

Comparison of the MMPI–2 Personality Psychopathology Five (PSY-5), the NEO-PI, and the NEO-PI-R

Timothy J. Trull and J. David Useda
University of Missouri—Columbia

Paul T. Costa, Jr., and Robert R. McCrae
National Institute on Aging

This study examined relations between Minnesota Multiphasic Personality Inventory Personality Psychopathology Five (PSY-5; A. R. Harkness, J. L. McNulty, & Y. S. Ben-Porath, 1995), NEO Personality Inventory (NEO-PI; P. T. Costa & R. R. McCrae, 1985), and the revised NEO-PI (NEO-PI-R; P. T. Costa & R. R. McCrae, 1992b) scores in community ($N = 170$) and clinical ($N = 57$) samples. In the clinical sample, the temporal stability of the scales and their associations with personality disorder symptom counts were also assessed. Correlations between the two instruments demonstrated meaningful relations between the two sets of constructs in both samples. Both instruments showed substantial stability over 6 months, and both were significant and substantial predictors of symptom counts for most personality disorders. The data support the reinterpretation of personality disorders in terms of underlying dimensions of personality.

The broad factors of Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness appear to explain most of the common variance among normal personality traits (Digman, 1990), and this Five-Factor Model (FFM) has also been used to understand clinical assessment and psychopathology (Costa & McCrae, 1992a; Costa & Widiger, 1994). The most obvious application of the FFM to psychopathology constructs is to the personality disorders listed in the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)* (American Psychiatric Association, 1994). Several studies have empirically demonstrated the relevance of the FFM to personality disorder constructs in both clinical (Soldz, Budman, Demby, & Merry, 1993; Trull, 1992) and nonclinical (Costa & McCrae, 1990; Wiggins & Pincus, 1989) samples. Further, the FFM is quickly gaining prominence as a dimensional alternative to the categorical psychiatric diagnostic system (Costa & Widiger, 1994).

Recently, A. R. Harkness, McNulty, and Ben-Porath (1995) reported on the development of a set of Minnesota Multiphasic Personality Inventory—2 (MMPI–2) scales (the Personality

Psychopathology Five; PSY-5) to measure five personality constructs that they believe are relevant to psychopathology. These constructs emerged from structural analyses of the semantic similarities between topics covered in the revised third edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R)* personality disorders and normal personality trait terms (A. R. Harkness, 1992; A. R. Harkness & McNulty, 1994). The PSY-5 scales were labeled as follows: Aggressiveness, Psychoticism, Constraint, Negative Emotionality, and Positive Emotionality. A. R. Harkness et al. (1995) noted a number of similarities and differences between the PSY-5 constructs and those represented in the FFM. As for similarities, both the PSY-5 and measures of the FFM (e.g., the Revised NEO Personality Inventory; Costa & McCrae, 1992b) include scales to measure dispositions toward the experience of negative affectivity/emotionality as well as positive emotionality and social engagement. Thus, the PSY-5 constructs of Negative Emotionality and Positive Emotionality are similar to their FFM counterparts of Neuroticism and Extraversion. However, A. R. Harkness et al. (1995) argued that PSY-5 Aggressiveness, Constraint, and Psychoticism constructs do not have direct FFM counterparts. First, PSY-5 Aggressiveness was said to tap more extreme traits related to low Agreeableness, but low PSY-5 Aggressiveness is not indicative of high Agreeableness. Second, low levels of PSY-5 Constraint (impulsiveness and risk-taking) do not correspond to low Conscientiousness (disorganization and carelessness). And, finally, PSY-5 Psychoticism has no FFM counterpart according to Harkness et al. These propositions have not been subjected to direct empirical test.

A. R. Harkness et al. (1995) argued that the PSY-5 scales should be more relevant to personality disorder pathology than are FFM scales such as the NEO Personality Inventory (NEO-PI). As A. R. Harkness and McNulty (1994) noted, the FFM originated in analyses of lay trait terms taken from the dictionary, whereas the PSY-5 constructs were derived from an analysis of personality disorder symptoms. Further, the PSY-5 scales were created from the items of the MMPI–2 (Butcher, Dahl-

Timothy J. Trull and J. David Useda, Department of Psychology, University of Missouri—Columbia; Paul T. Costa, Jr., and Robert R. McCrae, Gerontology Research Center, National Institute on Aging, Baltimore, Maryland.

Study 2 was supported by National Institutes of Health Behavioral Research Support Grant RR 07053 and by a Summer Research Fellowship from the University of Missouri—Columbia. Portions of this article were reported at the Midwinter Meeting of the Society for Personality Assessment, April 1994, in Chicago. Study 1 contains new analyses of data previously analyzed in Costa, Busch, Zonderman, and McCrae (1986) and Costa and McCrae (1992b). Study 2 contains new analyses of data previously analyzed in Trull (1992). We are grateful to Angela Z. Vieth for her comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Timothy J. Trull, Department of Psychology, University of Missouri, 210 McAlester Hall, Columbia, Missouri 65211. Electronic mail may be sent via Internet to psytrull@mizzou1.missouri.edu.

strom, Graham, Tellegen, & Kaemmer, 1989), which was designed as a measure of psychopathology. Given these differences, it would not be surprising if the PSY-5 scales were better predictors of personality pathology. However, no direct comparison between the usefulness of the PSY-5 and NEO-PI in the assessment of personality disorders has been conducted.

In this article, we present results from two studies that examined the relationship between the MMPI-2 PSY-5 scales and the NEO-PI and Revised NEO-PI (NEO-PI-R). Study 1 presents data obtained from a community-based nonclinical sample, and Study 2 presents data obtained from a clinical sample of psychiatric outpatients. Collectively, these two studies addressed the following issues: (a) What are the relations between PSY-5 scores and NEO-PI-R scores? (b) How stable are mean levels and individual differences for the two sets of scales among individuals undergoing psychotherapy? (c) Do PSY-5 or NEO-PI scale scores account for a significant amount of variance in personality disorder scores (assessed via interview and self-report), and which individual PSY-5 or NEO-PI scales are significant correlates of the personality disorder constructs?

Study 1

Method

The purpose of the first study was to assess the relationship between PSY-5 scales and the factors and facets of the NEO-PI-R (Costa & McCrae, 1992b) in order to evaluate the PSY-5 constructs in terms of the FFM of personality. Participants were 111 men and 59 women in the Baltimore Longitudinal Study of Aging (BLSA; Shock et al., 1984) ranging in age from 20 to 84. The mean age of the sample at administration of the MMPI was 61.1 ($SD = 14.9$). BLSA participants were generally healthy and well-educated community-dwelling volunteers. Mean years of education were 16.6 ($SD = 2.3$), and 80% of the sample had college or advanced degrees. Although the BLSA currently has 11% minority participation, at the time that these data were collected, over 96% of the sample was White.

The PSY-5 scales were rationally constructed from the MMPI-2 item pool via a series of procedures that included the nomination of items to measure the PSY-5 constructs by trained item selectors, as well as the deletion of items in which keying (true vs. false) was ambiguous or of items that were potentially confusing because of the wording. The final PSY-5 scales contain no overlapping items, and total scale length ranges from 18 to 34 items. A. R. Harkness et al. (1995) reported internal consistency coefficients for the PSY-5 scales that ranged from .68 to .88 across clinical and collegiate samples, and preliminary support for the construct validity of these scales was also provided.

In Study 1, slightly shortened versions of the PSY-5 scales were scored from the original version of the MMPI; a total of 119 of the 139 items (86%) were used. Specifically, 15 of the 18 Aggressiveness items, 20 of the 25 Psychoticism items, 27 of the 29 Constraint items, 27 of the 33 Negative Emotionality items, and 30 of the 34 Positive Emotionality items were used in Study 1. The internal consistencies for these scales in this sample were .62 (Aggressiveness), .50 (Psychoticism), .72 (Constraint), .81 (Negative Emotionality), and .74 (Positive Emotionality).¹ The MMPI was completed at the Gerontology Research Center in Baltimore between 1981 and 1985 (Costa, Busch, Zonderman, & McCrae, 1986).

The FFM was assessed by the NEO-PI-R (Costa & McCrae, 1992b). This is a 240-item questionnaire that measures six specific traits or facets for each of the five domains. The instrument is published in parallel forms for self-reports (Form S) and observer ratings (Form R). Internal consistencies for the 8-item facet scales range from .56 to .81; for the 48-item

domain scales, they range from .86 to .92. Orthogonal factor scores, derived from a weighted combination of facet scale scores standardized within gender, are recommended as the best operationalization of the FFM (Costa & McCrae, 1992b). In nonclinical adult samples, NEO-PI-R domain scores are stable over time, with uncorrected 7-year retest correlations ranging from .63 to .81 in peer ratings (Costa & McCrae, 1992c). In the present study, the instrument was completed in two parts in 1986 and 1990; both sections were mailed to participants and completed at home. In addition to self-reports, one or more peer ratings were also available on a subsample of 55 participants (Costa & McCrae, 1992c).

Results and Discussion

Table 1 presents zero-order correlations between the PSY-5 scale scores and the NEO-PI-R factor and facet scores.² As indicated, NEO-PI-R Neuroticism and Extraversion scores were highly correlated with PSY-5 Negative Emotionality and Positive Emotionality scores, respectively. Further, all the facets of these respective NEO-PI-R factors correlated significantly with their PSY-5 counterparts. PSY-5 Aggressiveness appears to tap both Extraversion (Gregariousness, Assertiveness, Activity, and Excitement-Seeking facets) as well as low Agreeableness (low scores on Straightforwardness, Compliance, and Modesty facets). PSY-5 Constraint correlated positively with NEO-PI-R Agreeableness (Trust, Straightforwardness, Compliance, and Modesty) and Conscientiousness (Dutifulness and Deliberation). Constraint was negatively correlated with Extraversion (Excitement-Seeking and Positive Emotions) and Openness (Fantasy, Feelings, Actions, and Values). Finally, PSY-5 Psychoticism was positively correlated with NEO-PI-R Neuroticism (Anxiety, Angry Hostility, Depression, Self-Consciousness, and Vulnerability). It was also positively related to Openness to Fantasy and Feelings, and negatively related to Trust, Straightforwardness, Altruism, and Self-Discipline, although all these correlations were small in magnitude. Many of the findings listed above were replicated using single peer ratings on the NEO-PI-R (see Table 1); correlations between the two instruments thus cannot be attributed wholly to self-presentation or other artifacts of a shared self-report method.

The correlations in Table 1 are probably attenuated to some extent by the lengthy time interval between administration of the two instruments. In view of this, the magnitude of correlations is noteworthy and adds to the growing body of literature showing that measures of psychopathology are closely related to measures of normal personality traits (Costa & McCrae, 1992a). The pattern of correlates tends to support the construct validity of the

¹ To estimate the relationship between the short and full versions of the PSY-5 scales, we calculated zero-order correlations between the two versions for each PSY-5 construct at each of the three assessment occasions in our clinical sample (Study 2). The correlation coefficients were uniformly high, ranging from .93 to .99. Further, we calculated internal consistency coefficients for both the short and full versions of each PSY-5 scale in the clinical sample. In each case, the internal consistency coefficients for the short and full versions were quite comparable: .73 versus .73 for Positive Emotionality, .57 versus .58 for Aggressiveness, .72 versus .72 for Constraint, .83 versus .85 for Negative Emotionality, and .75 versus .81 for Psychoticism.

² The full matrix of intercorrelations among all these variables is available from Timothy J. Trull.

Table 1
Correlations Between Personality Psychopathology Five Scales and Revised NEO Personality Inventory (NEO-PI-R) Factors and Facet Scales

NEO-PI-R scale	Personality Psychopathology Five scale				
	POSEMO	AGGRES	CONSTR	NEGEMO	PSYCHOT
Factors					
Neuroticism	-.21**	-.10	.01	.60****	.30***
Extraversion	.59****	.29****	-.31***	.08	-.06
Openness	.12	.03	-.18**	-.09	.11
Agreeableness	-.12	-.32****	.32****	-.13	-.12
Conscientiousness	-.01	.07	.27****	-.22**	-.06
Neuroticism facets					
Anxiety	-.25**	-.16*	.06	.60****	.35***
Angry Hostility	-.11	.16**	-.13	.59****	.16*
Depression	-.18*	-.08	.02	.51***	.20**
Self-Consciousness	-.21**	-.22**	.00	.42***	.17*
Impulsiveness	.00	.04	-.27****	.37****	.12
Vulnerability	-.23**	-.22**	.00	.44***	.21**
Extraversion facets					
Warmth	.40***	.05	.00	-.11	-.08
Gregariousness	.52****	.17*	-.12	-.10	-.12
Assertiveness	.31***	.35****	-.08	-.13	-.10
Activity	.36***	.32***	-.13	-.08	-.08
Excitement Seeking	.45****	.19**	-.40****	.15	.13
Positive Emotions	.41***	.14	-.27****	-.04	-.01
Openness facets					
Fantasy	.12	.02	-.32****	.17*	.23***
Aesthetics	.14	.09	-.10	-.18*	.04
Feelings	.23**	.16	-.20**	.15*	.18*
Actions	.29***	.03	-.18*	-.18*	-.06
Ideas	.12	.00	-.05	-.10	.06
Values	.07	-.02	-.31****	-.05	.04
Agreeableness facets					
Trust	.09	-.05	.17**	-.24**	-.20*
Straightforwardness	-.19**	-.21***	.35****	-.11	-.20****
Altruism	.13	-.07	.13	-.14	-.19*
Compliance	-.04	-.39****	.26****	-.25**	-.10
Modesty	-.10	-.19**	.17**	.04	.04
Tender-Mindedness	-.03	-.02	.05	.12	-.03
Conscientiousness facets					
Competence	.17*	.16*	.12	-.34***	-.11
Order	.12	.03	.04	-.27***	.02
Dutifulness	-.06	-.04	.30****	-.28****	-.14
Achievement Striving	.16*	.14	.12	-.19*	-.07
Self-Discipline	.11	.08	.14	-.36***	-.21**
Deliberation	-.23***	-.17*	.36****	-.23**	-.03

Note. $N = 170$. POSEMO = Positive Emotionality and Extraversion; AGGRES = Aggressiveness; CONSTR = Constraint; NEGEMO = Negative Emotionality and Neuroticism; PSYCHOT = Psychoticism.

* Replicated association ($p < .05$) using single peer ratings ($n = 104$ ratings of 55 targets) on Form R of the NEO-PI-R.

* $p < .05$. ** $p < .01$. *** $p < .001$.

PSY-5 scales, most obviously in the correspondence between Negative Emotionality and Neuroticism and between Positive Emotionality and Extraversion. It is also reasonable to find that Aggressiveness is related to both Assertiveness and low Compliance, and Constraint is related to Deliberation and low Excitement Seeking. These combinations of elements, however, represent different constructs from those found in the Five-Factor Model.

Table 1 also makes it clear that there are distinctive dimensions in each instrument. The NEO-PI-R does not measure Psychoticism (a scale that might also be interpreted, especially in a

normal sample, as an artifact of endorsement frequency and evaluation; cf. A. R. Harkness & McNulty, 1994) whereas the PSY-5 does not have any scale that correlates strongly ($>.40$) with either Openness to Experience or Conscientiousness. These scales represent areas in which the two instruments might complement one another.

Study 2

Although the findings from Study 1 suggest many meaningful relationships between the NEO-PI-R and the PSY-5 scales in a

nonclinical sample, they do not address several other issues that are relevant to a comparison between the FFM of personality and the PSY-5 constructs. First, can the relationships obtained in Study 1 be replicated in a clinical sample? Second, how do the PSY-5 and NEO-PI instruments compare with respect to their test-retest reliabilities and their stability over time? Third, how do these two instruments compare with respect to their relationships to personality disorder pathology? Study 2 presents results that address these three issues.

Method

In this study, participants for a 6-month longitudinal study on personality features were recruited from recent admissions to two outpatient clinics in Columbia, Missouri, that served a mixture of community residents, university staff, and students. Exclusionary criteria included age of less than 18 years; presence of a chronic substance abuse or a chronic organic condition; current treatment with medication for a *DSM-III-R* (American Psychiatric Association, 1987) condition; and presence of acutely suicidal, acutely psychotic, or bipolar features. These diagnostic exclusionary criteria were employed to minimize the potential effects of acute distress on test scores and to ensure valid responses from the participants. Therefore, the final sample was composed of outpatients with anxiety disorders, mild depressive disorders, dysthymia, or personality disorders. There were three assessment occasions in this study: Entry or Time 1 (T1; $n = 57$), 3-month follow-up or Time 2 (T2; $n = 51$), and 6-month follow-up or Time 3 (T3; $n = 44$). At T1, 72% of the sample was female, and at T3, 73% of the sample was female. The mean age of the 57 participants was 28.12 ($SD = 7.90$), the mean years of education were 15.32 ($SD = 2.42$), and 30% of the sample had college or advanced degrees. Ninety-seven percent was White, and undergraduate or graduate students comprised 53% of the sample. More details regarding the study's procedures and characteristics of the sample can be found in Trull and Goodwin (1993).

Of relevance to the present article, each participant completed the NEO-PI (Costa & McCrae, 1985) and the MMPI-2 PSY-5 (A. R. Harkness et al., 1995) at each of three assessment occasions, along with a number of other self-report measures. The NEO-PI (Costa & McCrae, 1985) is an earlier version of the NEO-PI-R and yields facet scales only for Neuroticism, Extraversion, and Openness. Agreeableness and Conscientiousness are measured at a global domain level with 18-item scales. In addition, results from two measures of *DSM-III-R* personality disorder pathology were available for all participants at Time 1: the Structured Interview for *DSM-III-R* Personality (SIDP-R; Pfohl, Blun, Zimmerman, & Stangl, 1989) and the Personality Diagnostic Questionnaire—Revised (PDQ-R; Hyler & Rieder, 1987). (Personality disorder scales can also be created from MMPI-2 items, but these were not analyzed in the present study because of item overlap with PSY-5 scales.)

The SIDP-R is a frequently used semistructured interview that assesses *DSM-III-R* (American Psychiatric Association, 1987) personality disorder criteria. Several studies have reported impressive interrater reliabilities for both personality disorder diagnoses and symptom counts derived from the SIDP-R (e.g., see Zimmerman & Coryell, 1989). Five researchers (three women and two men) served as SIDP-R interviewers in our study. All interviewers were unaware of the self-report scores of the participants, and all interviewers had received extensive training in the administration and scoring of the SIDP-R. Reliability checks were conducted for 18 randomly selected T1 SIDP-R interviews; a second researcher reviewed the audiotapes of these interviews and provided independent ratings. Intraclass correlation coefficients (ICCs) were computed, comparing the independent ratings of the number of criteria met for each personality disorder as assessed by

the SIDP-R. The mean ICC for the *DSM-III-R* personality disorder criteria sets was .81 (range .61 to .95).

According to the SIDP-R, 39% of the sample received at least one *DSM-III-R* personality disorder diagnosis. The most prevalent diagnoses were Histrionic (21%), Borderline (12%), Obsessive-Compulsive (12%), and Paranoid (11%) personality disorders. The least frequent diagnoses were the Schizoid (2%), Schizotypal (2%), and Sadistic (2%).

The PDQ-R is a 152-item inventory keyed to the personality disorder criteria presented in *DSM-III-R* and a score representing the number of criteria met for each disorder can be calculated. Total PDQ-R scores have been shown to discriminate well between those patients judged to display significant personality disorder traits (i.e., at least one personality disorder diagnosis) and those without prominent traits of a personality disorder (Hyler et al., 1988).

In addition, each participant completed the Beck Depression Inventory (BDI; Beck, 1978) and the Beck Anxiety Inventory (BAI; Beck, 1987) at Time 1. These measures were included in order to provide an estimate of the acute mood state (i.e., depression and anxiety) of the participants.

In Study 2, three major sets of analyses were conducted. First, zero-order correlations between PSY-5 scale scores and NEO-PI scores were computed at Time 1 to assess which significant relations identified in Study 1 were replicated in our clinical sample. Second, analyses to assess the stability of PSY-5 and NEO-PI scores over time were conducted. These analyses are important because it is widely believed that personality scale scores are likely to be distorted by mood or by psychiatric disorder (Hirschfeld, Klerman, Clayton, & Keller, 1983). However, if either PSY-5 or NEO-PI scores are grossly unstable, their usefulness in personality disorder assessment would be limited. To date, no study has assessed the stability of either PSY-5 scores or NEO-PI scores in a clinical sample. Third, the relations between personality disorder scores and scores on the PSY-5 scales and the NEO-PI scales, respectively, were assessed via two series of multiple regression analyses.

Results and Discussion

PSY-5 and NEO-PI relations. Table 2 presents the zero-order correlations between raw scores on the PSY-5 scales, and NEO-PI domain and facet *T* scores (standardized within gender; 1989 adult norms). As in Study 1, Negative Emotionality and Neuroticism were highly positively correlated as were Positive Emotionality and Extraversion. Further, the significant relationships between Aggressiveness and low Agreeableness, between Constraint and both Agreeableness and Conscientiousness, and between Psychoticism and Neuroticism were replicated in this clinical sample.

Mean level changes and test-retest stability. The PSY-5 and NEO-PI were completed by all available participants at each of the three assessment occasions (Time 1, $n = 57$; Time 2, $n = 51$; Time 3, $n = 44$). Table 3 presents the descriptive statistics for each scale at each assessment occasion for the 44 participants who completed all three assessments. At initial assessment, they scored as a group in the high range on NEO-PI Neuroticism and in the low range on Agreeableness and Conscientiousness, consistent with previous studies of clinical samples (Miller, 1991). Participants, were also slightly above average on Openness. Repeated measures analyses of variance (ANOVAs) were conducted to identify those scales whose mean score changed significantly over time. Over the course of the follow-up, there were significant declines in Negative Emotionality, Psychoticism, and Neuroticism scores and a significant increase in Agreeableness scores. These changes were modest in magni-

Table 2
Correlations Between the MMPI-2 Personality Psychopathology Five (PSY-5) and the NEO-PI Dimension and Facet T Scores at Time 1 (n = 57)

Scales	Personality Psychopathology Five scale				
	POSEMO	AGGRES	CONSTR	NEGEMO	PSYCHOT
NEO-PI Scale					
Neuroticism	-.40**	-.00	-.16	.67***	.32**
Extraversion	.71***	.21	.14	-.07	.08
Openness	.34**	.15	-.12	-.05	-.01
Agreeableness	.16	-.47***	.35**	-.54***	-.43***
Conscientiousness	.10	.18	.37**	-.08	-.05
Neuroticism facets					
Anxiety	-.31*	.02	-.16	.58***	.27*
Hostility	-.02	.31**	-.12	.65***	.46***
Depression	-.44***	-.10	-.04	.49***	.17
Self-Consciousness	-.48***	-.24*	-.07	.42**	.12
Impulsiveness	-.24*	.03	-.11	.46***	.22
Vulnerability	-.29*	-.05	-.23	.55***	.25
Extraversion facets					
Warmth	.55***	-.05	.28*	-.12	.01
Gregariousness	.39**	-.14	-.02	-.16	-.22
Assertiveness	.31*	.53***	.04	-.02	.06
Activity	.30*	.21	.24*	-.01	.18
Excitement Seeking	.30*	.14	-.34**	.21	.10
Positive Emotions	.62***	.01	.30*	-.21	.04
Openness facets					
Fantasy	.05	.31*	-.29*	.11	.04
Aesthetics	.27*	.11	.06	-.03	.03
Feelings	.43***	.06	.14	-.07	-.03
Actions	.11	-.21	-.08	-.13	-.12
Ideas	.21	.22*	-.03	-.02	.09
Values	.10	-.01	-.17	-.14	-.19

Note. MMPI-2 = Minnesota Multiphasic Personality Inventory-2; NEO-PI = NEO Personality Inventory; POSEMO = Positive Emotionality and Extraversion; AGGRES = Aggressiveness; CONSTR = Constraint; NEGEMO = Negative Emotionality and Neuroticism; PSYCHOT = Psychoticism.

* $p < .05$. ** $p < .01$. *** $p < .001$.

tude, however, with the largest being approximately 0.5 *SD*. As a group, participants appear to have become slightly better adjusted over the 6-month period (cf. K. L. Harkness, Bagby, Joffe, & Levitt, 1994). No significant mean changes were seen in Positive Emotionality or Extraversion, or in Aggressiveness, Constraint, Openness, or Conscientiousness.

Because both the PSY-5 and the NEO-PI constructs are regarded as personality traits, a fair amount of stability over time is expected for these scores. Therefore, we calculated test-retest indices for PSY-5 and NEO-PI scores in order to assess their reliability over time. Table 4 presents three separate test-retest correlation coefficients for each scale: the correlation between Time 1 and Time 2 scores (T1-T2); between Time 2 and Time 3 scores (T2-T3); and between Time 1 and Time 3 scores (T1-T3). As indicated, there was some variability in individual scale test-retest correlations depending on the time frame involved. In general, the highest test-retest correlation coefficients were obtained for the PSY-5 Constraint and Negative Emotionality scales and the NEO-PI Openness and Conscientiousness scales. The lowest test-retest correlations were obtained for the PSY-5 Aggressiveness and Psychoticism scales.

By way of comparison, McCrae and Costa (1983) reported 6-month test-retest correlations of .87, .91, and .86 for Neuro-

ticism, Extraversion, and Openness, respectively, in a nonclinical sample. The corresponding T1-T3 correlations (6-month test-retest) in Table 4 are lower than these values, suggesting that these scores may be somewhat affected by the course of psychiatric disorder(s) or treatment. However, these test-retest correlations, especially for the NEO-PI scales, are still quite high, indicating substantial stability of trait scores even in this clinical sample.

Relationship to personality disorders. At study entry, all participants completed an Axis II structured interview (the SIDP-R) as well as an Axis II self-report inventory (the PDQ-R).³ Zero-order correlations showed predictable patterns of relations between the personality measures and the personality disorder measures. Across both personality disorder measures, NEO-PI Neuroticism was positively related to Schizotypal, Obsessive-Compulsive, Histrionic, Dependent, Narcissistic,

³ One participant failed to complete a substantial proportion of the PDQ-R items. Therefore, complete data were available from only 56 participants at T1. Information regarding the means and standard deviations of the SIDP-R and PDQ-R symptom counts for those participants who completed the entire study ($n = 44$) is presented in Trull and Goodwin (1993).

Table 3
Descriptive Statistics for the Personality Psychopathology Five (PSY-5) Scales and the NEO Personality Inventory (NEO-PI) Scales

Scale	<i>M</i>	<i>SD</i>	Range
PSY-5			
Positive Emotionality			
T1	20.37	4.46	4–30
T2	21.07	4.70	6–33
T3	21.74	5.25	1–31
Aggressiveness			
T1	6.93	2.49	2–14
T2	6.93	2.66	1–15
T3	6.77	2.29	2–11
Constraint			
T1	17.67	4.08	5–25
T2	17.60	4.04	9–26
T3	17.42	4.30	9–28
Negative Emotionality***			
T1	15.70	6.35	3–30
T2	14.65	6.60	1–27
T3	13.26	6.18	3–24
Psychoticism*			
T1	5.42	4.02	0–20
T2	4.84	3.10	0–14
T3	3.95	2.45	0–11
NEO-PI			
Neuroticism***			
T1	63.82	11.56	37–85
T2	62.70	12.19	40–92
T3	58.86	10.77	40–86
Extraversion			
T1	50.17	10.69	22–72
T2	52.28	12.67	14–96
T3	51.49	11.36	21–75
Openness			
T1	56.28	9.54	26–77
T2	55.30	9.63	21–79
T3	55.17	13.03	5–76
Agreeableness**			
T1	42.97	13.90	15–76
T2	44.87	11.46	17–69
T3	46.89	11.96	20–71
Conscientiousness			
T1	42.95	11.13	5–68
T2	43.20	10.44	9–66
T3	42.91	10.36	17–65

Note. *N* = 44. NEO-PI scores are *T* scores standardized within gender (1989 adult norms). T1 = Time 1 (Entry); T2 = Time 2 (3-month follow-up); T3 = Time 3 (6-month follow-up). Asterisks indicate a significant change in mean levels of scores for that scale over time: * *p* < .05. ** *p* < .01. *** *p* < .001.

Avoidant, Borderline, and Self-Defeating personality disorders. Extraversion was negatively related to Schizoid and Avoidant symptom counts, and Openness was negatively related to Schizoid personality disorder. Agreeableness was negatively related to symptoms of many disorders as measured by both the SIDP-R and PDQ-R: Paranoid, Schizoid, Schizotypal, Obsessive-Compulsive, Antisocial, Narcissistic, Borderline, Passive Aggressive, and Sadistic. Finally, low Conscientiousness was related to Antisocial symptom counts.

Table 4
Test-Retest Correlations for the Personality Psychopathology Five (PSY-5) and NEO Personality Inventory (NEO-PI) Scales

Scale	Test-retest correlations		
	T1-T2	T2-T3	T1-T3
PSY-5			
Positive Emotionality	.76	.81	.70
Aggressiveness	.65	.44	.62
Constraint	.84	.85	.86
Negative Emotionality	.78	.81	.84
Psychoticism	.83	.64	.67
Mean	.77	.71	.74
NEO-PI			
Neuroticism	.88	.89	.79
Extraversion	.80	.80	.79
Openness	.86	.85	.84
Agreeableness	.83	.78	.76
Conscientiousness	.89	.86	.84
Mean	.85	.84	.80

Note. *N* = 44. Test-retest correlations are zero-order correlations. T1 = Time 1 (study entry); T2 = Time 2 (3-month follow-up); T3 = Time 3 (6-month follow-up).

Most of these correlations are consistent with prior studies and hypotheses (Costa & Widiger, 1994).

PSY-5 scales also showed many predictable correlations with the personality disorder measures. Positive Emotionality was negatively related to Schizoid and Avoidant symptom counts, and Aggressiveness was positively correlated with Paranoid, Antisocial, Sadistic, and Self-Defeating personality disorders. Low Constraint was associated with Antisocial, Borderline, and Sadistic symptom counts, and Psychoticism was positively correlated with Borderline and Self-Defeating personality disorders. Finally, high Negative Emotionality was associated with many disorders: Paranoid, Schizotypal, Obsessive-Compulsive, Histrionic, Dependent, Narcissistic, Borderline, Passive Aggressive, and Self-Defeating.

To compare the two personality instruments as predictors of personality disorder symptoms, two sets of regression analyses were conducted. In the first set of analyses, Axis II symptom counts derived from the SIDP-R (i.e., the number of *DSM-III-R* criteria met) were simultaneously regressed onto the scale scores from the PSY-5 or the NEO-PI. A significant *R*² indicated that scales from the respective personality instrument accounted for a significant amount of the variance in the symptom count of that particular Axis II disorder. The results for the SIDP-R symptom counts are presented in Table 5. A second set of regression analyses was identical to the first except that symptom count scores were derived from the PDQ-R, a self-report inventory. These results are presented in Table 6. Finally, both sets of regression analyses were repeated with the exception that BDI and BAI scores were partialled first from the Axis II symptom counts in order to remove variance attributable to acute mood state (i.e., depression and anxiety). These results also appear in Tables 5 and 6.⁴

⁴ As one reviewer pointed out, our analyses partialling BDI and BAI scores from Axis II symptom counts may be quite conservative in that

Table 5
Regression Results for Each SIDP-R Axis II Symptom
Count in Which Predictors Were Either PSY-5 Scales
or NEO-PI Scales

SIDP-R symptom count	PSY-5		NEO-PI	
	R ²	R ² change ^a	R ²	R ² change ^a
Paranoid	.33**	.21*	.25*	.22*
Schizoid	.33**	.32**	.28**	.27**
Schizotypal	.24*	.16	.18	.12
Obsessive-Compulsive	.28**	.21*	.37***	.32**
Histrionic	.21*	.11	.24*	.16
Dependent	.22*	.09	.15	.06
Antisocial	.42***	.40***	.18	.17
Narcissistic	.24*	.11	.17	.10
Avoidant	.22*	.14	.18	.11
Borderline	.51***	.28***	.43***	.26**
Passive Aggressive	.21*	.21*	.21*	.20*
Sadistic	.24*	.23*	.21*	.22*
Self-Defeating	.33***	.23*	.28**	.23**
Mean variance accounted for	.29	.21	.24	.19

Note. $N = 56$. SIDP-R = Structured Interview for DSM-III-R Personality; PSY-5 = Personality Psychopathology Five; NEO-PI = NEO Personality Inventory.

^a Change in R² after entering Beck Depression Inventory and Beck Anxiety Inventory scores in a previous step.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Examining the SIDP-R data first, we found that multiple regression analyses indicated the PSY-5 scales were significantly related to all 13 Axis II symptom counts, whereas the NEO-PI scales were significantly related to 8. After controlling for depression and anxiety scores, PSY-5 scales predicted symptom counts for eight personality disorders; NEO-PI scales also predicted symptom counts for seven of these disorders. Overall, both instruments predicted a significant amount of variance in the majority of interview-based personality disorder symptom scores, even after controlling for mood state. The most conspicuous difference between the two instruments occurred for Antisocial personality disorder, where the PSY-5 scales jointly accounted for twice as much variance as the NEO-PI scales.

Table 6 shows results of regressions predicting self-reported symptom counts derived from the PDQ-R. Probably because of shared method, these associations are in general larger than those presented in Table 5, and almost all of the associations are significant. After controlling for mood state, 7 of the PSY-5 analyses and all 13 of the NEO-PI analyses showed significant effects.

Results for the two personality disorder instruments differed in some details, but, as the summary of replicated simple cor-

relations demonstrated, there was also substantial cross-measure consistency. This is particularly noteworthy because alternative measures of personality disorders typically show only modest convergence; in the present sample, the median correlation between corresponding SIDP-R and PDQ-R symptom counts was only .43 (Trull & Larson, 1994). Judged against these relatively low validity coefficients, the magnitude of the predictions in Tables 5 and 6 are noteworthy: Personality traits, as assessed by either the PSY-5 or the NEO-PI, appear to measure a substantial part of the valid variance in personality disorder scales.

Correlations with NEO-PI facet scales. Correlations between NEO-PI Neuroticism, Extraversion, and Openness facet scales and the two sets of personality disorder symptom counts showed a variety of predictable associations.⁵ For example, across both personality disorder measures, the number of Avoidant symptoms was positively correlated with the facets of Depression and Self-Consciousness, and negatively correlated with Warmth, Assertiveness, and Positive Emotions. Schizoid symptom counts were negatively associated with Warmth, Positive Emotions, and Openness to Aesthetics and Feelings. Borderline symptoms were associated with Anxiety, Hostility, Depression, Impulsiveness, and Vulnerability, and with the Excitement-Seeking facet of Extraversion. An even more differentiated set of predictions might be provided by the NEO-PI-R, which includes facet scales for all five domains.

General Discussion

This research examined two alternative systems for personality assessment in two different studies. In the first study, shortened versions of PSY-5 scales were correlated with self-reports and peer ratings on the NEO-PI-R in a community sample; in the second, the full PSY-5 scales were correlated with unrevised NEO-PI scales in a smaller clinical sample. Despite these methodological differences, strikingly similar results were found. PSY-5 Positive Emotionality and Negative Emotionality strongly resembled the NEO-PI dimensions of Extraversion and Neuroticism, respectively. The other PSY-5 scales did not show a one-to-one correspondence with NEO-PI dimensions, but they were related to specific NEO-PI facet scales: Aggressiveness appears to combine some aspects of low Agreeableness and high Extraversion; Constraint may be characterized by high Agreeableness and high Conscientiousness; and Psychoticism was positively related to Neuroticism and negatively related to some facets of Agreeableness. None of the PSY-5 scales was consistently related to Openness. It appears that these two instruments—one representing personality pathology constructs and the other measuring general dimensions of personality—assess overlapping but not identical sets of constructs.

Study 2 provided additional comparisons of the two instruments, examining their temporal stability and relation to personality disorder symptoms. Although psychiatric disorders and psychotherapeutic interventions might be thought to have

these measures likely represent some combination of state and trait negative affectivity or neuroticism (Clark & Watson, 1991). Despite this caveat, we believe that these supplementary analyses and results are useful because they represent an attempt to control for acute generalized distress, a state-like condition that is quite prevalent among patients presenting for treatment (as at Time 1 in Study 2).

⁵ The full matrix of intercorrelations among NEO-PI facet scores and the two sets of personality disorder symptom counts is available from Timothy J. Trull.

Table 6
Regression Results for Each PDQ-R Axis II Symptom Count in Which Predictors Were Either PSY-5 Scales or NEO-PI Scales

PDQ-R symptom count	PSY-5		NEO-PI	
	R ²	R ² change ^a	R ²	R ² change ^a
Paranoid	.26**	.14	.32**	.21*
Schizoid	.24*	.25*	.35***	.36**
Schizotypal	.38***	.30**	.27**	.19*
Obsessive-Compulsive	.24*	.14	.32**	.23*
Histrionic	.42***	.33***	.56***	.48***
Dependent	.20*	.13	.36***	.28**
Antisocial	.36***	.30**	.28**	.22*
Narcissistic	.47***	.29**	.56***	.40***
Avoidant	.23*	.16	.26**	.20*
Borderline	.64***	.43***	.55***	.37***
Passive Aggressive	.19	.18	.40***	.42***
Sadistic	.25*	.27**	.22*	.24*
Self-Defeating	.21*	.16	.39***	.31**
Mean variance accounted for	.31	.24	.37	.30

Note. N = 56. PDQ-R = Personality Diagnostic Questionnaire—Revised; PSY-5 = Personality Psychopathology Five; NEO-PI = NEO Personality Inventory.

^a Change in R² after entering Beck Depression Inventory and Beck Anxiety Inventory scores in a previous step.

* p < .05. ** p < .01. *** p < .001.

profound effects on self-reported personality traits, in fact, both mean levels and individual differences on NEO-PI and PSY-5 scales were generally stable. Further, as hypothesized, these trait measures were systematically related to personality disorder symptom counts, whether based on interviews or on self-reports.

Consider Borderline personality disorder. Tables 5 and 6 show that symptoms related to this disorder were predicted by both PSY-5 and NEO-PI scales, even after controlling for state anxiety and depression. An examination of zero-order correlations showed that Borderline symptoms were specifically (and predictably) related to NEO-PI Neuroticism and low Agreeableness, and to PSY-5 Negative Emotionality, Psychoticism, and low Constraint. This wide array of personality correlates may help explain why Borderline Personality Disorder is the most frequent Axis II diagnosis in clinical samples (e.g., Widiger & Rogers, 1989): Extreme standing on any of several different personality dimensions appears to be associated with the manifestation of Borderline personality disorder symptoms.

The general pattern of findings relating personality traits to personality disorder symptoms is consistent with the premise that personality psychopathology can be understood largely in terms of common dimensions of personality. In particular, these data support many of the hypotheses relating personality disorders to traits defining the FFM. For example, Widiger, Trull, Clarkin, Sanderson, and Costa (1994) hypothesized that Avoidant personality disorder would be characterized by high Self-Consciousness and low Warmth, and that Schizoid personality disorder would be characterized by low Positive Emotions and Openness to Feelings. Analyses of NEO-PI facet scales supported all of these hypotheses. Such findings provide an empir-

ical basis for arguments that personality psychopathology might be assessed and interpreted from a dimensional perspective (Costa & Widiger, 1994).

But which dimensional system should be preferred? *DSM-IV* notes that "there have been many different attempts to identify the most fundamental dimensions that underlie the entire domain of normal and pathological personality functioning" (American Psychiatric Association, 1994, p. 633). In this research, two approaches were compared: the FFM and PSY-5 systems. Although the results of a single small study cannot be conclusive, the present data suggest that both instruments work reasonably well.

In one respect these results are surprising. One might have expected the PSY-5 scales to be much better predictors of symptoms because they were composed of items from a major measure of psychopathology, the MMPI-2, and were specifically developed with personality disorder psychopathology in mind. For example, the PSY-5 Constraint scale includes items about lying, stealing, and trouble with the law; its high correlation with measures of Antisocial personality disorder are thus understandable. By contrast, the NEO-PI was developed as a measure of normal personality, with no attempt to assess specific dimensions of personality relevant to psychopathology.

In another respect, however, the usefulness of the NEO-PI is not surprising. The FFM, on which it was explicitly based, has been elaborated and refined over a period of 30 years as a comprehensive model of personality (McCrae & John, 1992). The PSY-5 scales are much more recent in origin and have not yet benefitted from extensive research. Refinements to the scales might show improved usefulness for the constructs they are intended to measure.

At a practical level, PSY-5 scales can be scored from the items of the MMPI-2, an instrument that is already widely used in clinical practice. We would encourage researchers and clinicians who employ the MMPI-2 to score and examine these scales as possible predictors of Axis II diagnoses. However, we would also recommend supplementing measures of psychopathology like the MMPI-2 with measures of the FFM, such as the NEO-PI-R. As the present research demonstrates, FFM dimensions are clearly relevant to the diagnosis of personality psychopathology. In addition, they are also useful in providing a full personality profile, including strengths as well as weaknesses, that can be helpful in establishing rapport, selecting approaches to therapy, and anticipating the course and outcome of treatment (Costa & McCrae, 1992a; Miller, 1991).

Comparisons of the NEO-PI-R and PSY-5 in additional clinical samples (including inpatient samples) are clearly needed. The present research offered a global comparison by using scales from each instrument separately as predictors of personality disorder symptom counts. With larger samples, more detailed comparisons could be made to determine whether, when used together, either instrument offers incremental validity. Such tests would be particularly valuable with regard to PSY-5 Psychoticism and NEO-PI-R Openness and Conscientiousness, dimensions that appear to be underrepresented in the other instrument. It would also be of interest to examine both instruments as predictors of Axis I disorders (cf. Trull & Sher, 1994). Personality traits are pervasive consistencies in cognitive, affect-

tive, and behavioral styles, and as such are likely to be relevant to a wide range of clinical phenomena.

References

- 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.
- Beck, A. T. (1978). *Beck Depression Inventory*. Philadelphia: Center for Cognitive Therapy.
- Beck, A. T. (1987). *Beck Anxiety Inventory*. Philadelphia: Center for Cognitive Therapy.
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory—2 (MMPI-2): Manual for administration and scoring*. Minneapolis: University of Minnesota Press.
- Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100*, 316–336.
- Costa, P. T., Jr., Busch, C. M., Zonderman, A. B., & McCrae, R. R. (1986). Correlations of MMPI factor scales with measures of the five-factor model of personality. *Journal of Personality Assessment, 50*, 640–650.
- Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., Jr., & McCrae, R. R. (1990). Personality disorders and the five-factor model of personality. *Journal of Personality Disorders, 4*, 362–371.
- Costa, P. T., Jr., & McCrae, R. R. (1992a). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment, 4*, 5–13.
- Costa, P. T., Jr., & McCrae, R. R. (1992b). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., Jr., & McCrae, R. R. (1992c). Trait psychology comes of age. In T. B. Sonderegger (Ed.), *Nebraska symposium on motivation: Psychology and aging* (pp. 169–204). Lincoln, NE: University of Nebraska Press.
- Costa, P. T., Jr., & Widiger, T. A. (1994). *Personality disorders and the five-factor model of personality*. Washington, DC: American Psychological Association.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417–470.
- Harkness, A. R. (1992). Fundamental topics in the personality disorders: Candidate trait dimensions from lower regions of the hierarchy. *Psychological Assessment, 4*, 251–259.
- Harkness, A. R., & McNulty, J. L. (1994). The Personality Psychopathology Five (PSY-5): Issues from the pages of a diagnostic manual instead of a dictionary. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 291–315). New York: Springer.
- Harkness, A. R., McNulty, J. L., & Ben-Porath, Y. S. (1995). The Personality Psychopathology Five (PSY-5): Constructs and MMPI-2 scales. *Psychological Assessment, 7*, 104–114.
- Harkness, K. L., Bagby, R. M., Joffe, R. T., & Levitt, A. (1994, July). *Major depression, double depression and the Neuroticism personality dimension*. Paper presented at the 7th European Conference of Personality, Madrid.
- Hirschfeld, R. M. A., Klerman, G., Clayton, P. J., & Keller, M. B. (1983). Personality and depression: Empirical findings. *Archives of General Psychiatry, 40*, 993–998.
- Hyler, S. E., & Rieder, R. O. (1987). *PDQ-R personality questionnaire*. New York: New York State Psychiatric Institute.
- Hyler, S., Rieder, R., Williams, J., Spitzer, R., Hendler, J., & Lyons, M. (1988). The Personality Diagnostic Questionnaire: Development and preliminary results. *Journal of Personality Disorders, 2*, 229–237.
- McCrae, R. R., & Costa, P. T., Jr. (1983). Joint factors in self-reports and ratings: Neuroticism, extraversion, and openness to experience. *Personality and Individual Differences, 4*, 245–255.
- McCrae, R. R., & John, O. P. (1992). An introduction to the Five-Factor Model and its applications. *Journal of Personality, 60*, 175–215.
- Miller, T. R. (1991). The psychotherapeutic utility of the five-factor model of personality: A clinician's experience. *Journal of Personality Assessment, 57*, 415–433.
- Pfohl, B., Blum, N., Zimmerman, M., & Stangl, D. (1989). *Structured interview for DSM-III-R personality: SIDP-R*. Iowa City, IA: Author.
- Shock, N. W., Greulich, R. C., Andres, R., Arenberg, D., Costa, P. T., Jr., Lakatta, E. G., & Tobin, J. D. (1984). *Normal human aging: The Baltimore Longitudinal Study of Aging* (NIH Publication No. 84-2450). Bethesda, MD: National Institutes of Health.
- Soldz, S., Budman, S., Demby, A., & Merry, J. (1993). Representation of personality disorders in circumplex and five-factor space: Explorations with a clinical sample. *Psychological Assessment, 5*, 41–52.
- Trull, T. J. (1992). *DSM-III-R personality disorders and the Five-Factor Model of personality: An empirical comparison*. *Journal of Abnormal Psychology, 101*, 553–560.
- Trull, T. J., & Goodwin, A. H. (1993). Relationship between mood changes and the report of personality disorder symptoms. *Journal of Personality Assessment, 61*, 99–111.
- Trull, T. J., & Larson, S. L. (1994). External validity of two personality disorder inventories. *Journal of Personality Disorders, 8*, 96–103.
- 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.
- Widiger, T. A., & Rogers, J. H. (1989). Prevalence and comorbidity of personality disorders. *Psychiatric Annals, 19*, 132–136.
- Widiger, T. A., Trull, T. J., Clarkin, J. F., Sanderson, C., & Costa, P. T., Jr. (1994). A description of the *DSM-III-R* and *DSM-IV* personality disorders with the Five-Factor Model of personality. In P. T. Costa, Jr., & T. A. Widiger (Eds.), *Personality disorders and the Five-Factor Model of personality* (pp. 41–56). Washington, DC: American Psychological Association.
- Wiggins, J. S., & Pincus, A. L. (1989). Conceptions of personality disorders and dimensions of personality. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 1*, 305–316.
- Zimmerman, M., & Coryell, W. (1989). The reliability of personality disorder diagnoses in a nonpatient sample. *Journal of Personality Disorders, 3*, 53–57.

Received August 31, 1994

Revision received May 2, 1995

Accepted May 2, 1995 ■