

## A First Look at the Structured Clinical Interview for *DSM-IV* Personality Disorders Screening Questionnaire: More Than Just a Screener?

Ralph L. Piedmont, Martin F. Sherman, Nancy C. Sherman, and Joseph E. G. Williams

*This study examined the psychometrics of the Structured Clinical Interview for DSM-IV Personality Disorders Personality Questionnaire (SCID-IIP) self-report personality questionnaire. The responses to the instrument were found reliable and evidenced good self-other convergence. Correlations with external criteria showed the SCID-IIP to contain broad personological information. The scale's basis in DSM-IV nosology makes it a potentially important tool for clinical research.*

Following the establishment of criterion sets for personality disorders (e.g., *Diagnostic and Statistical Manual of Mental Disorders*, third edition [*DSM-III*; American Psychiatric Association, 1980]), clinicians and researchers have attempted to construct instruments that would yield responses that are valid and reliable for assessing these disorders. Because Axis II dynamics are likely to have an effect on the occurrence, expression, course, and treatment of Axis I disorders, having an efficient and reliable means of assessing characterological qualities becomes clinically imperative (e.g., Shea, Widiger, & Klein, 1992; Widiger & Rogers, 1989; Woody, McLellan, Luborsky, & O'Brien, 1985). The field has responded enthusiastically to this need by creating a wide spectrum of both interviewer-based (e.g., Personality Disorder Examination; Loranger, Susman, Oldham, & Russakoff, 1987) and self-report instruments (e.g., Personality Diagnostic Questionnaire-Revised; Hyler, Rieder, Williams, & Spitzer, 1987). In comparison with the structured/semi-structured interview, in general, self-report instruments have tended to produce responses that tended to be as reliable and valid (Guthrie & Mobley, 1994; Trull & Goodwin, 1993) as well as being more economical to use. Although the current *DSM* system is interviewer based, the ease of administration and logistical parsimony of self-reports continue to make them popular in the field.

The most straightforward self-report instrument that is also directly tied to the *Diagnostic and Statistical Manual of Mental Health Disorders*, fourth edition (*DSM-IV*; American Psychiatric Association [APA], 1994) criteria is the Structured Clinical Interview for *DSM-IV* Personality Disorders Personality Questionnaire (SCID-IIP; First, Gibbon, Spitzer, Williams, & Benjamin, 1997). The SCID-IIP consists of 119 items (which are the diagnostic criteria for the 12 different Axis II categories), and patients are asked to endorse the presence or absence of specific symptoms. There is no specific scoring of responses per se; clinicians merely use the endorsed items as areas that need to be pursued in detail when the structured clinical interview is conducted. Given the ease of administration and its ties to the diagnostic criteria, this scale would seem to be an ideal measure for research on personality disorders. However, First et al. (1997) have made it clear that they "do not recommend using the Personality Questionnaire as a stand-alone instrument for any pur-

Ralph L. Piedmont, Department of Pastoral Counseling, and Martin F. Sherman, Department of Psychology, Loyola College; Nancy C. Sherman, Villa Julie College; Joseph E. G. Williams, Department of Psychology, Eastern Illinois University, Charleston. Partial support for this study was provided by the Institute for Religious and Psychological Research. Portions of this study were presented at the 1998 Annual Meeting of the Eastern Psychological Association, Providence, Rhode Island. The authors would like to thank Mark Leach for his helpful comments on an earlier version of this article. Correspondence concerning this article should be addressed to Ralph L. Piedmont, Department of Pastoral Counseling, Loyola College in Maryland, 8890 McGaw Road, Columbia, MD 21045 (e-mail: rpiedmont@loyola.edu).

pose other than as a rough screening device" (p. 7). As a result, little attention has gone into evaluating this measure. To date, no research has been done with the *DSM-IV* version of the instrument, although some studies have examined the *Diagnostic and Statistical Manual of Mental Disorders*, third edition, revised (*DSM-III-R*; APA, 1987) version.

Ekselius, Linstrom, von Knorring, Bodlund, and Kullgren (1994) compared the results of the SCID-II interviews of 69 patients with the results of the SCID-IIP and concluded that the SCID-IIP may be a useful screener for personality disorders and has the potential to be an independent diagnostic measure in epidemiological studies. Ouimette and Klein (1995), using a nonclinical sample (97 college students), demonstrated that responses to the personality disorder scales of the SCID-IIP were relatively stable over time (median retest correlation was .69 over a 10-week period), not influenced by state depression, and relatively concordant with informant reports (median informant-subject correlation was .35). They concluded that their results showed promise for the SCID-IIP and suggested that additional research was warranted. Piedmont and Sherman (1998) provided clear construct validity evidence for the responses to the SCID-IIP in a sample of polysubstance abusers. Cross-time, cross-method, cross-observer correlations showed the SCID-IIP scale to correlate moderately (from .20 to .58) with measures of symptomatology, personality, and counselor ratings. If the object of using the SCID-IIP is to serve as a general personality disorder assessment, in which the focus would not be on diagnosis, these results are very encouraging. Carey (1994) came to the same conclusion when she studied 25 severely mentally ill outpatients who were diagnosed with Axis I disorders. She indicated that the SCID-IIP "may be a useful source of clinical data; to the extent that the items reflect patients' views of their typical patterns of thinking, feeling, and relating to others, such information may have value independent of its contribution to the diagnostic process" (Carey, 1994, p. 671).

Thus, the extant literature on the SCID-IIP screener suggests the possibility that the instrument may have larger potential value for the field than originally thought. Responses to the instrument have demonstrated useful construct validity by capturing salient personological qualities, making it useful in nonclinical contexts where subthreshold clinical dynamics may be of greater interest (e.g., college students, employment settings; see McKeeman & Erickson, 1997). The initial proscription not to use the instrument alone may have been premature. Although a body of literature exists on the *DSM-III-R* version of the instrument, nothing yet exists for its *DSM-IV* cousin. The purpose of the current report was to provide a first look at the reliability and validity of responses to this instrument across both normal and clinical samples.

Given our previous research with the *DSM-III-R* version of the SCID-IIP (e.g., Piedmont & Sherman, 1998), we identified six important issues that formed the basis for the present study. First, there was tremendous overlap among the various SCID-IIP scales. Such redundancy has important diagnostic implications and reflects the difficulty in making differential diagnoses. In our previous study we regressed each SCID-IIP scale on all the others. The resulting multiple  $R^2$  represented each scale's communality with the others. In that study we found that the average  $R^2$  was .53, indicating that over half of each scale's total variance was shared with the linear combination of the others. In that study of substance abusers, the average reliability of the responses to the scales was .63, thus over 90% of the reliable variance was shared in common among the scales. Given the *DSM-IV*'s efforts at making inclusion criteria for the personality disorders more orthogonal, some evaluation of interscale overlap is warranted to determine if a better diagnostic separation has been achieved. Second, in our previous study, the Antisocial scale appeared to be the most independent of the other scales. It was not clear whether this was a result of the dichotomous response format for those items (and the resulting reduced variability of the scale) or because the items queried about behaviors occurring before age 15 (which made them less relevant to adult functioning). The *DSM-IV* version contains items tapping into behaviors both before and after age 15. We therefore created two Antisocial scales, one pertaining to behaviors before age 15, the other after age 15. Furthermore, we converted the items in this latter scale into a Likert-type format. We hoped that this would enable us to better determine the psychometric characteristics of responses to these two scales. Third, although scores on the screener correlated moderately with counselor ratings of personality, no data exist which would allow the evaluation of the *DSM-IV* version's

self-peer convergence. Such cross-observer convergence would provide a level of consensual validation for the constructs being assessed, demonstrating that they represent relatively overt, broad aspects of functioning that can be reliably assessed by knowledgeable observers. The present study obtained observer ratings on the SCID-IIP scales to examine this issue. Fourth, this study looked at the potential role of acquiescence in distorting scores in a college sample. Given that all the items on the screener questionnaire are framed in the positive direction (i.e., to score high on a scale individuals need to endorse an item), acquiescence becomes a real concern. Such a response distortion may be another source for the overdiagnosing consistently found with this scale. Correlating SCID-IIP scores with an acquiescence index would provide some insights into this issue.

A fifth goal of this study was to evaluate the construct validity of the responses of the *DSM-IV* version of the SCID-IIP screener by correlating them with scores on the Five-Factor Model (FFM) of personality. There is a burgeoning research literature showing strong empirical links between characterological impairment and personality functioning (e.g., Dreary, Peter, Austin, & Gibson, 1998; Trull, 1992). The FFM domains have also been shown to evidence substantial, theoretically consistent correlations with the diagnostic dimensions of each scale (Widiger & Costa, 1994, review several other similar studies). On the basis of a review of studies that correlated a variety of measures of Axis II functioning with the domains of the FFM (e.g., Coolidge et al., 1994; Costa & McCrae, 1990; Piedmont, Sherman, & Murphy, 1999; Soldz, Budman, Demby, & Merry, 1993; Trull, 1992), consistent patterns of association were identified between the disorder scales and each of the personality dimensions. These consistencies will serve as a set of hypotheses for the present study. It was expected that Neuroticism would correlate positively with the Avoidant, Dependent, Obsessive-Compulsive, Passive-Aggressive, Depressive, Paranoid, Schizotypal, Histrionic, Narcissistic, and Borderline disorders. Extraversion was expected to correlate positively with the Obsessive-Compulsive, Histrionic, and Narcissistic disorder scales and negatively with the Avoidant, Depressive, Schizotypal, and Schizoid disorder scales. Agreeableness was expected to correlate negatively with the Passive-Aggressive, Paranoid, Schizotypal, Narcissistic, Borderline, and Antisocial disorder scales. Conscientiousness was expected to correlate negatively with the Dependent, Passive-Aggressive, Histrionic, Borderline, and Antisocial disorder scales. No predictions were made for Openness to Experience, which is usually found to be independent of Axis II dynamics. These 30 expectations are used to evaluate the construct validity of responses to the SCID-IIP scales.

The final goal of this study was to examine SCID-IIP performance in both college and clinical samples. The latter consisted of an outpatient sample of individuals in treatment for substance abuse. Three questions were addressed. First, would the SCID-IIP evidence more extreme responses than the college sample (i.e., a greater endorsement of pathology). Second, would participants' responses evidence levels of internal reliability? Finally, would scores on the SCID-IIP continue to correlate with measures of personality and symptom experience in expected ways?

## METHOD

### Participants

*College sample.* Participants consisted of 218 women and 84 men, ages 17 to 54 ( $M = 20$ ,  $SD = 4.1$ ), who were either undergraduate or graduate students. More than 90% were White, 5% were African American, 2% were Asian, and 1% each were Hispanic, other, and had missing data. All participants volunteered. Undergraduate participants received course credit for their participation.

Of the participants, 141 obtained peer ratings (128 had two raters, 13 had only a single rater). Of these 269 raters, 173 were women and 96 were men, with an average age of 27 (range: 14 to 64). On average, raters had known the participants 9 years (range: 3 months to 38 years). Raters were also asked how well they felt they knew the participant on a Likert-type scale from 1 (*not at all*) to 7 (*excellent*); the average rating was 6.3 (range: 3 to 7). Obviously raters knew their ratees quite well. In fact, for the undergraduates, many appeared to be parents or relatives.

*Clinical sample.* Participants consisted of 57 men and 16 women, 19 to 66 years of age ( $M = 41$ ), who had consecutive admissions to an urban, outpatient substance abuse treatment program. The majority were African American (85%), with Caucasians (14%) and Asians (1%) constituting the remainder. The most frequently abused substances were alcohol, cocaine, heroin, and marijuana. All members of the program were volunteers, who had been recommended by local shelters and social agencies. To be accepted to the program, an individual had to be drug free for at least 30 days. Any use of substances during the program was grounds for immediate dismissal. Although the program did not perform any of its own drug tests to ensure drug-free compliance, these individuals were already living in controlled environments (e.g., shelters, halfway houses) where drug testing was being conducted on a random basis. On average, participants had been unemployed for the previous 18 months (range: 1 month to 8+ years) and had their last rehab experience 6 months prior to entering the current program. These individuals had an average of three prior detox experiences. Twenty-seven percent were on probation at the time of entering the program.

## Measures

*Revised NEO Personality Inventory (NEO-PI-R).* Developed by Costa and McCrae (1992a), this 240-item questionnaire was developed through rational and factor analytic methods to measure the five major factors of personality: Neuroticism (N), Extraversion (E), Openness to Experience (O), Agreeableness (A), and Conscientiousness (C). For each factor, there are six facet scales that are designed to capture more specific traits. Items are answered on a 5-point scale ranging from 1 (*strongly agree*) to 5 (*strongly disagree*), and scales are balanced to control for the effects of acquiescence. The items themselves are simple statements describing general tendencies (e.g., "It's often hard for me to make up my mind"; "I often crave excitement"). Normative internal consistency estimates of responses for the Form S scales for adults range from .59 to .92 (Costa & McCrae, 1992a). Six-year stability coefficients for responses range from .68 to .83 in both self-reports and peer ratings for N, E, and O. Three-year retest coefficients of responses of .63 and .79 were seen for brief versions of the A and C domain scales (Costa & McCrae, 1988). Responses to the NEO-PI-R have been validated in studies with other self-reports (Costa, McCrae, & Dye, 1991; McCrae & Costa, 1992; Piedmont & Weinstein, 1993). Scores have shown evidence of convergent and discriminant validity across instruments, methods, and observers and have been related to a number of life outcomes including frequency of somatic complaints, ability to cope with stress, and burnout (see Piedmont, 1998).

Because the items on the NEO-PI-R are balanced to control for acquiescent responding, it served as a means of indexing this response distortion. Prior to scoring the NEO-PI-R, the number of "agree" and "strongly agree" responses were calculated for each participant. This score was then regressed on the participant's normative scores on the five personality domains and the standardized residual was saved. This index represents an individual's tendency to agree with an item independent of the item's content, and research has shown it to be a valid measure of acquiescence (Costa & McCrae, 1992a; McCrae, Herbst, & Costa, 2001).

*Brief Symptom Inventory (BSI).* Developed by Derogatis (1993), this 53-item, self-report inventory is designed to capture psychological symptom patterns over nine primary, clinically relevant dimensions and three global indices. Each item is responded to on a 4-point Likert-type scale ranging from 0 (*not at all*) to 4 (*extremely*). For the purposes of this research, only the Global Severity Index, which is the sum of the nine symptom clusters divided by the total number of responses, was used. This index provides an overall index of symptomological distress. Derogatis, Rickels, and Rock (1976) have shown scores on the BSI to converge well with scores on the Minnesota Multiphasic Personality Inventory. Other validity research has shown that responses to the BSI can be useful in detecting symptomological distress in clients in a drug treatment context (Buckner & Mandell, 1990; Royse & Drude, 1984). Following guidelines provided in the manual, the scores from participants in this study were converted to T scores based on norms provided for the adult nonpatient sample.

*SCID-IIP*: This scale consists of 133 items and represents the actual questions used in the screener questionnaire. The first 118 items are responded to on a Likert-type scale ranging from 1 (*strongly agree*) to 5 (*strongly disagree*). The remaining 15 items are responded to on a simple Yes/No scale. These items were derived directly from the diagnostic criteria for each personality disorder listed in the *DSM-IV*. The one addition we made to this instrument concerned the Antisocial Personality Disorder subscale. The screener questionnaire uses only items that pertain to the respondent's life prior to age 15. These were the 15 Yes/No items. However, we have found that this manner of presenting the Antisocial items may not be psychometrically useful (Piedmont et al., 1999). As such, we added additional *DSM-IV* diagnostic items that pertain to adult behavior as well. Thus, there were two Antisocial subscales used in this study, one that focused on behavior prior to age 15 (and used a Yes/No response format), and another that focused on adult behavior (and used a Likert-type format). Both samples completed this version of the instrument.

These items were also phrased in the third person for use as an observer rating. "I" and "my" were changed to "he/she" and "his/hers." The same response formats were kept as used with the self-reports. Participants were asked to obtain two peer ratings. When there were two scores, they were averaged to form an aggregated score, which was used in all analyses. Observer ratings were only obtained for the student sample.

### Procedure

All undergraduate volunteer students completed the self-report materials during class time. Graduate students completed their forms on their own time. The sequencing of the scales was counterbalanced to control for any order effects. Individuals were asked to find two people who had known them for at least 3 months to rate them on the *SCID-IIP* questionnaire. Raters were provided with materials in a sealable envelope. Once completed, the ratings were sealed and returned to the experimenters (authors of the current study). Participants did not have access to the ratings. All materials were completed anonymously.

For the clinical sample, prior to beginning treatment in the program, clients completed the assessment materials as part of an initial screening session. Again, the sequencing of the scales was counterbalanced to control for any order effects. Usually this was accomplished in one sitting, although at times clients were allowed to complete the forms at home and to bring them back at the time of the first day of the program. All information was collected and completed before the treatment program began.

### RESULTS

Given the correlational nature of this study, in an effort to control for experiment-wise alpha levels, only effects significant at the .01 level or less are considered significant.

In comparing mean level scores on these scales between the student and clinical samples, a one-way multivariate analysis of covariance was performed, in which sample group was the independent variable, scores on the *SCID-IIP* scales were the dependent variables, and gender was the covariate. A significant effect for sample group was obtained, Wilks's lambda = .29, multivariate  $F(13, 255) = 67.59, p < .001$ . An inspection of the univariate effects showed that the clinical sample scored significantly higher than the college student sample on all the *SCID-IIP* scales.

Table 1 presents descriptive statistics and consistency of responding for each of the *SCID-IIP* personality disorder scales for both samples. Alpha reliabilities for the college self-report responses ranged from .53 for Obsessive-Compulsive to .86 for Borderline ( $M = .74$ ), whereas for the clinical sample a comparable level was also observed (mean  $\alpha = .73$ ). Reliability estimates for the observer ratings are consistent with the self-reports: Values ranged from a low of .60 for Schizoid to a high of .91 for Borderline ( $M = .81$ ). Despite the polythetic nature of the diagnostic categories, responses to these scales evidenced a fair degree of cohesiveness. A high level of internal consistency was also found for the observer ratings as well. Finally, for the college sample, self-other convergence was examined and substantial agreement across these two data sets was found (mean  $r = .43$ ). These values compared favorably with self-peer values found using normal personality

**TABLE 1**  
**Descriptive Statistics and Alpha Reliabilities for the Self-Report and**  
**Observer Rating Versions of the Structured Clinical Interview for DSM-IV**  
**Personality Disorders Personality Questionnaire (SCID-IIP) Screener Scales**

SCID-IIP Scale	Total Sample		Self-Report $\alpha$	Correlation With AI	Multiple $R^2$	Observer Rating $\alpha$ (N = 140)	Self-Other Correlation (N = 139)
	M	SD					
College sample (N = 302)							
Avoidant	11.71	4.5	.70	.11	.44***	.83	.49***
Dependent	10.55	4.3	.63	.11	.37***	.72	.34***
Obsessive-Compulsive	17.58	4.5	.53	.14	.28***	.66	.36***
Passive-Aggressive	12.00	4.3	.68	.08	.57***	.82	.30***
Depressive	11.50	5.8	.84	.11	.66***	.88	.38***
Paranoid	11.76	5.9	.82	.12	.60***	.86	.54***
Schizotypal	13.66	6.5	.77	.16**	.42***	.82	.34***
Schizoid	6.41	3.2	.56	.18**	.21***	.60	.31***
Histrionic	11.87	4.8	.74	.15**	.35***	.82	.54***
Narcissistic	24.92	8.5	.82	.11	.57***	.88	.30***
Borderline	21.30	9.8	.86	.04	.65***	.91	.49***
Antisocial 1 (adult)	15.07	8.3	.80	.06	.60***	.85	.53***
Antisocial 2 (youth)	1.23	2.1	.82	-.04	.35***	.86	.62***
Clinical sample (N = 68)							
Avoidant	19.46	5.1	.83	.22	.53***		
Dependent	18.74	4.6	.72	.14	.56***		
Obsessive-Compulsive	26.12	4.1	.50	.37**	.29		
Passive-Aggressive	20.81	4.1	.66	.25	.64***		
Depressive	21.84	5.6	.81	.27	.69***		
Paranoid	23.85	4.7	.71	.35**	.44***		
Schizotypal	27.79	5.6	.72	.31**	.42***		
Schizoid	15.15	3.5	.56	.32**	.33		
Histrionic	18.01	3.9	.64	.33**	.48***		
Narcissistic	45.10	8.3	.81	.46***	.64***		
Borderline	43.19	9.0	.83	.34**	.72***		
Antisocial 1 (adult)	43.60	9.0	.82	.33**	.58***		
Antisocial 2 (youth)	4.09	4.1	.89	.17	.46***		

Note. AI = Acquiescence Index.  
 \*\* $p < .01$ . \*\*\* $p < .001$ , two-tailed.

dimensions (e.g., Piedmont, 1994) and with research using the DSM-III-R version of the SCID-IIP scales (Ouimette & Klein, 1995).

Given that all the items on the scale are phrased in a positive direction, an evaluation of possible acquiescence effects was undertaken. An acquiescence index (AI) was developed from the NEO-PI-R, which is balanced to control for such effects. Prior to scoring the NEO-PI-R, the number of items to which individuals responded *agree* or *strongly agree* were counted. This score represents one's tendency to agree with items irrespective of their content. This total was then regressed on the final domain T scores for the five personality domains. The standardized residual from this analysis served as the AI. This index represents respondents' tendency to agree with items independent of their personality profile. The AI was then correlated with the SCID-IIP scales, and the results are presented in Table 1. As can be seen, for the college sample, 5 of the 13 scales showed significant correlations with the AI. However, these associations are all quite small, suggesting

that acquiescence may have played a small role in influencing these scores. For the clinical sample, correlations with the AI were significant for 10 of the 13 scales, suggesting that the positive wording of the symptom items may lead to an overendorsement of pathology in a clinical sample.

In order to examine the amount of overlap among these scales, a series of multiple regression analyses were conducted where each SCID-IIP scale was systematically regressed on the remaining scales. The resulting multiple  $R^2$ 's are presented in Table 1. As can be seen, for the college sample there is a high degree of overlap among these measures, with most of these scales sharing more than half of their reliable variance with the other scales (mean  $R^2 = .47$ ). An inspection of the regression coefficients from each analysis indicated that the shared variance was spread across scales from the different *DSM* personality disorder clusters. Each scale related to from 3 to 7 other scales. For the clinical sample, there is a comparable level of overlap (mean  $R^2 = .52$ ), and this finding continues to underscore the high degree of overlap among these scales. Some of this overlap may be a function of using only a single homogenous type of sample. It is also possible that this overlap reflects real redundancy among these diagnostic clusters.

Table 2 presents the partial correlations between the SCID-IIP screener scales and the five NEO-PI-R domain scales controlling for acquiescence. As can be seen, there are numerous, significant associations. For the college sample, of the 47 (out of 65) significant correlations, 29 (62%) were replicated with the SCID-IIP observer ratings. The hypothesized correlations with the FFM domains are indicated in bold. Of the 30 expected correlations, all but two were obtained. Thus, the *DSM-IV* version of the SCID-IIP screener correlated with these personality domains in a manner consistent with previous research. For the clinical sample, of the 30 expected correlations, 16 (53%) were found. The fewer number of expected significant correlations may be a function of the much smaller sample size. Nonetheless, there are a number of consistencies between these data and those found with the college sample. Both data sets show the correlation of (low) Extraversion in the Dependent and Depressive symptom scales; the association of Agreeableness on Obsessive-Compulsiveness and Depressive scales; and the correlation of Conscientiousness on the Avoidant, Dependent, and Depressive scales.

Finally, Table 2 presents the correlations between the personality disorder scales and self-reported scores on the BSI. As found in our previous study (Piedmont & Sherman, 1998), there are numerous relationships between the two measures, suggesting that individuals manifesting higher characterological symptoms also experienced higher levels of Axis I-type distress. This is particularly true for the Borderline, Paranoid, and Depressive disorder scales. This is true for both samples.

## DISCUSSION

Overall these results indicate that responses to the SCID-IIP questionnaire evidence a great deal of construct validity. The pattern of results with the FFM dimensions suggested that the SCID-IIP screener seems to be "personologically" equivalent to other Axis II disorder scales. The fact that many of these convergent correlations were replicated with the observer ratings lends confidence that the self-report findings are not merely the product of an acquiescence effect. Further, the level of cross-observer convergence found in the student sample on the SCID-IIP scales provides more confirmation of the value of this self-report measure of psychopathology. Given that the magnitude of cross-observer agreement was at a level consistently found with measures of normal personality suggest that the SCID-IIP screener may assess broad qualities of the individual that can be reliably identified by observers. The value of this finding is that it supports the usage of the SCID-IIP screener for more than just diagnostic assessment. The scale clearly captures salient "personological" qualities of individuals that may be useful in a number of nonclinical contexts (such as employment screening and personality research).

Finding significantly higher scores across all the symptom dimensions in the clinical sample provides additional validity to the responses of the SCID-IIP screener. Individuals presenting for treatment should manifest more symptomological problems than a general college sample. This provides some indirect support for the utility of the SCID-IIP as a screening device; scores vary in expected ways across psychologically different groups.

TABLE 2

Partial Correlation Between the Self-Reported Structured Clinical Interview for DSM-IV Personality Disorders Personality Questionnaire (SCID-IIP) Screener Scales and the Revised NEO Personality Inventory (NEO-PI-R) Domain Scores Controlling for Acquiescence

SCID-IIP Scale	NEO-PI-R Domain					BSI Global Symptom Index
	N	E	O	A	C	
College sample (N = 302)						
Avoidant	.50****	-.52****	-.19***	-.08	-.17**	.33***
Dependent	.43****	-.12	-.17**	-.03	-.32****	.32***
Obsessive-Compulsive	.32****	-.09	-.08	-.30***	.24****	.26***
Passive-Aggressive	.60****	-.23***	-.06	-.47****	-.33****	.44***
Depressive	.64****	-.41****	.04	-.36****	-.23****	.43***
Paranoid	.52****	-.31****	.00	-.55****	-.21***	.43***
Schizotypal	.29****	-.19***	.21****	-.31****	-.20****	.30***
Schizoid	.10	-.36****	-.11	-.22***	-.15	.13
Histrionic	.19***	.37****	.14	-.22****	-.17***	.13
Narcissistic	.33****	.00	.06	-.56****	-.16**	.32***
Borderline	.66****	-.22***	.07	-.48****	-.38****	.49***
Antisocial 1 (adult)	.32****	-.01	.10	-.51****	-.42****	.30***
Antisocial 2 (youth)	.21***	-.05	.15**	-.39****	-.23****	.22**
Clinical sample (N = 68)						
Avoidant	.38**	-.63****	-.24	.08	-.46****	.19
Dependent	.30**	-.26	-.02	-.26	-.42****	.39***
Obsessive-Compulsive	.05	-.23	-.09	-.32**	.09	.11
Passive-Aggressive	.54***	-.22	-.02	-.44***	-.44***	.44***
Depressive	.70***	-.48****	-.01	-.33**	-.64****	.58***
Paranoid	.38***	-.07	-.01	-.59***	-.14	.40***
Schizotypal	.10	-.17	.04	-.20	-.01	.29
Schizoid	-.15	-.43****	-.25	.06	.08	-.17
Histrionic	-.01	.24	-.11	-.25	.05	.18
Narcissistic	.14	.07	-.14	-.46****	-.09	.25
Borderline	.50***	-.05	.28**	-.36**	-.42****	.50***
Antisocial 1 (adult)	.31**	-.01	.17	-.27	-.46****	.28**
Antisocial 2 (youth)	.13	.00	-.03	-.16	-.20	.11

Note. N = Neuroticism; E = Extraversion; O = Openness to Experience; A = Agreeableness; C = Conscientiousness; BSI = Brief Symptom Inventory. Correlations in bold were predicted based on previous research.

\*Correlation replicated with SCID-IIP observer ratings.

\*\* $p < .01$ . \*\*\* $p < .001$ , two-tailed.

That the SCID-IIP scales correlated with the personality dimensions of the FFM in theoretically consistent ways provides not only sound construct validity evidence to the responses but also raises important theoretical issues surrounding conceptualizations of Axis II dynamics. Numerous researchers have argued that dimensional personality models, like the FFM, offer several conceptual and empirical advantages over the more traditional categorical approach to diagnosis taken by the DSM (Helmes & Jackson, 1994; Livesley, Schroeder, Jackson, & Jang, 1994; Widiger & Frances, 1994). Westen and Arkowitz-Westen (1998) argued that the DSM categorical system, in order to more fully address the range of pathology encountered in clinical practice, needs to be replaced by a dimensional system like the FFM. This would enable clinicians to evaluate pathology less as a "present/absent" quality and more as a functional assessment of impairment. The data presented in this study provide additional support for this approach. The consistent pattern of correlates with the FFM dimensions underscores the relatedness between normal and abnormal aspects of personality functioning. That the dimensions of the FFM have been linked to the types of specific personal problems individuals experience (Piedmont & Ciarrocchi, 1999; Piedmont, Sherman, &



Barrickman, 2000) suggests the hypothesis that Axis II disorders may be conceived as representing specific types of problems in living or as failures in adaptation (see Costa & McCrae, 1992b). Whatever the outcome from this line of reasoning, researchers in this area may find the SCID-IIP questionnaire to be an efficient and effective instrument to use in testing these types of formulations.

With regard to the Antisocial scale, both the Adult and Child versions appeared to manifest useful psychometric properties. The Adult version had much better variability and higher correlations with the other measures. On the other hand, it had a much higher level of overlap with the other SCID-IIP scales, and the Child version had a slightly higher alpha reliability. Thus, the use of the Likert-type format with the Adult version does not seem to have improved significantly the qualities of the scale in these samples.

One important area in which the results of these two samples differed concerned the associations between the SCID-IIP scales and the AI, where greater overlap appeared in the clinical sample. Thus, future research needs to examine to what extent the singular reliance on items phrased in one direction may lead to an overendorsement of symptoms. Response distortion may be more of a problem in a clinical sample than with a nonclinical sample, where one's psychological distress may predispose one to perceive a wide array of behaviors as pathological. Complicating this process is the presence of items which are themselves very redundant. Endorsing one item may lead to endorsing several others across different diagnostic categories. Therefore, acquiescence may seriously moderate one's ability to make accurate diagnostic judgments. Nonclinical samples may be able to make finer discriminations across the items, given their overall lower levels of distress. One limitation of the present study concerns differences in race and education levels (mostly Caucasian college-level sample vs. a mostly African American high school level sample), which may be one reason for this discrepancy. The impact of cultural and educational differences on symptom reporting and test responding patterns represents an area for future research to examine more closely. Another limitation is the different percentages of men and women in the two samples; the observed discrepancies may reflect actual gender differences in symptom reporting. Nonetheless, new items that are negatively reflected need to be developed and tested through a direct comparison to the current item set.

Another area to examine concerns the overlap among the various symptom scales. The data in both samples show that responses to these items share much of their reliable variance in common. These results are consistent with similar data obtained with the *DSM-III-R* version of the instrument (e.g., Piedmont & Sherman, 1998). Thus, despite efforts to create more orthogonal categories, the items seem to share much in common. This creates many obvious problems for making diagnostic assignments, because individuals scoring high on one scale are likely to score high on others. An inspection of the SCID-IIP's correlations with the FFM shows that these items share much in common with Neuroticism, indicating a noticeable amount of personological univocality. The taxonomic nature of the FFM may be helpful in identifying and creating new items with more diverse content.

Although the data from this study provide very encouraging empirical support for the SCID-IIP, more evidence is clearly needed. Future research will need to accomplish the following. First, demonstrate concurrent validity for responses to the scale, showing that it can accurately estimate the personality disorder diagnoses currently held by individuals in a clinical sample. Second, examine how racial, educational, and gender variables affect scores on the instrument. Third, conduct factor analyses of the responses to determine the underlying structure of the scale. Clearly the current data suggest that the SCID-IIP does not contain 12 distinct dimensions. However, would one find dimensions consistent with the current *DSM* disorder clusters? Fourth, examine the impact of acquiescence on scores in a clinical sample. Can the scale be improved by developing counterindicative items for each diagnostic scale? Would such items also facilitate differential diagnosis? It would also be fruitful to determine whether indices of acquiescence can be used to adjust scores on the scales. And fifth, which is related to the previous issue, examine the amount of empirical overlap among the scales. In the college student sample 63% of the reliable variance was held in common (range: 43% to 84%), whereas in the clinical sample 71% was overlapping (range: 52% to 97%). It appears that the *DSM-IV's* efforts at making the inclusion criteria for these Axis II groups to be more orthogonal has not been very successful. Such overlap works against accurate classification.

## CONCLUSION

The developers of the SCID-IIP screening questionnaire explicitly decried using the measure as a stand-alone clinical tool. The weight of evidence from this study as well as the data from evaluations of earlier versions of the scale indicate that this proscription may be too severe. The responses to the SCID-IIP questionnaire were found to be reliable and valid for use in both clinical and nonclinical samples. The SCID-IIP may be more than just a screening tool. Its correlation to measures of personality and symptom experience show it to capture salient psychological dimensions of individuals. It is interesting to note that symptom statements carry so much personological weight. The challenge to use these items in broader research contexts is certainly warranted by the findings of this study. The SCID-IIP questionnaire has properties that commend its use in a variety of nonclinical contexts. The scale is easy to use, and with 133 items is shorter than comparable instruments that are commercially available. Furthermore, the instrument is freely available, and the fact that the items are directly tied to the DSM diagnostic criteria makes research on the personality disorders with large samples eminently feasible. The SCID-IIP's overlap with the FFM makes it a potentially useful tool for counselors who are interested in obtaining a brief assessment of their clients on dimensions shown to be clinically salient (e.g., Miller, 1991). Finally, these data also show the value of an observer rating form to the SCID-IIP, opening the door for broader empirical assessment paradigms. The SCID-IIP questionnaire may have much wider value to the field than as just a screener.

## REFERENCES

- 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., revised). Washington, DC: Author.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Buckner, C. J., & Mandell, W. (1990). Risk factors for depressive symptomatology in a drug-using population. *American Journal of Public Health, 80*, 580-585.
- Carey, K. B. (1994). Use of the Structured Clinical Interview for DSM-III-R Personality Questionnaire in the presence of severe Axis I disorders: A cautionary note. *Journal of Mental Disease, 182*, 669-671.
- Coolidge, F. L., Becker, L. A., DiRito, D. C., Durham, R. L., Kinlaw, M. M., & Philbrick, P. B. (1994). On the relationship of the five-factor personality model to personality disorders: Four reservations. *Psychological Reports, 75*, 11-21.
- Costa, P. T., Jr., & McCrae, R. R. (1988). Personality in adulthood: A 6-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology, 54*, 853-863.
- 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). *Revised NEO Personality Inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., Jr., & McCrae, R. R. (1992b). The Five-Factor Model of personality and its relevance to personality disorders. *Journal of Personality Disorders, 6*, 343-359.
- Costa, P. T., Jr., McCrae, R. R., & Dye, D. A. (1991). Facet scales for Agreeableness and Conscientiousness: A revision of the NEO Personality Inventory. *Personality and Individual Differences, 12*, 887-898.
- Derogatis, L. R. (1993). *Brief Symptom Inventory, manual*. Minneapolis, MN: National Computer Systems.
- Derogatis, L. R., Rickels, K., & Rock, A. (1976). The SCL-90 and the MMPI: A step in the validation of a new self-report scale. *British Journal of Psychiatry, 128*, 280-289.
- Dreary, I. J., Peter, A., Austin, E., & Gibson, G. (1998). Personality traits and personality disorders. *British Journal of Psychology, 89*, 647-661.
- Ekselius, L., Linstrom, E., von Knorring, L., Bodlund, O., & Kullgren, G. (1994). SCID II interviews and the SCID screen questionnaire as diagnostic tools for personality disorders in DSM-III-R. *Acta Psychiatrica Scandinavica, 90*, 120-123.
- First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *User's guide for the Structured Clinical Interview for DSM-IV Axis I Personality Disorders (SCID-I)*. Washington, DC: American Psychiatric Press.

- Guthrie, P. C., & Mobley, B. D. (1994). A comparison of the differential diagnostic efficiency of three personality disorder inventories. *Journal of Clinical Psychology, 50*, 656-665.
- Helmes, E., & Jackson, D. N. (1994). Evaluating normal and abnormal personality using the same set of constructs. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 341-360). New York: Springer.
- Hyer, S. E., Rieder, R. O., Williams, J. B. W., & Spitzer, R. L. (1987). *Personality Diagnostic Questionnaire (PDS-R)*. New York: New York State Psychiatric Institute.
- Livesley, W. J., Schroeder, M. L., Jackson, D. N., & Jang, K. L. (1994). Categorical distinctions in the study of personality disorder: Implications for classification. *Journal of Abnormal Psychology, 103*, 6-17.
- Loranger, A. W., Susman, V. L., Oldham, J., & Russakoff, L. M. (1987). The Personality Disorder Examination: A preliminary report. *Journal of Personality Disorders, 1*, 1-13.
- McCrae, R. R., & Costa, P. T., Jr. (1992). Discriminant validity of the NEO PI-R facet scales. *Educational and Psychological Measurement, 52*, 229-237.
- McCrae, R. R., Herbst, J. H., & Costa, P. T., Jr. (2001). Effects of acquiescence on personality factor structures. In R. Riemann, F. Ostendorf, & F. Spinath (Eds.), *Festschrift in honor of Alois Angleitner* (pp. 217-232). Berlin: Pabst Science Publishers.
- McKeeman, J. L., & Erickson, M. T. (1997). Self and informant ratings of SCID-II personality disorder items for nonreferred college women: Effects of item and participant characteristics. *Journal of Clinical Psychology, 53*, 523-533.
- Miller, T. (1991). The psychotherapeutic utility of the five-factor model of personality: A clinician's experience. *Journal of Personality Assessment, 57*, 415-433.
- Ouimette, P. C., & Klein, D. N. (1995). Test-retest stability, mood-state dependence, and informant-subject concordance of the SCID-Axis II Questionnaire in a non-clinical sample. *Journal of Personality Disorders, 9*, 105-111.
- Piedmont, R. L. (1994). Validation of the NEO PI-R observer form for college students: Toward a paradigm for studying personality development. *Assessment, 1*, 259-268.
- Piedmont, R. L. (1998). *The Revised NEO Personality Inventory: Clinical and research applications*. New York: Plenum Press.
- Piedmont, R. L., & Ciarrocchi, J. W. (1999). The utility of the Revised NEO Personality Inventory in an outpatient, drug rehabilitation context. *Psychology of Addictive Behaviors, 13*, 213-226.
- Piedmont, R. L., & Sherman, M. F. (1998, April). *Psychometric utility of the SCID-IIP screening questionnaire*. Paper presented at the Annual Convention of the Eastern Psychological Association, Providence, RI.
- Piedmont, R. L., Sherman, M. F., & Barrickman, L. (2000). Brief psychosocial assessment of a clinical sample: An evaluation of the Personal Problems Checklist for Adults. *Assessment, 7*, 177-187.
- Piedmont, R. L., Sherman, M. F., & Murphy, P. (1999). *A psychometric evaluation of the SCID-IIP self-report screening questionnaire*. Unpublished manuscript.
- Piedmont, R. L., & Weinstein, H. P. (1993). A psychometric evaluation of the new NEO PI-R facets scales for Agreeableness and Conscientiousness. *Journal of Personality Assessment, 60*, 302-318.
- Royse, D., & Drude, K. (1984). Screening drug abuse clients with the Brief Symptom Inventory. *International Journal of Addictions, 19*, 849-857.
- Shea, M. T., Widiger, T. A., & Klein, M. H. (1992). Comorbidity of personality disorders and depression: Implications for treatment. *Journal of Consulting and Clinical Psychology, 60*, 857-868.
- 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.
- Westen, D., & Arkowitz-Westen, L. (1998). Limitations of Axis II diagnosing personality pathology in clinical practice. *American Journal of Psychiatry, 155*, 1767-1771.
- Widiger, T. A., & Costa, P. T., Jr. (1994). Personality and personality disorders. *Journal of Abnormal Psychology, 103*, 78-91.
- Widiger, T. A., & Frances, A. (1994). Toward a dimensional model for the personality disorders. In P. T. Costa, Jr., & T. A. Widiger (Eds.), *Personality disorders and the Five-Factor Model of personality* (pp. 19-40). Washington, DC: American Psychological Association.
- Widiger, T. A., & Rogers, J. H. (1989). Prevalence and comorbidity of personality disorders. *Psychiatric Annals, 19*, 132-136.
- Woody, G. E., McLellan, A. T., Luborsky, L., & O'Brien, C. P. (1985). Sociopathy and psychotherapy outcome. *Archives of General Psychiatry, 42*, 1081-1086.