regier et al., 1984). We modified the DIS to use only those items that were criteria for *DSM–III–R* classifications, to omit lifetime prevalence questions, and to score items as 0 (*no*), 1 (*sometimes*), and 2 (*yes, definitely*). In identifying disorder, only scores of 2 were used to indicate a positive response (commensurate with a 5 in the original DIS). Our diagnoses were made in the same way that diagnoses were made in the ECA studies: They were based on the then-current *DSM* criteria for each disorder (i.e., *DSM–III–R*) and were given regardless of whether an additional coexisting diagnosis was given (cf. Leaf, Myers, & McEvoy, 1991).

Of the sample, 44% met the requisite *DSM–III–R* criteria for a 12-month disorder at age 18. Although this estimate may seem high, it is consistent with prevalence data for this age group from the ECA studies (Robins & Regier, 1991) and from the National Comorbidity Survey (Kessler et al., 1994). For the current study, we used data regarding 10 disorders that were found to have high prevalence rates in our sample at age 18 and that were also assessed at age 21: Major depressive episode (17.3% of the sample members who completed the DIS at age 18 met the criteria for this diagnosis), dysthymia (3.2%), generalized anxiety disorder (1.8%), agoraphobia (4.8%), social phobia (13.8%), simple phobia (7.4%), obsessive–compulsive disorder (4.6%), conduct disorder (8.1%), marijuana dependence (6.6%), and alcohol dependence (16.3%). As reported previously by Feehan, McGee, Nada Raja, and Williams (1994), high rates of comorbidity were observed among these disorders; of the persons in our sample who met criteria for at least one disorder at age 18, 46% also met criteria for one or more additional disorders. An extensive report on the mental health of the Dunedin sample at age 18 may be found in Feehan et al. (1994).

Disorders at age 21. At age 21, the same DIS administered at age 18 was used to obtain diagnoses of mental disorder in the previous 12 months. Diagnoses were made on the basis of the *DSM–III–R* criteria for each disorder, were given regardless of whether an additional coexisting diagnosis was given, and were made without knowledge of the individual's diagnostic status at age 18.

Of the sample, 40% met the requisite *DSM–III–R* criteria for a 12-month disorder at age 21. For the current study, we used data regarding 10 diagnoses that were found to have high prevalence rates in our sample at age 21 and that were also assessed at age 18: Major depressive episode (16.8% of the sample members who completed the DIS at age 21 met the criteria for this diagnosis), dysthymia (3.0%), generalized anxiety disorder (1.9%), agoraphobia (3.8%), social phobia (9.7%), simple phobia (8.4%), obsessive–compulsive disorder (7.1%), anti-social personality disorder (3.2%), marijuana dependence (9.6%), and alcohol dependence (9.8%). As reported by Newman et al. (1996), high rates of comorbidity were observed among these disorders; of the persons in our sample who met criteria for at least one disorder at age 21, 47.3% also met criteria for one or more additional disorders. An extensive report on the mental health of the Dunedin sample at age 21 may be found in Newman et al. (1996); information regarding zero-order relations among variables derived from the DIS at ages 18 and 21 may be found in Krueger, Caspi, Moffitt, Silva, and McGee (1996).

Data Analysis

The PRELIS computer program, version 1.20 (Jöreskog & Sörbom, 1988), was used to create tetrachoric correlation matrices and asymptotic covariance matrices for the 10 diagnostic variables, both separately at ages 18 and 21, and simultaneously at both ages. Because of the dichotomous nature of the diagnostic variables, we analyzed tetrachoric correlations among the 10 disorders. Our CFA models were tested using the LISREL computer program, version 7.20 (Jöreskog & Sörbom, 1989). The models were estimated using the weighted least squares (WLS) procedure. WLS is the appropriate choice for the analysis of dichotomous, ordinal variables such as psychiatric diagnoses because, unlike the commonly used maximum likelihood procedure, WLS does not involve the assumption that the measured variables have a multivariate normal distribution in the population (Jöreskog & Sörbom, 1989, p. 202). However, WLS (as implemented in LISREL) does require complete data on every case; hence, cases were removed listwise for each analysis we performed (at age 18, $N = 930$; at age 21, $N = 937$; and for both ages combined, $N = 882$). Our latent factors were assigned units of measurement by fixing one path linking each latent factor to a measured variable at 1.0. We evaluated the fit of our models using the chi-square goodness-of-fit statistic, the goodness-of-fit index (GFI), the adjusted goodness-of-fit index (AGFI), root mean square residual (RMSR), and the Bayesian information criterion (BIC). The BIC statistic is computed as $\chi^2 - df \ln N$, where χ^2 is the chi-square fit statistic for the model, df is the corresponding degrees of freedom, and N is the sample size (Raftery, 1995). Increasingly negative values of BIC correspond to increasingly better fitting models, and, in comparing two models, differences in BIC larger than 10 represent very strong evidence in favor of the model with the smaller BIC value (Raftery, 1995).

For a given sample, BIC balances two important quantities in assessing the fit of a model: the discrepancy between the sample and the fitted correlation matrices and the number of parameter estimates required to achieve this fit. This can be seen clearly in BIC's formula: The chi-square value, an index of the discrepancy between the sample and the fitted correlation matrices, is offset by an evaluation of the number of parameter estimates required to achieve this fit (i.e., the df, as models with more free parameters have smaller df values). Thus, BIC prefers parsimonious models (i.e., those with larger df values), as long as this parsimony is not achieved at the expense of accurately reproducing the sample correlation matrix. BIC thus quantifies the idea that, in comparing scientific theories (and their representations as statistical models), we should prefer the theory that achieves the best verisimilitude while postulating the minimum number of theoretical entities.

Results

Mental Disorders at Age 18

What factors underlie the 10 disorders at age 18? To answer this question, we fit one-factor, two-factor-oblique, and four-factor-oblique models to the age-18 diagnostic data ($N = 930$). Fit indices for all three models can be seen in Table 1.[3] In the

[3] To identify the four-factor model in single-wave data (i.e., within age 18 and within age 21), it was necessary to assume that conduct disorder at age 18 and antisocial personality disorder at age 21 were measured without error. This is because these variables are the only indicators of a latent antisocial behavior factor, and, for single-indicator latent variables, it is not possible to allow both the variance of the latent variable and the error in its indicator to be freely estimated. To make comparisons between the one-, two-, and four-factor models fair (in the sense that the only thing that changed between the models was the number of factors postulated), conduct disorder and antisocial personality disorder were assumed to be measured without error in all analyses reported in Table 1.

Table 1
Fit Indices for One-, Two-, and Four-Factor Models of 10 DSM–III–R *Diagnoses at Ages 18 and 21*

Age/model	Fit index						
	χ^2	*df*	*p*	GFI	AGFI	RMSR	BIC
Age 18 (n = 930)							
One factor	168.13	36	.00	.97	.95	.17	−77.94
Two factor	41.61	35	.21	.99	.99	.09	−197.62
Four factor	31.98	30	.37	.99	.99	.08	−173.08
Age 21 (n = 937)							
One factor	156.28	36	.00	.97	.96	.20	−90.06
Two factor	59.15	35	.01	.99	.98	.11	−180.34
Four factor	46.24	30	.03	.99	.98	.10	−159.04

Note. *DSM–III–R* = *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.); χ^2 = chi-square goodness-of-fit index; *df* = degrees of freedom for chi-square; *p* = *p* value for chi-square; GFI = goodness-of-fit index; AGFI = adjusted goodness-of-fit index; RMSR = root mean squared residual; BIC = Bayesian information criterion.

one-factor model, all 10 disorders were specified as indicators of a single latent factor. Although this model does not provide an unreasonable fit to the data at age 18 (as suggested by the relatively high GFI and AGFI), there may still be room for improvement (as suggested by the large and significant chi-square value and the relatively larger RMSR). We therefore compared the fit of this one-factor model with the fit of a two-factor model in which major depressive episode, dysthymia, generalized anxiety disorder, agoraphobia, social phobia, simple phobia, and obsessive–compulsive disorder were indicators of a latent internalizing factor, and conduct disorder, marijuana dependence, and alcohol dependence were indicators of a latent externalizing factor. The two-factor model represented a substantial improvement over the one-factor model, yielding a very small and nonsignificant chi-square, an essentially perfect fit according to GFI and AGFI, a relatively small RMSR, and a much more negative value for BIC than was generated by the one-factor model.

Do we need more than two factors to offer a good account of the correlations observed among the 10 disorders at age 18? The excellent fit of the two-factor model suggests not; nevertheless, to address this question directly, we fit a four-factor model in which major depressive episode and dysthymia were indicators of a latent affective disorder factor; generalized anxiety disorder, agoraphobia, social phobia, simple phobia, and obsessive–compulsive disorder were indicators of a latent anxiety disorder factor; marijuana dependence and alcohol dependence were indicators of a latent substance dependence factor; and conduct disorder was an indicator of a latent antisocial behavior factor. Judged in terms of its absolute goodness-of-fit (i.e., in terms of the absolute chi-square, GFI, AGFI, and RMSR values), the four-factor model fits well; however, it appears to be overparameterized. That is, it includes more parameters than are necessary to fit the correlations among the 10 disorders at age 18. This state of affairs is revealed clearly by the BIC statistic, which strongly prefers the two- to the four-factor model. What BIC reveals is that the drop in chi-square between the two- and four-factor models was well-offset by the increase in the number of parameters estimated.

Another way of asking if we need more than two distinct factors is to examine the correlations between the latent factors in the four-factor model. The correlation between the affective and anxiety factors is estimated at 1.0, and the correlation between the substance dependence and antisocial behavior factors is estimated at .89.[4] These high correlations are further evidence in favor of the two-factor model, in which the affective and anxiety factors were collapsed into a single internalizing factor, and the substance dependence and antisocial behavior factors were collapsed into a single externalizing factor.

Taken as a whole, these results indicate that a model positing two oblique latent factors provided the best fit to the diagnostic data at age 18. Thus, to create a final model for the 10 diagnoses at age 18, we refit the two-factor model, allowing for error in the conduct disorder indicator. Figure 1 shows this final model and its standardized parameter estimates. The model fits very well, $\chi^2(34, N = 930) = 35.42$, $p = .40$, GFI = .99, AGFI = .99, RMSR = .09, BIC = −196.98. Standard errors for estimated parameters ranged from .04 to .12, and *t* values for these parameters ranged from 2.73 to 12.10 (all significant at $p < .01$, two-tailed).

Mental Disorders at Age 21

Did internalizing and externalizing factors also underlie the 10 disorders at age 21? Table 1 shows the fit of the two-factor CFA model for the 10 disorders at age 21 (N = 937). The model again fits very well, yielding a small chi-square, a nearly perfect GFI and AGFI, and a relatively small RMSR. This model can again be compared with a one-factor model in which all 10 disorders are specified as indicators of a single latent factor; for the one-factor model, the fit declined substantially, as revealed

[4] The correlation of 1.0 between the affective and anxiety factors meant that the covariance matrix for the latent factors in the four-factor model at age 18 was not positive definite. This is further evidence that this model is overparameterized—that there is no need to split the internalizing factor into affective and anxiety subfactors (cf. Wothke, 1993).

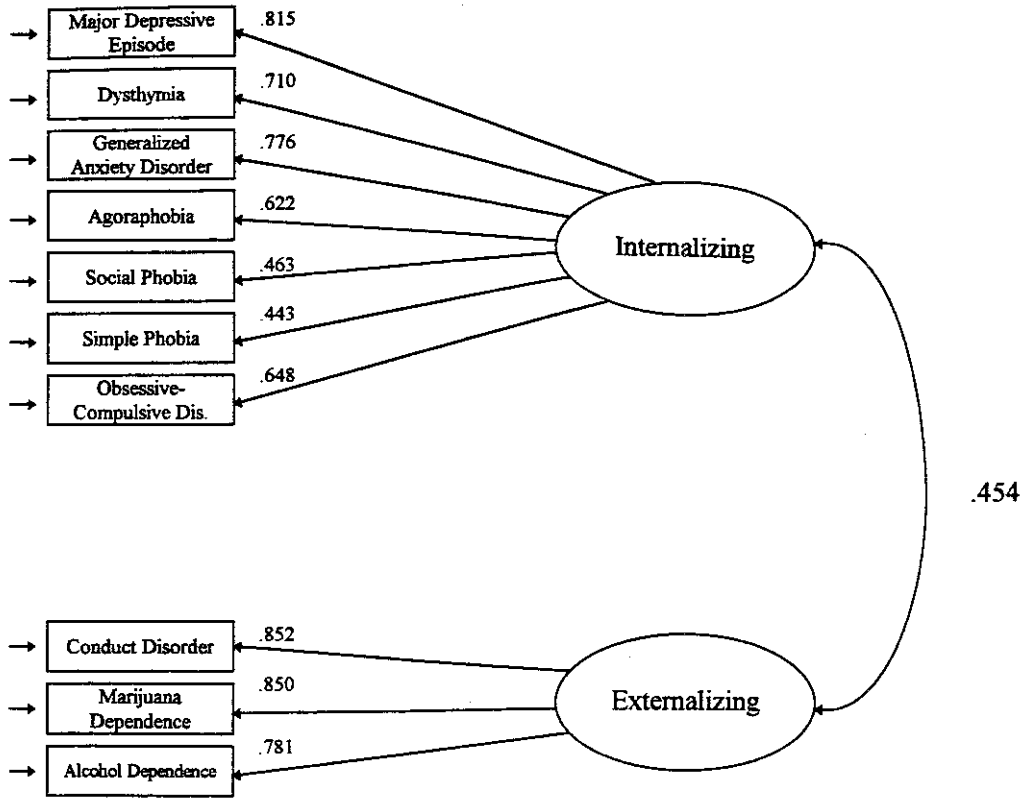

Figure 1. Fitted two-factor model of 10 *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.) mental disorders at age 18. All coefficients are standardized. Dis. = disorder.

by the larger chi-square and RMSR and the substantially less negative BIC score.

Do we need more than two factors to offer a good account of the correlations observed among the 10 disorders at age 21? Judged in terms of its absolute goodness-of-fit, the four-factor model at age 21 fits well (see Table 1); however, it again appears to be overparameterized. This is revealed by the BIC statistic, which strongly prefers the two- to the four-factor model at age 21, as was the case at age 18. The drop in chi-square between the two- and four-factor models was once again offset by the increase in the number of parameters estimated. The correlation between the affective and anxiety factors is estimated at .90, and the correlation between the substance dependence and antisocial behavior factors is estimated at .72. These high correlations are further evidence in favor of the two-factor model in the age-21 data.

Taken as a whole, these results indicate that a model positing two oblique latent factors provided the best fit to the diagnostic data at age 21. Thus, to create a final model for the 10 diagnoses at age 21, we refit the two-factor model, allowing for error in the antisocial personality disorder indicator. Figure 2 shows this final model and its standardized parameter estimates. The model fits very well, $\chi^2(34, N = 937) = 50.48$, $p = .03$, GFI = .99, AGFI = .99, RMSR = .10, BIC = −182.17. Standard errors for estimated parameters ranged from .04 to .20, and t values for these parameters ranged from 3.58 to 9.25 (all significant at $p < .01$, two-tailed).

Stability of Mental Disorder From Age 18 to Age 21

A two-factor model, positing the existence of internalizing and externalizing factors, appears to fit the data at both ages 18 and 21. Were these internalizing and externalizing factors stable from age 18 to age 21? To determine the stability of mental disorder from age 18 to age 21, we estimated both CFA models simultaneously and added a structural model positing paths linking the latent factors measured at both ages. This model can be seen in Figure 3 ($N = 882$). Initially, the model was estimated allowing for the two cross-lagged structural paths (i.e., paths from Internalizing 18 to Externalizing 21, and from Externalizing 18 to Internalizing 21) and all 10 potential autocorrelations (correlations between errors in the same diagnoses measured at both age 18 and age 21). For this initial model, $\chi^2(154, N = 882) = 170.87$, $p = .17$, GFI = .99, AGFI = .98, RMSR = .17, and BIC = −873.59. However, neither of the two cross-lagged paths were significant (for the path from Internalizing 18 to Externalizing 21, $t = -.89$; for the path from Externalizing 18 to Internalizing 21, $t = -1.20$), and 7 of the 10 autocorrelations were not significant (these 7 t values ranged from −.49 to +1.54). Hence, we trimmed these nonsignificant parameters from the model, producing the final model seen in Figure 3.

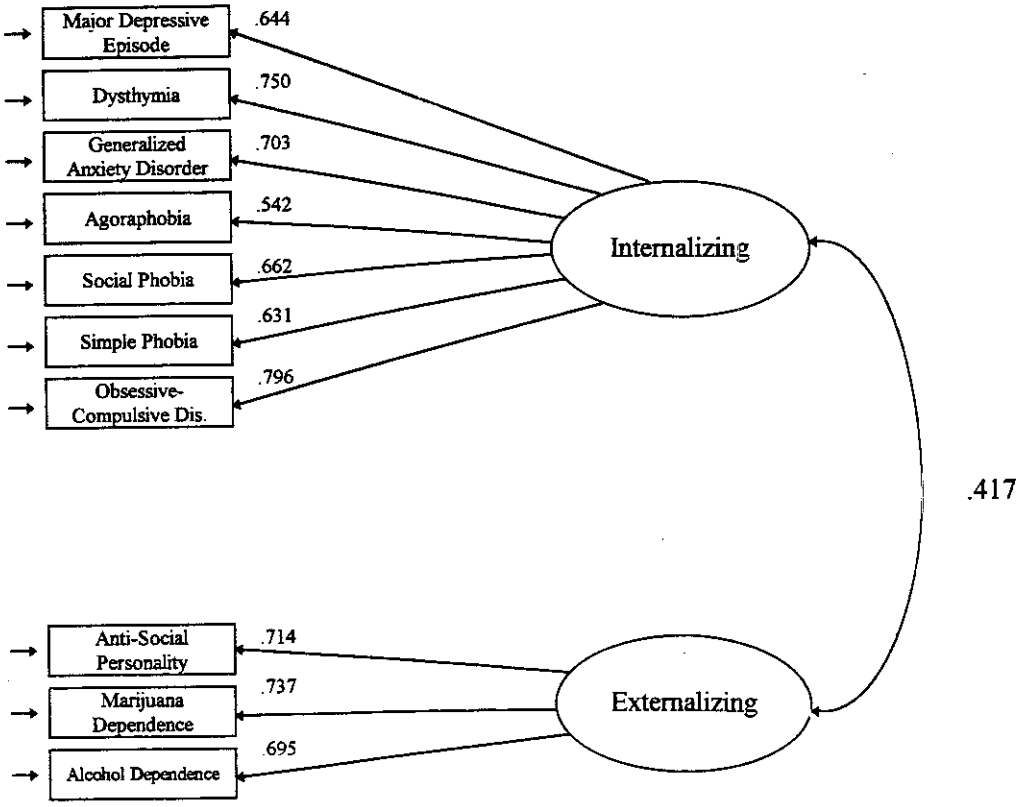

Figure 2. Fitted two-factor model of 10 *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.) mental disorders at age 21. All coefficients are standardized. Dis. = disorder.

The final model provides a good fit to the data, $\chi^2(163, N = 882) = 181.90$, $p = .15$; GFI = .98, AGFI = .98, RMSR = .17, and BIC = −923.60. Standard errors for estimated parameters ranged from .03 to .14, and t values for these parameters ranged from 2.20 to 12.34 (all significant at $p < .05$, two-tailed). The path linking internalizing disorder at age 18 with internalizing disorder at age 21 is estimated at .69, and the path linking externalizing disorder at age 18 with externalizing disorder at age 21 is estimated at .86. To determine if these paths were significantly different from one another, we reran the model in Figure 3, forcing both paths to be equal. This produced a $\chi^2(164, N = 882) = 188.98$, $p = .09$, GFI = .98, AGFI = .98, RMSR = .17, BIC = −923.30. The chi-square value is slightly larger than the value for the original, unconstrained model, $\chi^2(1, N = 882) = 7.08$, $p < .01$. This means that the paths linking the disorder factors at age 18 with their counterpart disorder factors at age 21 are statistically different: The externalizing disorder factor is slightly more stable than the internalizing disorder factor. Nonetheless, both paths are large in absolute magnitude, indicating substantial stability in the rank ordering of our research participants on both factors across the 3-year period from age 18 to age 21.

Summary

The results confirmed the presence of an oblique, two-factor structure underlying 10 *DSM–III–R* disorders assessed independently at two points in time in our sample. In addition, the results from the structural component of the stability model (the two middle arrows in Figure 3) indicated that, to a substantial extent, persons in our sample retained their relative positions on the two latent factors from age 18 to age 21, with externalizing disorders demonstrating significantly greater stability than internalizing disorders.

Discussion

This article investigated the factor structure underlying common *DSM–III–R* mental disorders in an effort to "elucidate the broad, higher-order structure of phenotypic psychopathology" (Clark et al., 1995, p. 131). A two-factor structure, consisting of internalizing and externalizing factors, offered the best account of the correlations observed among 10 common disorders taken separately at two different ages (18 and 21); the two-factor model was found to be superior to both one- and four-factor models. In addition, the findings indicated substantial differential stability (i.e., preservation of our research participants' rank orders; Caspi & Bem, 1990) across the 3-year period from age 18 to age 21.

Despite the strengths of the current study (i.e., the large, representative sample and the use of a well-known, standardized diagnostic interview), three specific shortcomings should be noted. First, we focused on a limited number of *DSM* disorders. We chose these specific disorders because they had high preva-

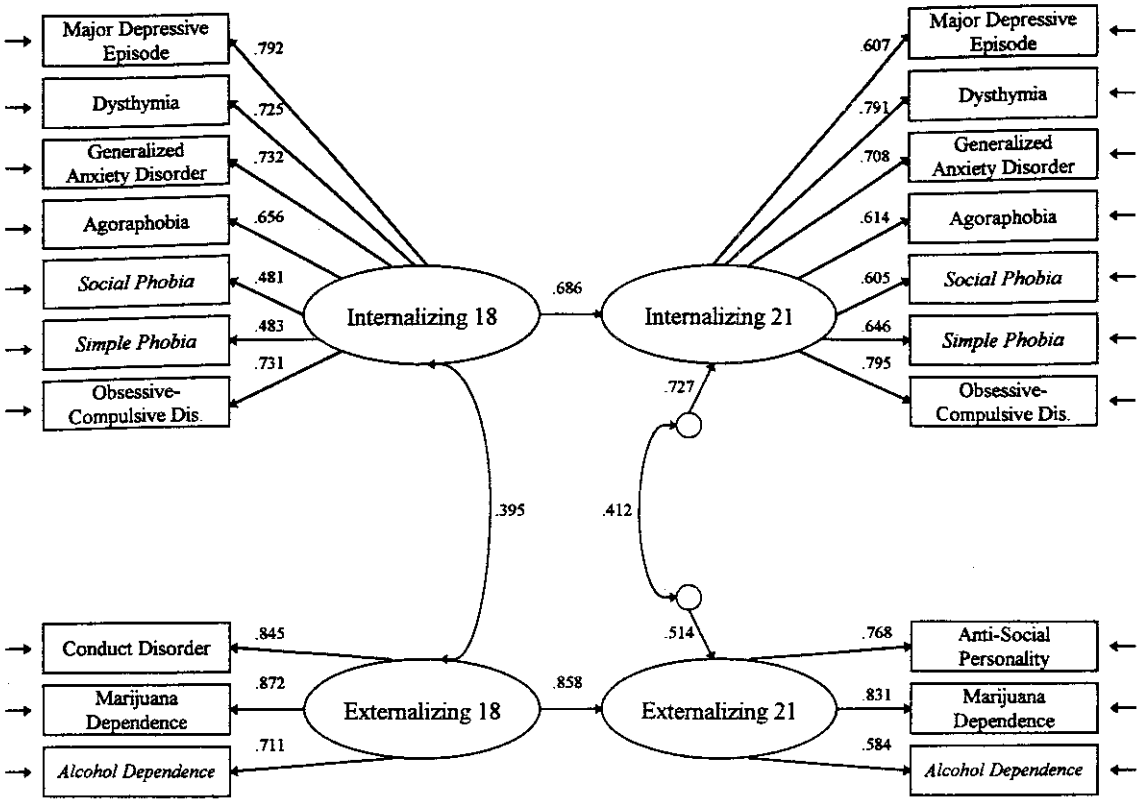

Figure 3. Fitted structural model of the stability of two factors underlying 10 *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.) mental disorders, measured separately at ages 18 and 21. All coefficients are standardized. Errors in indicators with italicized names (*Social Phobia, Simple Phobia,* and *Alcohol Dependence*) were allowed to covary across the age 18 and age 21 assessments. The small circles just beneath Internalizing 21 and just above Externalizing 21 represent disturbance terms in the structural equations. Dis. = disorder.

lence rates in our general population sample. This allowed for the accurate estimation of the correlations among the disorders in the general population. Nevertheless, our epidemiological sampling strategy limited our coverage of the broader psychopathology domain that may be covered by clinical samples. Future research must determine if the two-factor model can incorporate additional *DSM* disorders (e.g., anorexia nervosa and panic disorder) that may appear at higher rates in clinical samples.

In addition, our discovery of a two-factor structure underlying 10 *DSM* disorders does not preclude the possibility that these two factors are constituted of a number of lower order symptom facets. That is, our results are not incompatible with the possibility that a more differentiated (albeit oblique) set of intermediary factors lies between measures of specific symptoms and the two higher order factors delineated here. Future research must determine if the two-factor model can accommodate symptom-level data or if such data would be better accommodated by a model in which the symptoms of common mental disorders constitute a first-order set of factors, the intercorrelations of which are, in turn, accounted for by second-order internalizing and externalizing factors.

Second, the age range of our sample was restricted to late adolescence and early adulthood. Future research must determine if the two-factor model will also hold in older age groups, in which the distribution of diagnoses may be different (e.g., in samples in which the participants have passed through the age of risk for other major mental disorders of adulthood, such as the adult psychoses). Third, our assessments of mental disorder were made using information obtained only through self-report interviews. Although this is the standard methodology for obtaining diagnoses of mental disorder, it is well-known that monomethod assessments confound method variance and true-score variance. Future research must ascertain the robustness of the two-factor structure when data regarding mental disorders are obtained using other methods (e.g., biochemical assays) and sources (e.g., informant reports).

Despite these concerns, our findings offer potential clarification of two robust empirical phenomena that have been observed in large-scale epidemiological studies of mental disorder: comorbidity and the positive association between comorbidity and disorder severity (Clark et al., 1995; Kessler et al., 1994; Robins & Regier, 1991). In our study, *DSM–III–R* disorders were successfully modeled as indicators of underlying factors. If a number of *DSM–III–R* disorders are all valid indicators of a given latent dimension, then the presence of comorbidity (understood as covariance) among these disorders should be observed;

from this perspective, comorbidity is not a nuisance, it is a phenomenon that follows logically from the existence of latent continuous structures underlying multiple manifest diagnoses (cf. Maser & Cloninger, 1990). Similarly, the CFA model predicts the observed positive association between severity and comorbidity: If two or more disorders are conceptualized as valid indicators of an underlying dimension, cases that are more likely to meet the criteria for one of these disorders (because of severe symptomatology) should also be more likely to meet the criteria for another of these disorders. Our results may be further understood by considering the nature of internalizing and externalizing patterns of behavior and by considering the implications of our findings for research on the predictors and correlates of psychopathology.

What Are Internalizing and Externalizing?

One possible interpretation of the stable internalizing and externalizing factors identified in this study involves viewing internalizing and externalizing as basic orientations toward the world. From this viewpoint, many disorders described in the categorical neo-Kraepelinian tradition are indicators of more essential difficulties—difficulties that give rise to the disordered states well operationalized in recent *DSM* manuals and in clinical interviews, such as the DIS. This idea has its roots in the writings of Karen Horney (1945), who postulated the existence of a basic anxiety in all persons that can be modulated in various ways. One form of modulation, moving away from the world, is conceptually similar to the internalizing factor identified in our study. Each internalizing disorder studied here contains an element of withdrawal from the external world—whether into the negative, self-referential thought patterns of a major depressive episode (Haaga, Dyck, & Ernst, 1991) or away from the world entirely, as in agoraphobia. Similarly, another form of modulation, moving against the world, parallels our externalizing factor. Each externalizing disorder we studied places the individual at odds with society, whether in the criminal behavior and disregard for others that accompany antisocial personality disorder (American Psychiatric Association, 1994) or in the criminal lifestyle that often accompanies substance dependence (Zucker & Gomberg, 1986).

In our study, these styles showed impressive stability across the 3-year period from age 18 to age 21: The internalizing factor at age 18 was correlated .69 with the internalizing factor at 21, and the externalizing factor at age 18 was correlated .86 with the externalizing factor at age 21. In addition to this 3-year stability in the psychopathology domain, internalizing and externalizing styles show coherence across multiple contexts and across longer periods in the life course. For example, Caspi, Elder, and Bem (1987, 1988) showed that boys who were moving away from the world at ages 8, 9, and 10 (by being shy and retiring) were slow to fully enter the adult roles of husband and father, and were slow to enter into a stable career. In addition, boys who were moving against the world at ages 8, 9, and 10 (by engaging in frequent temper tantrums) became men who experienced considerable occupational instability in their adult lives. If further developmental research shows that these two styles represent core psychopathological processes that are stable and consistent, psychopathology research may profit from focusing attention on the origins of these core processes, in addition to studying their varied manifestations (e.g., as *DSM* disorders).

Our internalizing and externalizing disorder factors may also map onto higher order trait dimensions of adult personality. Specifically, both internalizing (anxiety and affective) and externalizing (substance dependence and antisocial) disorders have been associated with high neuroticism–negative emotionality, whereas externalizing disorders have been uniquely associated with low conscientiousness–constraint (Clark, Watson, & Mineka, 1994; Krueger et al., 1996; Sher & Trull, 1994; Trull & Sher, 1994). These patterns of relations suggest that internalizing and externalizing may correspond, respectively, with neuroticism–negative emotionality and conscientiousness–constraint. One strategy for evaluating this speculation directly would involve examining personality and psychopathology variables jointly in the same factor analyses. A joint factor analysis may reveal distinct personality and psychopathology dimensions, or it may reveal that personality and psychopathology indicators, are, in fact, measuring the same latent variables. A second strategy would use behavior genetic methods to determine the extent to which the etiologies of both normal and abnormal variation (i.e., personality and psychopathology) overlap. For example, Kendler, Neale, Kessler, Heath, and Eaves (1993) showed that most of the correlation they observed between the personality trait of neuroticism and the liability to a diagnosis of major depression was attributable to genetic factors involved in both neuroticism and major depression.

Implications for Research on Psychopathology

Many predictors of psychopathology are relatively nonspecific; this lack of specificity may reflect the oblique, two-factor structure that appears to underlie affective, anxiety, substance dependence, and antisocial behavior disorders. For example, the developmental antecedents of children's antisocial and anxious–depressive disorders, although not identical, are more similar than they are different (e.g., Henry, Moffitt, Robins, Earls, & Silva, 1993; White, Moffitt, Earls, Robins, & Silva, 1990). Two possible implications can be drawn from this observation: First, our inability to achieve greater predictive specificity results from imprecision in our current nosological systems, or, second, our inability to achieve greater predictive specificity results from the involvement of common factors in apparently different varieties of mental disorder.

The divergent research implications of these two perspectives are not trivial. Consider, for example, the now extensive literature on the comorbidity of anxiety and depression (Kendall & Watson, 1989; Maser & Cloninger, 1990; Mineka, Watson, & Clark, in press). Although it is possible to achieve some separation of the two syndromes by emphasizing their differences (Watson, Weber, et al., 1995), the largest portion of variance in both syndromes is shared and is not unique (Watson, Clark, et al., 1995). Should research in this and other areas of research in which comorbidity has captured extensive attention (e.g., the comorbidity of attention deficit and conduct disorder in children and adolescents; Fergusson, Horwood, & Lloyd, 1991) focus on achieving better separation of putative disorders, or should the reasons for disorder covariance be more extensively explored?

The current study suggests that the latter perspective deserves serious consideration. Perhaps the predictors of psychopathology are relatively nonspecific because they, like the disorders they predict, are reflective of certain core psychopathological processes. The implication that follows is that research on the correlates and predictors of psychopathology may benefit from seeking the most robust predictors of multiple disorders (i.e., predictors that correlate most highly with variables such as our internalizing and externalizing factors), as opposed to seeking variables that predict disorder X better than they predict disorder Y. Robust predictors of multiple disorders may provide clues about the nature of hypothetical core psychopathological processes (such as the internalizing and externalizing interactional styles outlined above) that may underlie multiple mental disorders.

Along these same lines, our study places emphasis on the potential advantages of continuous approaches to the measurement and classification of mental disorders.[5] Various writers (e.g., Lewin, 1935) have noted the conceptual advances that continuous variables have presaged in other sciences. A recent example can be found in the field of molecular genetics, where the older one-gene/one-disorder hypothesis is gradually giving way to a quantitative trait loci conception, in which multiple genes, none of which are singularly necessary and sufficient for the development of a disorder, are regarded as contributing additively and interchangeably to determine an individual's placement on a continuum of vulnerability (Plomin, 1995). Psychopathology research may benefit from following this example and from considering carefully the conceptual and empirical advantages offered by a continuous approach to the classification and study of mental disorders. As noted by Clark et al. (1995) with regard to mental disorders, "it is time to halt the general call for dimensional systems and to begin the hard work of developing specific dimensional proposals in targeted domains" (p. 147). We hope that our work will be regarded as a modest first step in this direction.

[5] In placing this emphasis, we do not mean to suggest that the continuous model is optimal for all psychopathologies (cf. Meehl, 1992). For example, recent work has suggested the presence of discontinuities in distributions of schizotypal signs (Korfine & Lenzenweger, 1995; Lenzenweger & Korfine, 1992), suggesting the possibility that psychosis proneness, as just one example, may be accurately characterized as a latent dichotomy.

References

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.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Blashfield, R. K. (1984). *The classification of psychopathology: Neo-Kraepelinian and quantitative approaches.* New York: Plenum.

Caspi, A., & Bem, D. J. (1990). Personality continuity and change across the life course. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 549–575). New York: Guilford Press.

Caspi, A., Elder, G. H., & Bem, D. J. (1987). Moving against the world: Life-course patterns of explosive children. *Developmental Psychology, 23,* 308–313.

Caspi, A., Elder, G. H., & Bem, D. J. (1988). Moving away from the world: Life-course patterns of shy children. *Developmental Psychology, 24,* 824–831.

Cattell, R. B. (1965). *The scientific analysis of personality.* Chicago: Aldine.

Clark, L. A., Livesley, W. J., Schroeder, M. L., & Irish, S. L. (1996). Convergence of two systems for assessing specific traits of personality disorder. *Psychological Assessment, 8,* 294–303.

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.

Costa, P. T., Jr., & Widiger, T. (Eds.). (1994). *Personality disorders and the five-factor model of personality.* Washington, DC: American Psychological Association.

Eysenck, H. J. (1960). *The structure of human personality.* London: Routledge & Kegan Paul.

Feehan, M., McGee, R., Nada Raja, S., & Williams, S. M. (1994). *DSM–III–R* disorders in New Zealand 18-year-olds. *Australian and New Zealand Journal of Psychiatry, 28,* 87–99.

Feinstein, A. R. (1970). The pre-therapeutic classification of co-morbidity in chronic disease. *Journal of Chronic Diseases, 23,* 455–468.

Fergusson, D. M., Horwood, L. J., & Lloyd, M. (1991). Confirmatory factor models of attention deficit and conduct disorder. *Journal of Child Psychology and Psychiatry, 32,* 257–274.

Fergusson, D. M., Horwood, L. J., & Lynskey, M. T. (1995). The stability of disruptive childhood behaviors. *Journal of Abnormal Child Psychology, 23,* 379–396.

Graham, J. R. (1993). *MMPI–2: Assessing personality and psychopathology* (2nd ed.). New York: Oxford University Press.

Haaga, D. A. F., Dyck, M. J., & Ernst, D. (1991). Empirical status of cognitive theory of depression. *Psychological Bulletin, 110,* 215–236.

Henry, B., Moffitt, T., Robins, L. N., Earls, F., & Silva, P. (1993). Early family predictors of child and adolescent antisocial behaviour: Who are the mothers of delinquents? *Criminal Behaviour and Mental Health, 3,* 97–118.

Horney, K. (1945). *Our inner conflicts.* New York: Norton.

Jensen, A. R. (1980). *Bias in mental testing.* New York: Free Press.

Jöreskog, K. G., & Sörbom, D. (1988). PRELIS: A program for multivariate data screening and data summarization. (Version 1.20) [Computer software]. Chicago: Scientific Software.

Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7: A guide to the program and applications.* Chicago: SPSS.

Kendall, P. C., & Watson, D. (Eds.). (1989). *Anxiety and depression: Distinctive and overlapping features.* San Diego, CA: Academic Press.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1993). A longitudinal twin study of personality and major depression in women. *Archives of General Psychiatry, 50,* 853–862.

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 Study. *Archives of General Psychiatry, 51,* 8–19.

Korfine, L., & Lenzenweger, M. F. (1995). The taxonicity of schizotypy: A replication. *Journal of Abnormal Psychology, 104,* 26–31.

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.

Leaf, P. J., Myers, J. K., & McEvoy, L. T. (1991). Procedures used in the epidemiologic catchment area study. In L. N. Robins & D. A. Regier (Eds.), *Psychiatric disorders in America* (pp. 11–32). New York: Free Press.

Lenzenweger, M. F., & Korfine, L. (1992). Confirming the latent structure and base rate of schizotypy: A taxometric analysis. *Journal of Abnormal Psychology, 101,* 567–571.

Lewin, K. (1935). The conflict between Aristotelian and Galilean modes of thought in contemporary psychology. In *A dynamic theory of personality: Selected papers* (pp. 1–42). New York: McGraw-Hill.

Maser, J. D., & Cloninger, C. R. (1990). *Comorbidity of mood and anxiety disorders.* Washington, DC: American Psychiatric Association Press.

Mednick, S. A. (1978). Berkson's fallacy and high-risk research. In L. C. Wynne (Ed.), *The nature of schizophrenia* (pp. 442–452). New York: Wiley.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology, 46,* 806–834.

Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality, 60,* 117–174.

Menninger, K. (1963). *The vital balance.* New York: Viking Press.

Millon, T. (1991). Classification in psychopathology: Rationale, alternatives, and standards. *Journal of Abnormal Psychology, 100,* 245–261.

Mineka, S., Watson, D., & Clark, L. A. (in press). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology.*

Newman, D. L., Moffitt, T. E., Caspi, A., Magdol, L., Silva, P. A., & Stanton, W. R. (1996). Psychiatric disorder in a birth cohort of young adults: Prevalence, co-morbidity, clinical significance, and new case incidence from age 11 to 21. *Journal of Consulting and Clinical Psychology, 64,* 552–562.

Oatley, K. (1988). Life events, social cognition, and depression. In S. Fisher & J. Reason (Eds.), *Handbook of life stress, cognition, and health* (pp. 543–557). New York: Wiley.

Ollendick, T. H., & King, N. J. (1994). Diagnosis, assessment, and treatment of internalizing problems in children: The role of longitudinal data. *Journal of Consulting and Clinical Psychology, 62,* 918–927.

Plomin, R. (1995). Molecular genetics and psychology. *Current Directions in Psychological Science, 4,* 114–117.

Quay, H. C. (1986). Classification. In H. C. Quay & J. S. Werry (Eds.), *Psychopathological disorders of childhood* (pp. 1–34). New York: Wiley.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25,* 111–163.

Regier, D. A., Myers, J. K., Kramer, M., Robins, L. N., Blazer, D. G., Hough, R. L., Eaton, W. W., & Locke, B. Z. (1984). The NIMH epidemiologic catchment area program: Historical context, major objectives, and study population characteristics. *Archives of General Psychiatry, 41,* 934–941.

Robins, L. N., Helzer, J. E., Cottler, L., & Goldring, E. (1989). *Diagnostic Interview Schedule, Version III–R.* Unpublished manuscript, Washington University, St. Louis, MO.

Robins, L. N., & Regier, D. A. (1991). *Psychiatric disorders in America.* New York: Free Press.

Sher, K. J., & Trull, T. J. (1994). Personality and disinhibitory psychopathology: Alcoholism and antisocial personality disorder. *Journal of Abnormal Psychology, 103,* 92–102.

Silva, P. A. (1990). The Dunedin Multidisciplinary Health and Development Study: A fifteen year longitudinal study. *Paediatric and Perinatal Epidemiology, 4,* 96–127.

Skinner, H. A. (1981). Toward the integration of classification theory and methods. *Journal of Abnormal Psychology, 90,* 68–87.

Strauss, M. E. (1995). Editorial. *Journal of Abnormal Psychology, 104,* 555–557.

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.

Watson, D., Clark, L. A., & Harkness, A. R. (1994). Structures of personality and their relevance to psychopathology. *Journal of Abnormal Psychology, 103,* 18–31.

Watson, D., Clark, L. A., Weber, K., Assenheimer, J. S., Strauss, M. E., & McCormick, R. A. (1995). Testing a tripartite model: II. Exploring the symptom structure of anxiety and depression in student, adult, and patient samples. *Journal of Abnormal Psychology, 104,* 15–25.

Watson, D., Weber, K., Assenheimer, J. S., Clark, L. A., Strauss, M. E., & McCormick, R. A. (1995). Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. *Journal of Abnormal Psychology, 104,* 3–14.

White, J. L., Moffitt, T. E., Earls, F., Robins, L. N., & Silva, P. A. (1990). How early can we tell? Predictors of childhood conduct disorder and adolescent delinquency. *Criminology, 28,* 507–533.

Wothke, W. (1993). Nonpositive definite matrices in structural modeling. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 256–293). Newbury Park, CA: Sage.

Zucker, R. A., & Gomberg, E. S. L. (1986). Etiology of alcoholism reconsidered: The case for a biopsychosocial process. *American Psychologist, 41,* 783–793.

Received January 23, 1996
Revision received December 12, 1996
Accepted September 3, 1997 ■