Examining the Reliability and Validity of Clinician Ratings on the Five-Factor Model Score Sheet

Lauren R. Few[1], **Joshua D. Miller**[1], **Jennifer Q. Morse**[2], **Kirsten E. Yaggi**[2], **Sarah K. Reynolds**[2], and **Paul A. Pilkonis**[2]

[1]University of Georgia, Athens, GA, USA

[2]Western Psychiatric Institute and Clinic, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

Abstract

Despite substantial research use, measures of the five-factor model (FFM) are infrequently used in clinical settings due, in part, to issues related to administration time and a reluctance to use self-report instruments. The current study examines the reliability and validity of the Five-Factor Model Score Sheet (FFMSS), which is a 30-item clinician rating form designed to assess the five domains and 30 facets of one conceptualization of the FFM. Studied in a sample of 130 outpatients, clinical raters demonstrated reasonably good interrater reliability across personality profiles and the domains manifested good internal consistency with the exception of Neuroticism. The FFMSS ratings also evinced expected relations with self-reported personality traits (e.g., FFMSS Extraversion and Schedule for Nonadaptive and Adaptive Personality Positive Temperament) and consensus-rated personality disorder symptoms (e.g., FFMSS Agreeableness and Narcissistic Personality Disorder). Finally, on average, the FFMSS domains were able to account for approximately 50% of the variance in domains of functioning (e.g., occupational, parental) and were even able to account for variance after controlling for Axis I and Axis II pathology. Given these findings, it is believed that the FFMSS holds promise for clinical use.

Keywords

DSM-5; personality disorders; assessment

One of the most widely researched and accepted models of personality is the five-factor model (FFM; Digman, 1990), which includes the following five broad dimensions: Neuroticism (vs. Emotional Stability), Extraversion (vs. Introversion), Openness (vs. Closedness to Experience), Agreeableness (vs. Antagonism), and Conscientiousness (vs. Disinhibition). In addition to these five broad domains, Costa and McCrae (1992) delineated six underlying facets subsumed by each of the FFM domains and assessed in their popular measure of the FFM, the Revised NEO Personality Inventory (NEO PI-R; Costa & McCrae, 1992). Although the FFM is a well-established model of personality that has proven helpful in integrating findings from various personality models, conceptualizing personality disorder (PD), and bridging the divide between "normal" and "abnormal personality," there is a smaller body of research addressing the clinical utility of this model. The limited research

Corresponding Author: Joshua D. Miller, Department of Psychology, University of Georgia, Athens, GA 30602-3013, USA jdmiller@uga.edu.

that is available, however, suggests that it has promise as it has been demonstrated that measures of the FFM predict treatment satisfaction and compliance (Miller, Pilkonis, & Mulvey, 2006), as well as functional impairment (e.g., Hopwood et al., 2009; Miller, Pilkonis, & Clifton, 2005). In addition, numerous studies have documented substantial relations between the FFM and Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (*DSM-IV*; American Psychiatric Association, 1994) PDs (see Costa &Widiger, 2002). Finally, clinicians view it as a useful model for conceptualizing personality pathology (Samuel & Widiger, 2006).

Despite its potential utility, clinicians have been slow to adopt measures of the FFM in their clinical work. There may be several reasons for this reluctance. First, the most popular measure of the FFM, the NEO PI-R (Costa & McCrae, 1992), is a 240-item measure that takes approximately 20 to 40 minutes to complete.[1] The NEO PI-R can be completed using self- or informant ratings or interview (Structured Interview for the Five-Factor Model [SIFFM]; Trull & Widiger, 1997), but the self-report methodology is most common. Although both self and informant ratings of personality demonstrate predictive validity (Klein, 2003), informant reports often provide stronger predictive or incremental validity over self-report data (Miller et al., 2005). Another issue that may limit the perceived clinical utility of FFM assessments is that these measures do not assess for impairment in relation to the various traits. This omission is significant as it has been argued that an explicit assessment of impairment is necessary when assessing PDs (Clark, 2007). Thus, despite the evidence that personality data derived from an FFM assessment show promise with regard to patient conceptualization, prediction of treatment outcomes, and level of impairment, the current assessment methodology may be regarded as too unwieldy and limited to warrant frequent use in clinical settings.

Given these concerns, several alternative assessment strategies have been developed that address some of these problems. For instance, the SIFFM (Trull & Widiger, 1997) is an interview-based method for assessing the traits consistent with the NEO PI-R that also includes an assessment of the dysfunction associated with each trait, which may increase its clinical utility. Unfortunately, the length of this interview may have limited its adoption in clinical settings. A recently developed measure, the FFM Rating Form (FFMRF; Widiger, 2004), has addressed concerns regarding brevity and the need to include lower order facets in other-rated assessment of the FFM. The FFMRF uses a single-item to assess each of the 30 facets of the FFM by including an identifying term for each facet (e.g., self-consciousness), along with two to four adjectives that describe both poles of each facet (e.g., timid, embarrassed vs. self-assured, glib, shameless). The FFMRF has been reliably used to describe prototypical cases of PD in terms of the FFM (Lynam & Widiger, 2001; Samuel & Widiger, 2004), and self-reports on the FFMRF demonstrated good convergent validity with self-reports on the NEO PI-R among undergraduate samples (Mullins-Sweatt, Jamerson, Samuel, Olson, & Widiger, 2006). The FFMRF has also been used to compare the clinical utility of the FFM with the *DSM-IV* model (Lowe & Widiger, 2009).

Another brief measure of the FFM is the FFM Score Sheet (FFMSS; Widiger & Spitzer, 2002), which is an earlier version of the FFMRF that also uses a single rating for each of the 30 personality facets captured in the NEO PI-R.[2] Although the content of both the FFMSS and FFMRF is largely the same, the two differ in that the FFMSS uses a scale ranging from

---

[1]Several abbreviated measures of the five-factor model (FFM) have been developed. These include the Big Five Inventory (BFI; John & Srivastava, 1999), the NEO-Five Factor Inventory (NEO-FFI; Costa & McCrae, 1992), Saucier's (1994) Mini Markers, and the 10-Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003). One limitation of these measures, however, is that they do not assess the facets of the FFM (Mullins-Sweatt, Jamerson, Samuel, Olson, & Widiger, 2006). This is problematic in that the FFM facets help differentiate between both Axis I and Axis II disorders and are better able to capture the different personality disorders (PDs; see Lynam & Widiger, 2001).

1 (*problematic, very low on the trait*) to 7 (*problematic, very high on the trait*; see the appendix), whereas the FFMRF uses a 5-point scale ranging from *extremely low* to *extremely high*. The FFMSS also allows for ratings that convey high or low levels of the trait that are not problematic (e.g., 3, *low on the trait*; 5, *high on the trait*). This difference is noteworthy because the FFMSS scale includes indications of poor adaptations at the poles of each FFM facet; this difference may be important in that previous research suggests that the NEO PI-R does not contain much content related to the maladaptive poles of certain dimensions (e.g., maladaptively high Agreeableness; Haigler & Widiger, 2001). Haigler and Widiger demonstrated that rather straightforward alterations of some of the NEO PI-R content to increase the description of maladaptivity at the high end of certain dimensions (i.e., Openness, Agreeableness, Conscientiousness) resulted in stronger correlations among certain domains (i.e., Conscientiousness) and PDs (i.e., Obsessive–Compulsive PD).

The goals of the current study were multifold. First, we tested the interrater reliability and internal consistency (domains only) of the ratings derived from the FFMSS. Second, we examined which FFM traits were more/less difficult for clinicians to rate reliably. Third, we examined the convergent and discriminant validity of the FFMSS ratings in relation to self-report personality traits and expert ratings of the *DSM-IV* PDs. Fourth, we examined the utility of the FFMSS ratings by examining their relations with expert ratings of functional impairment (e.g., interpersonal, occupational). We also tested whether the FFMSS ratings provided incremental validity in the statistical prediction of these impairment domains above and beyond that provided by ratings of depression, anxiety, and *DSM-IV* PD symptoms. Also unique to the current study is that it provided the first test of these single-item facet ratings for the assessment of the FFM in a clinical setting. In previous research with the FFMRF, the measure has been used to generate ratings of prototypical PDs (Samuel & Widiger, 2004), PD vignettes (Sprock, 2002), or self-report descriptions of undergraduates (Mullins-Sweatt et al., 2006). We expand on each of these goals below.

As noted above, in the current study, we investigated which FFM traits were most difficult to rate using the FFMSS. Research examining self- and other ratings of personality may be valuable in evaluating the potential difficulty for clinicians in rating FFM dimensions using the FFMSS. For example, in a clinical sample of predominantly depressed individuals, Bagby et al. (1998) found that Openness and Neuroticism were rated less reliably relative to the other FFM domains, and Conscientiousness and Agreeableness were rated most reliably, followed by Extraversion. A similar pattern of findings is hypothesized here.

In the current study, we also tested the convergent and discriminant validity of the FFMSS in relation to personality as measured by the Schedule for Nonadaptive and Adaptive Personality (SNAP; Clark, 1993) trait and temperament scales and expert rating of *DSM-IV* PD symptoms. The relations between the FFMSS and the SNAP were expected to be largely consistent with those found by Clark, Vorhies, and McEwen (2002). For example, FFMSS Extraversion should correlate significantly with SNAP Positive Temperament, Exhibitionism, Detachment, Entitlement, and Impulsivity, whereas FFMSS Neuroticism should be significantly related to SNAP Negative Temperament, Self-Harm, Mistrust, Dependency, and Aggression. FFMSS Conscientiousness should be significantly related to Disinhibition, Workaholism, Impulsivity, Manipulativeness, Dependency, and Propriety. The additional FFM factors with no explicit counterpart—Agreeableness and Openness—should also demonstrate significant correlations with certain SNAP scales. For instance, FFMSS Agreeableness should be related to SNAP Aggression, Mistrust, Manipulativeness,

---

[2]The Five-Factor Model Score Sheet ([FFMSS] vs. the FFM rating Form [FFMRF]) was the only measure available at the time that this study was initiated.

and Detachment. Finally, FFM Openness should be positively related to SNAP Eccentric Perceptions and Impulsivity and negatively related to Propriety.

With regard to the relations between the FFMSS and *DSM-IV* PDs, we expected that the findings should be consistent with theoretical conceptualizations of PDs from a FFM perspective (Lynam & Widiger, 2001). For instance, individuals with greater Borderline PD symptoms should be rated higher on the FFMSS Neuroticism facets of anxiety, anger–hostility, depression, impulsiveness, and vulnerability, and they should be rated lower on FFMSS Agreeableness facets of trust and compliance, in addition to the competence facet of Conscientiousness. To quantify these relations, we examined the similarity of the FFMSS–*DSM*-IV PD correlates found in the current sample with the hypothesized correlates between the FFM and *DSM-IV* PDs generated by expert raters (Lynam & Widiger, 2001).

Finally, we examined the association between FFMSS ratings and impairment (i.e., romantic relationships, parenting, other social relationships, occupational impairment, distress caused to significant others, and overall impairment). Consistent with previous results, we expected that the impairment domains would be positively related to Neuroticism and negatively related to Extraversion, Agreeableness, and Conscientiousness (e.g., Hopwood et al., 2009; Miller et al., 2005). The results from these analyses are important in testing the clinical utility of the FFMSS ratings; that is, do clinicians' ratings of these personality traits correspond to functional impairment in a variety of different domains? We also tested the incremental validity of the FFMSS by evaluating whether it accounted for additional variance in the prediction of impairment beyond Axis I symptoms (i.e., depression, anxiety) and Axis II PD symptoms. This analysis provided important information regarding whether the FFMSS ratings provide unique information that is not captured by the current *DSM* constructs as has been demonstrated previously with FFM information derived from other sources (i.e., Miller et al., 2006).

Method

Participants and Procedure

Participants included 130 outpatients (97 females; 33 males) from Western Psychiatric Institute and Clinic (WPIC) in Pittsburgh, Pennsylvania between the ages of 21 and 60 years ($M = 37.9$, $SD = 10.6$). The majority of the participants were White (74.6%); of the remaining participants, the largest percentage was that of African American (23.8%). A total of 50% were single, 26.9% were either separated or divorced, 16.1% were married, 5.4% lived with a long-term domestic partner, and 1.5% were widowed. With regard to education, the majority of the sample (81.5%) had some college education or higher (i.e., completed college; graduate/professional training). Axis I ratings were determined by expert consensus following a structured clinical interview. In all, 92 participants (71%) had a current mood disorder diagnosis and 61 (55%) had a current anxiety disorder. A total of 102 participants (78%) met criteria for a *DSM-IV* PD; the most prevalent PD diagnoses were Borderline (45%) and Avoidant (35%). On average, each participant met criteria for 1.9 PDs. Overall, the prevalence rates of the PDs were consistent with the initial purpose of the study, which was an investigation of interpersonal functioning in patients with Borderline PD and two comparison groups: patients with Avoidant PD and patients with Axis I disorders but no PD. During the initial recruitment of the sample, it became clear that individuals with many other PDs were responding to the study advertisements, and ultimately a fourth group of patients with other PDs was included. Therefore, the overall sample included a substantial degree of heterogeneity with regard to PD symptoms, although the Cluster A PDs were the least prevalent. Interested participants contacted the research staff and were screened via telephone to determine whether they met any of the following exclusion criteria: psychotic disorders, organic mental disorders, mental retardation, and major medical illnesses that

influence the central nervous system and might be associated with organic personality disturbance (see Morse et al., 2009, for more sample details).

Eligible participants provided written consent after all study procedures had been explained. Subsequently, they were assessed by a primary interviewer across a minimum of three sessions lasting approximately 2 to 3 hours each. Session 1 included the administration of the Structured Clinical Interview for *DSM-IV* Axis I Disorders (SCID-I; First, Gibbon, Spitzer, & Williams, 1997) and other measures of current symptomatology. In Session 2, a detailed social and developmental history was completed. During Session 3, the Structured Clinical Interview for *DSM-IV* Axis II PDs (SCID-II; First et al., 1997) was administered. The participants completed self-report measures (e.g., SNAP) in between assessment sessions.

The Longitudinal Expert Evaluation Using All Data (LEAD) method (Spitzer, 1983) was used in determining all consensus ratings. This method emphasizes the contribution of expert clinical judgment, but includes the use of multiple information sources in arriving at that judgment. These sources included assessment interviews with the patient as well as judgments of other professionals. Following the assessment sessions, the primary interviewer presented the case at a 3- to 4-hour diagnostic conference with colleagues from the research team. A minimum of three judges participated.[3] All available data were reviewed and discussed at the conference. Judges were given access to all data that had been collected: current and lifetime Axis I information, symptomatic status, social and developmental history, and personality features acknowledged on the Axis II interviews. The self-report personality data collected with the SNAP were not available for the LEAD ratings. The relevant data for the current study that were derived from the case conference include (a) consensus ratings of *DSM-IV* PD criteria and (b) consensus ratings on impairment variables. The FFMSS ratings were typically completed by the primary interviewer prior to the case conference, whereas the secondary FFMSS rater (where available) did so following the completion of the case conference (the FFMSS ratings were not discussed during the case conference). All FFMSS raters received basic training in relation to the FFM and orientation to the FFMSS, which involved reviewing the FFM facets using descriptions from Costa and McCrae's (1992) NEO PI-R manual. Raters also received copies of these descriptions that could be used as a reference when completing the FFMSS.

Measures

Five-Factor Model Score Sheet—The FFMSS (Widiger & Spitzer, 2002; see the appendix) is a one-page rating sheet consisting of 30 items representing each of the 30 facets of the FFM, as conceptualized in the NEO PI-R. These facets are organized with respect to the FFM domains, such that there are six items beneath a listed domain. Each item includes a list of two to four adjectives describing the trait. For example, high straightforwardness, a facet of Agreeableness, is described with the adjective "naive" whereas low straightforwardness is described using the adjective "deceptive." Each item is rated on a 1 (*problematic, very low on the trait*) to 7 (*problematic, very high on the trait*) scale. Of the 130 participants, 112 were rated by two raters.

Schedule for Nonadaptive and Adaptive Personality—The SNAP (Clark, 1993) is a 375-item, true–false inventory that assesses 15 traits relevant to PD: 12 lower order primary

---

[3]The team of interviewers and expert judges comprised licensed psychologists, senior doctoral students (i.e., completing internship training), and master's-level clinicians (this team was composed of both male and female raters). Expert raters, outside of the primary interviewer, did not typically have any direct contact with the participants. Outside of the second author and fifth authors (both of whom provided a very small percentage of the primary ratings used in most of the current analyses), the interviewers and judges tended to have limited familiarity with the FFM.

traits and three broad temperament dimensions—Negative Temperament, Positive Temperament, and Disinhibition (vs. Constraint). Alpha coefficients in the current study ranged from .71 (Disinhibition) to .91 (Aggression). The SNAP traits have shown to be related to *DSM-IV* PDs in theoretically expected directions (e.g., Reynolds & Clark, 2001).

Consensus ratings of DSM-IV PD criteria—These ratings were determined in each participant's case conference. A consensus rating of each *DSM-IV* PD symptom was determined using a 0-2 scale, with 0 indicating *absent*, 1 indicating *present*, and 2 indicating *strongly present*. Symptom counts for each participant were generated by adding all scores (i.e., 0, 1, 2) for each PD. Alpha coefficients for the PDs ranged from .53 (Dependent PD) to .88 (Avoidant PD), with a median of .73.

Ratings of depression and anxiety—These ratings were conducted with the Hamilton Depression Rating Scale (HAM-D; Hamilton, 1960) and the Hamilton Anxiety Rating Scale (HAM-A; Hamilton, 1959) by the primary interviewer for each participant. Studies have supported the reliability of HAM-D ratings by clinicians (Carroll, Fielding, & Blashki, 1973). The HAM-A has also been shown to be both reliable and valid when used with individuals with depressive disorders and anxiety disorders (Maier, Buller, Philipp, & Heuser, 1988).

Consensus ratings of impairment—Consensus ratings were determined separately for romantic relationships, parenting, other social relationships (e.g., friends, family members), occupational impairment, distress caused to significant others (e.g., friends, children), and overall impairment using a 1 (*exceptionally positive functioning*) to 9 (difficulties are persistent and pervasive, without clearly identifiable elements of functioning relevant to the domain) one-item scale with higher scores indicative of greater impairment.

Results

Prior to completing reliability and validity analyses, descriptive statistics for the FFMSS were examined. Means and standard deviations for the facets and domains are presented in Table 1. Intercorrelations between FFMSS domains were also computed. Several analyses were conducted to examine the reliability of the FFMSS. First, internal consistency of the FFMSS domains was examined. Second, individual profile agreement, or reliability of ratings for each participant on the FFMSS facets, was determined. Third, interrater reliability of ratings across the FFMSS domains and facets was examined.

Following the reliability analyses, several analyses were conducted to examine convergent and divergent validity. The FFMSS ratings (using only the primary raters' scores) were examined in relation to (a) SNAP scales, (b) consensus *DSM-IV* PD ratings, and (c) consensus impairment ratings across a variety of domains (e.g., romance, work). To control for Type I error, we lowered our significance level to p .001 for all analyses.

Internal Consistency and Interrater Reliability

Coefficient alphas for the FFMSS domains were .61, .88, .87, .86, and .92 for Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness, respectively. More detailed analysis of the Neuroticism domain revealed that the angry hostility and impulsiveness facets were either weakly or negatively related to the other four Neuroticism facets, thus decreasing the internal consistency of this domain. The angry hostility facet was most strongly correlated with Agreeableness ($r = -.66$) and the impulsiveness facet was most strongly correlated with Conscientiousness ($r = -.70$).

To examine individual profile agreement, a double-entry intraclass correlation coefficient (ICC$_{DE}$) was used as it has proven to be a superior measure for calculating profile agreement (see McCrae, 2008 for a review). The ICC$_{DE}$ was computed for 112 of the 130 participants (all those who had data from two raters). One benefit of this approach is that it takes into consideration absolute agreement of the ratings (rather than just the relative agreement) generated by each rater. To do this, each of the raters' 30 facet ratings for a given participant was entered. Then those same ratings were pasted into the other rater's column (e.g., Rater 1 has his or her 30 FFMSS ratings of a participant listed followed by Rater 2's 30 FFMSS ratings of that same participant in column 1; Rater 2 has his or her 30 FFMSS ratings of that participant listed followed by Rater 1's 30 FFMSS ratings of that same participant in column 2). Following this data manipulation, a Pearson correlation was computed between these columns (with 60 data points in each column). For the 112 participants, the interrater reliability coefficients for the entire individual FFM profiles ranged from –.31 to .92 with a median coefficient of .58.

We also examined which traits were more difficult to rate reliably using the same double-entry method (Table 1). Coefficients for the facets ranged from .30 (depression) to .62 (gregariousness, excitement-seeking, activity) with a median coefficient of .52. Coefficients for the domains ranged from .55 (Neuroticism) to .72 (Conscientiousness) with a median of .66.

Intercorrelations of FFMSS Domains

Correlations between the FFMSS domains ranged from –.36 (Neuroticism and Extraversion) to .49 (Extraversion and Openness; see Table 2). To determine whether the results were similar to those found in previous research with the NEO PI-R, the intercorrelations in the current study were compared with the intercorrelations of the NEO PI-R domains reported in the NEO PI-R normative sample (Costa & McCrae, 1992). To do this, the two sets of correlations were tested to see if there were any significant differences (i.e., test for independent rs; Cohen and Cohen, 1983). Of the 10 pairs of correlations (e.g., FFMSS Neuroticism and FFMSS Extraversion vs. NEO PI-R Neuroticism and NEO PI-R Extraversion), two were significantly different (z 3.44, p .001): Conscientiousness and Neuroticism (NEO PI-R, –.53; FFMSS, –.36) and Conscientiousness and Extraversion (NEO PI-R: .27; FFMSS: –.05) such that the FFMSS domains appear to be more orthogonal than the NEO PI-R domains.

Convergence of FFMSS Dimensions and SNAP Traits

Next, correlations between the FFMSS domains and the SNAP traits were examined (see Table 3). As expected, FFMSS Neuroticism, Extraversion, and Conscientiousness demonstrated significant correlations with their SNAP counterparts (Negative Temperament, $r = .45$; Positive Temperament, $r = .52$; Disinhibition, $r = –.32$). With regard to all the SNAP scales, FFMSS Neuroticism was significantly positively related to Self-Harm, Dependency, and Detachment and negatively related to SNAP Positive Temperament. FFMSS Extraversion was significantly positively related to SNAP Exhibitionism and negatively related to Detachment. Significant positive correlations emerged between FFMSS Openness and SNAP Eccentric Perceptions and Impulsivity, whereas significant negative correlations emerged between FFMSS Agreeableness and SNAP Aggression. Last, FFMSS Conscientiousness was positively related to SNAP Workaholism and negatively related to SNAP Negative Temperament, Manipulativeness, Aggression, Self-harm, and Eccentric Perceptions. Overall, FFMSS Neuroticism and Extraversion demonstrated the largest correlations with Positive Temperament (–.47 and .52, respectively). FFMSS Openness manifested its largest correlation with Eccentric Perceptions (.32), whereas FFMSS

Agreeableness and Conscientiousness evinced the largest correlations with Aggression (–.38) and Manipulativeness (–.34), respectively.

Relations Between FFMSS Traits and Consensus DSM-IV PD Ratings

Correlations were next computed between FFMSS traits and *DSM-IV* PD ratings (see Table 4). At the domain level, Neuroticism was significantly positively correlated with 3 of 10 PDs: Borderline, Avoidant, and Dependent. Extraversion manifested one significant positive correlation (Histrionic) and two significant negative correlations (Schizoid, Avoidant). Openness was significantly positively correlated with two PDs (i.e., Borderline, Histrionic) and significantly negatively correlated with Avoidant PD. Agreeableness manifested significant negative correlations with four of the *DSM-IV* PDs (i.e., Paranoid, Antisocial, Borderline, and Narcissistic). Last, Conscientiousness manifested three significant negative correlations (i.e., Antisocial, Borderline, and Histrionic) and was positively correlated with Obsessive–Compulsive PD. Domain and facet-level correlations for each of the 10 consensus PDs are presented in Table 4.

To quantify the match between the FFMSS profile for each *DSM-IV* PD with expert conceptualizations of *DSM-IV* PDs from a FFM perspective, the trait profiles for each consensus PD (i.e., correlations between the FFMSS and the consensus ratings for the *DSM-IV* PD symptoms) were compared with FFM descriptions of the PDs generated by experts (Lynam & Widiger, 2001). This involved correlating the set of Pearson *r*s between FFMSS traits and *DSM-IV* PDs with the expert FFM prototypes for each PD. These second-order correlations ranged from .47 (Paranoid) to .95 (OCPD) with a median of .85 (Table 4, final row). Overall, the FFMSS–*DSM-IV* PD correlates were quite similar to the expert descriptions of prototypical individuals with each PD.

Relations Between FFMSS Traits and Impairment

The utility of the FFMSS ratings was evaluated by examining the relations between FFMSS traits and consensus impairment ratings (see Table 5). As expected, FFMSS Conscientiousness (six of six significant *r*s), Agreeableness (five of six significant *r*s), and Neuroticism (five of six significant *r*s) were the most consistent FFM correlates of impairment. FFMSS Extraversion was negatively related to overall impairment and, more specifically, social impairment, whereas the Openness domain was unrelated to the impairment domains.

Also presented in Table 5 are results from simultaneous multiple regression analyses that were conducted to examine the amount of variance in the impairment variables accounted for by the FFMSS domains. The domains accounted for a significant amount of variance in each of the six impairment variables, ranging from 23% (romantic) to 60% (distress caused to significant others), with a mean of 46%. Next, hierarchical multiple regression analyses were conducted to examine the incremental validity of the FFMSS scales (entered at Step 2) beyond symptoms of depression, anxiety, and Axis II psychopathology (entered at Step 1). The FFMSS domains accounted for a significant amount of additional variance in four of the six impairment variables.[4] Specifically, the domains accounted for an additional 13% of the variance in distress caused to significant others and social impairment, an additional 15% of the variance in occupational impairment, and an additional 10% of the variance in the overall impairment score. Across these analyses, Conscientiousness (negative) was the most consistently significant FFM unique incremental predictor (i.e., four of four analyses).[5]

---

[4]There was some evidence of multicollinearity in these hierarchical models; however, we were less concerned with the significance of the individual betas than with the overall change in variance explained by the set of FFM domains.

Discussion

The goal of the current study was to evaluate the reliability and validity of the FFMSS (Widiger & Spitzer, 2002). Previous research has identified a number of reliable and valid measures of the FFM (e.g., NEO PI-R, NEO-FFI, Costa & McCrae, 1992; FFMRF, Widiger, 2004). There are various aspects of these measures, however, which may contribute to their limited use in clinical settings, such as administration time (e.g., NEO PI-R; SIFFM), self-report methodology (e.g., NEO-FFI; NEO PI-R), and limited breadth (i.e., assess domains only: NEO-FFI). Given the likelihood that the *DSM-V* may include a dimensional trait model of personality for use in the diagnosis and description of personality pathology, it will be important to develop a clinically reliable, valid, and useful (e.g., efficient) measure of traits such as those found in the FFM. The current study extended research in the assessment of the FFM by evaluating a clinician-rated, single-item rating form of the 30 FFM facets.

Internal Consistency and Inter-rater Reliability

In general, internal consistency was acceptable in the current study, with the exception of the Neuroticism domain (i.e., $\alpha = .61$). Interitem correlations indicated that the Neuroticism facets of angry hostility and impulsiveness were either minimally or negatively related to the other four Neuroticism facets. This finding is consistent with FFM research examining the factor structure of the FFM, in that NEO PI-R angry hostility manifests a significant secondary loading on the Agreeableness domain, whereas impulsiveness manifests a substantial secondary loading on Conscientiousness factors (Costa & McCrae, 1992). Raters may have had a difficult time conceptualizing the experience of anger as being distinct from acting toward others in an angry, hostile manner. Similarly, although neurotic impulsiveness (e.g., difficulty resisting cravings and urges) is different than impulsivity conceptualized via Conscientiousness (Whiteside & Lynam, 2001), it may have been difficult for clinical raters with limited knowledge of the FFM to make this nuanced distinction. The label of this Neuroticism facet (i.e., "Impulsiveness") may have inadvertently caused raters to associate it with the traits from Conscientiousness. It is interesting to note that Whiteside and Lynam (2001) named an impulsivity-related trait based on this facet "Urgency," which may be both more descriptive and help raters understand how it differs from less-affectively focused forms of impulsivity.

The median interrater reliability coefficient in the current study was .58 when examining profile agreement across raters for each participant. This is consistent with previous reports of profile agreement in studies examining ratings of PDs using the FFM facets. For example, Sprock (2002) reported an ICC of .67 when examining interrater reliability for clinician ratings of prototypical and nonprototypical PD cases for each of the three PD clusters. Overall, it is promising that adequate interrater reliability was obtained when examining profile agreement using only ratings from those who received minimal training with regard to these general personality traits. We believe that it is possible that more intensive, structured training could substantially increase the interrater reliability of this instrument. The limited training implemented in the current study provided a conservative test of the utility of the FFMSS. In fact, the interrater reliability and convergent validity found in the current study may be indicative of the lower bounds of these statistics.

Interrater reliability of the individual FFMSS facets and domains was also examined to determine which traits were more reliably rated by clinicians. Based on previous research examining self–other convergence in FFM ratings (e.g., Bagby et al., 1998), it was

---

[5]We also tested whether the relations between the FFM domains and the impairment outcomes were curvilinear; there were no instances of statistically significant curvilinearity.

hypothesized that Openness and Neuroticism would be the most difficult domains to rate reliably, whereas Conscientiousness, Agreeableness, and Extraversion would evince the greatest interrater reliability. Results supported these hypotheses, in that Neuroticism and Openness manifested the lowest interrater reliability, whereas Conscientiousness, Agreeableness, and Extraversion were the domains rated most reliably. These findings are also consistent with previous research examining the interrater reliability of "thin-slice" ratings of personality (Oltmanns, Friedman, Fiedler, & Turkheimer, 2004). Specifically, when rating military recruits, strangers rated Neuroticism least reliably and Extraversion most reliably.

At the facet level, the median ICC was .52. Despite fair to good reliability across the FFMSS domains and 27 of 30 facets, interrater reliability for the FFMSS in the current study was lower than the SIFFM (Trull et al., 1998), an interview-based measure of the FFM. This is not surprising as the SIFFM was designed to assess these traits directly, whereas the FFMSS was completed without direct assessment of these traits. The three facets that exhibited the poorest inter-rater reliability were depression (.30), aesthetics (.38), and achievement-striving (.37), which manifested similarly low correlations between two peer raters (Costa & McCrae, 1992). It is possible that depression was rated less reliably because one rater interviewed the patient, whereas the other made ratings on the basis of case conference information. Given the nature of depression, face-to-face contact with the patient may have influenced the ratings (e.g., assessment of flat or depressed affect; crying) and led to lower convergence. The lower interrater reliability for the traits of aesthetics and achievement-striving facets may be due to the fact that content related to these traits, particularly aesthetics, may not be typically gleaned from initial clinical interviews, making them more difficult to rate.

Intercorrelations

The FFMSS is also novel in that it allows for an assessment of maladaptive functioning at both poles of each domain. We found no evidence that the overall structure of the measure was substantially altered by the current methodology as the intercorrelations between the domains were relatively consistent with NEO PI-R intercorrelations from the normative data (Costa & McCrae, 1992), which suggests that the inclusion of maladaptive poles did not alter the manner in which the FFM domains relate to one another in a fundamental manner. In fact, the two correlations that differed significantly from the pattern reported in Costa and McCrae's (1992) normative data (i.e., Conscientiousness and Neuroticism; Conscientiousness and Extraversion) differed such that the FFMSS correlations were smaller than that found for the FFM suggesting that the FFMSS results in a more orthogonal pattern of interrelations.

Convergence Between FFMSS Traits and SNAP Personality Traits

Validity of the FFMSS was examined, in part, by testing its relations to personality traits, as measured by the self-reported scores on the SNAP. At the higher order level, the SNAP temperament dimensions were significantly correlated with FFM domains as predicted, although the magnitude of the correlations were smaller than we expected. The current relations are likely smaller because of the use of different assessment methodologies and, as such, they do not benefit from shared method variance. Consistent with Clark et al.'s (2002) findings, discriminant correlations were typically lower than convergent correlations.

FFMSS Traits and Personality Disorders

The FFM has gained prominence in the field of psychology not only for its descriptive utility but also for its validity and utility in the conceptualization and prediction of PDs. To examine the validity of the FFMSS, we examined whether the FFMSS was able to capture

the range of personality pathology in a manner comparable with other FFM measures. Consensus *DSM-IV* PD ratings were correlated with the FFMSS to generate FFM profiles of each *DSM-IV* PD. The results of these analyses provide relatively strong support for the use of the FFMSS in conceptualizing and predicting *DSM-IV* PDs. The FFMSS profiles for each PD evinced substantial overlap with expert conceptualizations of PD prototypes from a FFM perspective (median r = .85). The second-order correlations for Cluster A disorders, however, were markedly lower than for Cluster B and C disorders, which may be due to the decreased variability of these PDs in the current sample relative to the other PDs. Alternatively, some have argued that the FFM does not contain content that references oddity and other traits and symptoms that are germane to the Cluster A PDs and that such a dimension (i.e., oddity; peculiarity) is both distinct from FFM Openness and necessary for capturing disorders such as Schizotypal PD (Watson, Clark, & Chmielewski, 2008).

The relation between the FFMSS facets and Dependent PD and the expert FFM ratings was also small compared with the other Cluster C PDs. Miller and Lynam (2001) have previously argued that there is somewhat of a disconnect between the expert ratings of Dependent PD (Lynam & Widiger, 2001) and the empirical FFM–Dependent PD correlates. As a result, Miller and Lynam generated a revised FFM–Dependent PD prototype on the basis of a meta-analysis; the revised FFM Dependent PD profile, which deemphasizes the role of high Agreeableness and emphasizes the role of low Conscientiousness, manifested a stronger correlation with the current FFMSS–Dependent PD profile (r = .62) compared with the original prototype (r = .52).

The FFMSS also evinced a substantial correlation with Obsessive–Compulsive PD, which is atypical in the FFM-PD literature (e.g., Miller et al., 2004). This strong correlation is likely due to the inclusion of maladaptive extremes at the high end of Conscientiousness, which are thought to be central to this PD. Haigler and Widiger (2001) demonstrated that the failure of certain FFM measures (i.e., NEO PI-R) to capture certain PDs, such as Schizotypal, Dependent, and OCPD, is due to a lack of items that reference maladaptivity at the high ends of Openness, Dependent, and OCPD, respectively. These authors found more substantial correlations between these three PDs and three FFM domains after experimentally altering NEO PI-R items to describe problematically high levels of these traits. The FFMSS, which includes maladaptivity at both the low and high ends of each FFM domain, did a much better job at capturing OCPD because of the changes to Conscientiousness. The same pattern was not found, however, for Schizotypal and Dependent PD, which still manifested limited correlations with Openness and Agreeableness, despite the inclusion of maladaptive levels of these two FFM domains.

Relations Between FFMSS Traits and Impairment Variables

As expected, Agreeableness and Conscientiousness (negatively) and Neuroticism (positively) were moderately to strongly related to impairment across a number of functional domains, which is consistent with previous research (Hopwood et al., 2009; Miller et al., 2005). The ability of the FFMSS to account for significant variance in a variety of domains of functioning (i.e., mean R^2 = .46) provides important evidence of the utility of these ratings. It is significant that the personality domains most strongly related to impairment were those that are conceptually related to the general PD criteria included in the *DSM-IV* (i.e., negative affect = Neuroticism; interpersonal functioning: Agreeableness, Extraversion; impulse control = Conscientiousness). The FFMSS also demonstrated incremental validity over Axis I-related symptoms (i.e., anxiety, depression) and Axis II PD symptoms in predicting four domains of functioning. Specifically, the FFMSS domains accounted for an additional 10% to 15% of the variance in these four variables with Conscientiousness emerging as the most consistent predictor. These results provide support for using the FFMSS in clinical settings as it provides meaningful and novel information that is pertinent

to client functioning and that may be inadequately assessed by current *DSM-IV* diagnostic constructs.

Strengths, Limitations, and Conclusions

Two strengths of the current study deserve mention. The use of the LEAD (Spitzer, 1983) method in generating expert consensus ratings for PD and impairment variables is a valuable approach as it emphasizes the contribution of expert clinical judgment and includes the use of multiple data sources. An additional strength is the use of a clinical sample in exploring the reliability and validity of the FFMSS as similar studies have used college samples or case vignettes (Mullins-Sweatt et al., 2006; Samuel & Widiger, 2006). The use of a clinical sample allowed for a test of whether clinicians can reliably and validly rate clients using this assessment in a sample that is similar to those in which it might be used in the future.

There were limitations to consider in the current study as well. Although training time on the FFMSS was minimal, the amount of patient exposure time (or exposure to patient information via case conference) that is needed to provide reliable and valid ratings of patients using the FFMSS is unclear. If it takes several hours of assessment time to rate clients on the FFM facets reliably then the FFMSS may not exhibit an advantage in terms of efficiency beyond the NEO PI-R. Significant time, however, is often given to broad assessment in clinical settings; as such, the current rating form may be feasible as part of an overall assessment summary. Therefore, it may not require additional clinician investment to rate clients effectively using the FFMSS. Another potential limitation relates to the characteristics of the current sample. Because the recruiting procedures were initially aimed at patients with Borderline and Avoidant PDs and patients with the absence of a PD diagnosis, the sample was overly weighted toward these forms of personality pathology. Despite this, the FFMSS correlated with the *DSM-IV* PDs in a manner that is consistent with previous studies (e.g., Samuel & Widiger, 2008). Finally, outside of the SNAP, the constructs explored here were all scored by the same group of expert raters, which raises the possibility that some of the current effect sizes are inflated due to shared method variance.

In conclusion, the current results provide further support for the reliability and validity of the use of single-item assessments of the 30 facets and five domains pertinent to the FFM. In the current study, we demonstrated that the FFMSS can be used to rate patients with reasonable reliability following only limited training and that the FFMSS provides personality data that is relatively convergent with other self-report personality instruments, expert ratings of *DSM-IV* PDs from an FFM perspective, and clinician ratings of *DSM-IV* PDs. The ratings also demonstrated clinical utility in "predicting" concurrent impairment across a number of domains (e.g., romance, work). Future studies should examine further the reliability, validity, and utility of this measure and compare it to a similar rating form (FFMRF) to see which measure is superior for use in clinical and research settings.

Acknowledgments

Appendix

Appendix

Five-Factor Model Score Sheet (Widiger & Spitzer, 2002)

7 = Problematic very high on the trait

6 = Problematic high on the trait (clear presence of clinically significant impairments)

5 = High on the trait (higher than the average, typical person; may or may not have minor impairments)

4 = Neither high nor low on the trait

3 = Low on the trait (lower than the average, typical person; may or may not have minor impairments)

2 = Problematic low on the trait (clear presence of clinically significant impairments)

1 = Problematic very low on the trait

? = Unable to estimate

Circle number that applies or ?:

	Neuroticism								
?	**Anxiousness** (fearful, apprehensive)	7	6	5	4	3	2	1	(relaxed, unconcerned, cool)
?	**Angry hostility**: (angry, bitter)	7	6	5	4	3	2	1	(even-tempered)
?	**Depressiveness**: (pessimistic, glum)	7	6	5	4	3	2	1	(optimistic)
?	**Self-consciousness**: (timid, embarrassed)	7	6	5	4	3	2	1	(self-assured, glib, shameless)
?	**Impulsivity**: (tempted, urgency)	7	6	5	4	3	2	1	(controlled, restrained)
?	**Vulnerability**: (helpless, difficulty dealing with stress)	7	6	5	4	3	2	1	(stalwart, brave, fearless, unflappable)
	Extraversion Versus Introversion								
?	**Warmth**: (affectionate, attached)	7	6	5	4	3	2	1	(Cold, aloof, indifferent)
?	**Gregariousness**: (sociable, outgoing)	7	6	5	4	3	2	1	(Withdrawn, isolated)
?	**Assertiveness**: (dominant, forceful)	7	6	5	4	3	2	1	(Unassuming, quiet, resigned)
?	**Activity**: (vigorous, energetic, active)	7	6	5	4	3	2	1	(Passive, lethargic)
?	**Excitement-seeking**: (reckless, daring)	7	6	5	4	3	2	1	(Cautious, monotonous, dull)
?	**Positive emotions**: (high-spirited)	7	6	5	4	3	2	1	(Placid, anhedonic)
	Openness Versus Closedness to Experience								
?	**Fantasy**: (dreamer, unrealistic, imaginative)	7	6	5	4	3	2	1	(practical, concrete)
?	**Aesthetic**: (preoccupied, aberrant, aesthetic)	7	6	5	4	3	2	1	(unaesthetic, uninvolved)
?	**Feelings**: (sensitive, responsive)	7	6	5	4	3	2	1	(constricted, constricted, alexythymic)
?	**Actions**: (unpredictable, unconventional)	7	6	5	4	3	2	1	(routine, habitual, stubborn)
?	**Ideas**: (strange, odd, peculiar, creative)	7	6	5	4	3	2	1	(pragmatic, rigid)
?	**Values**: (permissive, broad-minded)	7	6	5	4	3	2	1	(traditional, inflexible, dogmatic)
	Agreeableness Versus Antagonism								
?	**Trust**: (gullible, trusting)	7	6	5	4	3	2	1	(skeptical, cynical, suspicious, paranoid)
?	**Straightforwardness**: (naive, honest)	7	6	5	4	3	2	1	(cunning, manipulative, deceptive)
?	**Altruism**: (sacrificial, giving)	7	6	5	4	3	2	1	(stingy, selfish, greedy, exploitative)
?	**Compliance**: (docile, cooperative)	7	6	5	4	3	2	1	(oppositional, combative, aggressive)
?	**Modesty**: (meek, self-effacing, humble)	7	6	5	4	3	2	1	(confident, boastful, arrogant)
?	**Tender-mindedness**: (soft, empathic)	7	6	5	4	3	2	1	(tough, callous, ruthless)
	Conscientiousness								

?	**Competence**: (perfectionistic, efficient)	7	6	5	4	3	2	1	(lax, negligent)
?	**Order**: (ordered, methodical, organized)	7	6	5	4	3	2	1	(haphazard, disorganized, sloppy)
?	**Dutifulness**: (rigid, reliable, dependable)	7	6	5	4	3	2	1	(casual, undependable, unethical)
?	**Achievement**: (workaholic, ambitious)	7	6	5	4	3	2	1	(aimless, desultory)
?	**Self-Discipline**: (dogged, devoted)	7	6	5	4	3	2	1	(hedonistic, negligent)
?	**Deliberation**: (ruminative, reflective)	7	6	5	4	3	2	1	(hasty, careless, rash)

Provide average of facet scores within each of the five domains to obtain global five-factor model description:

Description for Five Broad Domains of Personality

Neuroticism	7	6	5	4	3	2	1	Low neuroticism
Extraversion	7	6	5	4	3	2	1	Introversion
Openness	7	6	5	4	3	2	1	Closedness
Agreeableness	7	6	5	4	3	2	1	Antagonism
Conscientiousness	7	6	5	4	3	2	1	Low conscientiousness

References

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

Bagby RM, Rector NA, Bindseil K, Dickens SE, Levitan RD, Kennedy SH. Self-report ratings and informants' ratings of personalities of depressed outpatients. American Journal of Psychiatry. 1998; 155:437–438. [PubMed: 9501762]

Carroll BJ, Fielding JM, Blashki TG. Depression rating scales: A critical review. Archives of General Psychiatry. 1973; 28:361–366. [PubMed: 4688625]

Clark, LA. Manual for the Schedule for Nonadaptive and Adaptive Personality (SNAP). University of Minnesota Press; Minneapolis: 1993.

Clark LA. Assessment and diagnosis of personality disorder: Perennial issues and an emerging reconceptualization. Annual Review of Psychology. 2007; 58:227–257.

Clark, LA.; Vorhies, L.; McEwen, JL. Personality disorder symptomatology from the five-factor model perspective.. In: Costa, PT.; Widiger, TA., editors. Personality disorders and the five-factor model of personality. 2nd ed.. American Psychological Association; Washington, DC: 2002. p. 125-148.

Cohen, J.; Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. 2nd ed.. Erlbaum; Hillsdale, NJ: 1983.

Costa, PT.; McCrae, RR. Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) professional manual. Psychological Assessment Resources; Odessa, FL: 1992.

Costa, PT.; Widiger, TA., editors. Personality disorders and the five-factor model of personality. 2nd ed.. American Psychological Association; Washington, DC: 2002.

Digman JM. Personality structure: Emergence of the five-factor model. Annual Review of Psychology. 1990; 41:417–470.

First, MB.; Gibbon, M.; Spitzer, RL.; Williams, JBW. Structured clinical interview for DSM-IV Axis-II personality disorders. American Psychiatric Press; Washington DC: 1997.

Gosling SD, Rentfrow PJ, Swann WB Jr. A very brief measure of the Big-Five personality domains. Journal of Research in Personality. 2003; 37:504–528.

Haigler ED, Widiger TA. Experimental manipulation of NEO PI-R items. Journal of Personality Assessment. 2001; 77:339–358. [PubMed: 11693863]

Hamilton M. The assessment of anxiety states by rating. British Journal of Medical Psychology. 1959; 32:50–55. [PubMed: 13638508]

Hamilton M. A rating scale for depression. Journal of Neurology, Neurosurgery, and Psychiatry. 1960; 23:56–62.

Hopwood CJ, Morey LC, Ansell EB, Grilo CM, Sanislow CA, McGlashan TH, Skodol AE. The convergent and discriminant validity of the five-factor traits: Current and prospective social, work, and recreational dysfunction. Journal of Personality Disorders. 2009; 23:466–476. [PubMed: 19817628]

John, OP.; Srivastava, S. The big five trait taxonomy: History, measurement, and theoretical perspectives.. In: Pervin, LA.; John, OP., editors. Handbook of personality: Theory and research. 2nd ed.. Guilford Press; New York, NY: 1999. p. 102-138.

Klein D. Patients' versus informants' reports of personality disorders in predicting 7½-year outcome in outpatients with depressive disorders. Psychological Assessment. 2003; 15:216–222. [PubMed: 12847782]

Lowe JR, Widiger TA. Clinicians' judgments of clinical utility: A comparison of the *DSM-IV* with dimensional models of general personality. Journal of Personality Disorders. 2009; 23:211–229. [PubMed: 19538078]

Lynam DR, Widiger TA. Using the five-factor model to represent the *DSM-IV* personality disorders: An expert consensus approach. Journal of Abnormal Psychology. 2001; 110:401–412. [PubMed: 11502083]

Maier W, Buller R, Philipp M, Heuser I. The Hamilton Anxiety Scale: Reliability, validity, and sensitivity to change in anxiety and depressive disorders. Journal of Affective Disorders. 1988; 14:61–68. [PubMed: 2963053]

McCrae RR. A note on some measures of profile agreement. Journal of Personality Assessment. 2008; 90:105–109. [PubMed: 18444102]

Miller JD, Lynam DR. Structural models of personality and their relation to antisocial behavior: A meta-analytic review. Criminology. 2001; 39:765–798.

Miller JD, Pilkonis PA, Clifton A. Self- and other-reports of traits from the five-factor model: Relations to personality disorder. Journal of Personality Disorders. 2005; 19:400–419. [PubMed: 16178682]

Miller JD, Pilkonis PA, Mulvey EP. Treatment utilization and satisfaction: Examining the contributions of Axis II psychopathology and the five-factor model of personality. Journal of Personality Disorders. 2006; 4:369–387. [PubMed: 16901260]

Miller JD, Reynolds SK, Pilkonis PA. The validity of the five-factor model prototypes for personality disorders in two clinical samples. Psychological Assessment. 2004; 16:310–322. [PubMed: 15456386]

Morse JQ, Hill J, Pilkonis PA, Yaggi KE, Broyden N, Stepp SD, Feske U. Anger, preoccupied attachment, and domain disorganization in borderline personality disorder. Journal of Personality Disorders. 2009; 23:240–257. [PubMed: 19538080]

Mullins-Sweatt SN, Jamerson JE, Samuel DB, Olson DR, Widiger TA. Psychometric properties of an abbreviated instrument of the five-factor model. Assessment. 2006; 13:119–137. [PubMed: 16672728]

Oltmanns TF, Friedman JNW, Fiedler ER, Turkheimer E. Perceptions of personality disorders based on thin slices of behavior. Journal of Research in Personality. 2004; 38:216–229.

Reynolds SK, Clark LA. Predicting dimensions of personality disorder from domains and facets of the five-factor model. Journal of Personality. 2001; 69:199–222. [PubMed: 11339796]

Samuel DB, Widiger TA. Clinicians' personality descriptions of prototypic personality disorders. Journal of Personality Disorders. 2004; 18:286–308. [PubMed: 15237048]

Samuel DB, Widiger TA. Clinicians' judgments of clinical utility: A comparison of the *DSM-IV* and five-factor models. Journal of Abnormal Psychology. 2006; 115:298–308. [PubMed: 16737394]

Samuel DB, Widiger TA. A meta-analytic review of the relationships between the five-factor model and *DSM-IV-TR* personality disorders: A facet level analysis. Clinical Psychology Review. 2008; 28:1326–1342. [PubMed: 18708274]

Saucier G. Mini-markers: A brief version of Goldberg's unipolar big five markers. Journal of Personality Assessment. 1994; 63:506–516. [PubMed: 7844738]

Spitzer RL. Psychiatric diagnosis: Are clinicians still necessary? Comprehensive Psychiatry. 1983; 24:399–411. [PubMed: 6354575]

Sprock J. A comparative study of the dimensions and facets of the five-factor model in the diagnosis of cases of personality disorder. Journal of Personality Disorders. 2002; 16:402–423. [PubMed: 12489308]

Trull, TJ.; Widiger, TA. Structured interview for the five-factor model of personality (SIFFM): Professional manual. Psychological Assessment Resources; Odessa, FL: 1997.

Trull TJ, Widiger TA, Useda JD, Holcomb H, Doan B, Axelrod SR, Gershuny BS. A structured interview for the assessment of the five-factor model of personality. Psychological Assessment. 1998; 10:229–240.

Watson D, Clark LA, Chmielewski M. Structures of personality and their relevance to psychopathology. II. Further articulation of a comprehensive unified trait structure. Journal of Personality. 2008; 76:1545–1586. [PubMed: 19012658]

Whiteside SP, Lynam DR. The five factor model and impulsivity: Using a structural model of personality to understand impulsivity. Personality and Individual Differences. 2001; 30:669–689.

Widiger, TA. Five-Factor Model Rating Form. 2004. Unpublished measure

Widiger, TA.; Spitzer, RL. Five-Factor Model Score Sheet. 2002. Unpublished measure

Table 1

Descriptive Statistics and Interrater Reliability Coefficients for FFMSS Traits

	M	SD	ICC $_{DE}$	Conv r	Median Div r
Neuroticism (α = .61)	30.85	4.22	.55		
Anxiety	5.52	1.23	.48	.43*	.12
Angry hostility	4.88	1.27	.54	.30*	−.20
Depression	5.35	1.12	.30	.45*	−.13
Self-consciousness	5.06	1.46	.59	.38*	.01
Impulsiveness	4.69	1.38	.54	.06	.03
Vulnerability	5.35	1.10	.56	.54*	−.07
Extraversion (α = .88)	22.82	6.83	.68		
Warmth	3.89	1.44	.50	.54*	.18
Gregariousness	3.71	1.62	.62	.83*	.04
Assertiveness	4.08	1.40	.56	.60*	−.11
Activity	3.71	1.39	.62	.78*	.09
Excitement seeking	4.00	1.41	.62	.64*	−.14
Positive emotions	3.42	1.34	.61	.78*	.08
Openness (α = .87)	24.89	6.35	.58		
Fantasy	4.12	1.49	.51	.64*	.00
Aesthetics	4.24	1.25	.38	.73*	.06
Feelings	4.65	1.41	.41	.59*	.11
Actions	3.94	1.49	.47	.67*	−.06
Ideas	3.89	1.31	.45	.72*	.01
Values	4.06	1.24	.50	.64*	−.01
Agreeableness (α = .86)	23.85	5.98	.66		
Trust	3.60	1.52	.53	.55*	.12

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

	M	SD	ICC_{DE}	Conv r	Median Div r
Straightforwardness	4.25	1.22	.49	.74*	.06
Altruism	4.07	1.27	.47	.71*	.10
Compliance	3.78	1.34	.59	.71*	.04
Modesty	4.09	1.30	.60	.55*	.04
Tender-mindedness	4.06	1.08	.49	.71*	.16
Conscientiousness (α = .92)	23.07	6.52	.72		
Competence	3.90	1.34	.57	.79*	.01
Order	3.91	1.26	.45	.75*	−.01
Dutifulness	4.02	1.34	.52	.82*	.02
Achievement-striving	3.58	1.20	.37	.75*	.08
Self-discipline	3.64	1.26	.54	.79*	.01
Deliberation	4.02	1.34	.51	.76*	−.09

Note: FFMSS = Five-Factor Model Score Sheet; ICC_{DE} = double entry intraclass correlation coefficient; N = 112 for ICC_{DE}; Conv r = corrected item–total correlations between a specific facet and the "home" domain score without that facet (e.g., Compliance Conv r = correlation between Compliance and Agreeableness *without* the Compliance facet included in the total domain score); Median Div r = median correlation of each facet with facets from the other four FFM (five-factor model) domains. N = 130 for descriptive statistics;

*
p .001.

Table 2

Interrelations Between the Five-Factor Model Score Sheet Domains

	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
Neuroticism	—				
Extraversion	−.36[*]	—			
Openness	.11	.49[*]	—		
Agreeableness	−.07	.04	.02	—	
Conscientiousness	−.26	−.05	−.18	.41[*]	—

[*]
p .001.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Correlations Between Clinician-Rated FFMSS Traits and SNAP Scales

	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
Negative temperament	.45*	–.11	.10	–.06	–.30*
Mistrust	.23	.04	.02	–.22	–.28
Manipulativeness	.02	.03	.20	–.27	–.34*
Aggression	.13	.16	.13	–.38*	–.33*
Self-harm	.51*	–.06	.21	–.11	–.31*
Eccentric perceptions	.10	.19	.32*	–.16	–.31*
Dependency	.27*	–.05	.05	.08	–.15
Positive temperament	–.47*	.52*	.20	.05	.09
Exhibitionism	–.29	.42*	.25	–.03	.03
Entitlement	–.23	.18	.08	–.25	–.16
Detachment	.36*	–.47*	–.07	–.13	–.09
Disinhibition	.11	.12	.17	–.24	–.32*
Impulsivity	.10	.14	.30*	–.15	–.28
Propriety	.05	.08	–.07	.06	.05
Workaholism	.01	.07	.03	.06	.30*

Note: FFMSS = Five-Factor Model Score Sheet; SNAP = Schedule for Nonadaptive and Adaptive Personality. *N* = 125.

*
p .001.

Table 4

Correlations Between Clinician-Rated FFMSS Traits and Consensus *DSM-IV* Personality Disorder Ratings

	PPD	SPD	STPD	ASPD	BPD	HPD	NPD	AVPD	DPD	OCPD
Neuroticism	.11	.13	.14	.05	.42	.09	–.03	.35	.37	.02
Anxious	–.20	.22	.20	–.22	–.03	–.15	–.13	.40	.15	.20
Angry hostility	.39	.03	.15	.27	.44	.18	.29	–.06	.11	.09
Depressiveness	.07	.21	.00	–.15	.07	–.06	.02	.44	.15	.14
Self-consciousness	–.16	.18	–.03	–.20	–.04	–.17	–.26	.63	.20	.07
Impulsivity	.20	–.21	.08	.45	.60	.29	.08	–.31	.24	–.30
Vulnerability	.11	.07	.11	–.01	.46	.27	–.07	.15	.46	–.10
Extraversion	–.12	–.31	–.12	.17	.18	.36	.07	–.66	–.01	–.15
Warmth	–.32	–.32	–.23	–.09	.01	.19	–.24	–.39	.15	–.18
Gregariousness	–.13	–.27	–.17	.07	.05	.37	.03	–.62	.04	–.16
Assertiveness	.22	–.16	.11	.26	.28	.29	.29	–.64	–.16	–.02
Activity	–.20	–.19	–.01	.04	.04	.21	.13	–.50	–.07	–.01
Excitement seeking	.02	–.28	–.15	.42	.45	.37	.07	–.53	.03	–.27
Positive emotions	–.16	–.26	–.08	.13	.02	.28	.06	–.47	–.03	–.08
Openness	–.07	–.20	.21	.20	.36	.31	.21	–.36	–.02	–.24
Fantasy	.04	–.11	.24	.11	.20	.28	.43	–.31	–.09	–.09
Aesthetic	–.13	–.11	.14	.04	.21	.15	.25	–.20	–.12	–.03
Feelings	–.12	–.19	.03	.11	.26	.31	–.08	–.30	.16	–.24
Actions	.04	–.25	.17	.31	.41	.30	.11	–.44	.07	–.30
Ideas	–.05	–.07	.38	.12	.30	.21	.21	–.11	–.03	–.18
Values	–.14	–.16	.02	.24	.30	.17	.05	–.27	–.09	–.28
Agreeableness	–.49	–.05	–.22	–.41	–.33	–.15	–.49	.22	.12	–.08
Trust	–.44	–.05	–.17	–.30	–.20	.05	–.17	–.10	.23	–.05
Straightforwardness	–.35	–.03	–.16	–.43	–.39	–.12	–.35	.20	–.03	–.02
Altruism	–.34	.01	–.28	–.31	–.32	–.18	–.48	.14	–.04	–.02
Compliance	–.39	–.02	–.18	–.39	–.34	–.17	–.33	.23	.03	–.07
Modesty	–.29	–.05	–.12	–.20	–.14	–.26	–.64	.45	.18	–.12
Tender-mindedness	–.46	–.10	–.12	–.30	–.17	–.05	–.36	.16	.14	–.09
Conscientiousness	–.27	.10	.02	–.39	–.45	–.29	–.07	.18	–.21	.39

	PPD	SPD	STPD	ASPD	BPD	HPD	NPD	AVPD	DPD	OCPD
Competence	−.09	.03	.13	−.28	−.30	−.14	.05	.12	−.23	.42
Order	−.12	.13	.09	−.23	−.25	−.15	.03	.20	−.13	.32
Dutifulness	−.28	.03	−.01	−.38	−.41	−.27	−.15	.13	−.21	.28
Achievement	−.29	−.05	.02	−.32	−.39	−.19	.00	−.03	−.25	.36
Self-discipline	−.34	.15	−.08	−.37	−.46	−.30	−.09	.11	−.13	.28
Deliberation	−.26	.21	−.05	−.40	−.48	−.41	−.21	.36	−.13	.33
r With L&W	.47	.77*	.53	.90*	.84*	.93*	.86*	.87*	.52	.95*

Note: FFMSS = Five-Factor Model Score Sheet; *DSM-IV = Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition; PPD = paranoid personality disorder; SPD = schizoid personality disorder, STPD = schizotypal personality disorder; ASPD = antisocial personality disorder; BPD = borderline personality disorder; HPD = histrionic personality disorder; NPD = narcissistic personality disorder; AVPD = avoidant personality disorder; DPD = dependent personality disorder; OCPD = obsessive–compulsive personality disorder; *r* with L&W = *r* with Lynam and Widiger (2001) expert FFM (five-factor model) PD prototype ratings. Correlations |.28| are significant at the .001 level. *N* = 130.

Table 5

Correlations Between Clinician-Rated Five-Factor Model Score Sheet Traits and Impairment

	Overall	Romantic	Parental	Occupational	Social	Distress to Others
Neuroticism	.57[*]	.39[*]	.26	.45[*]	.54[*]	.52[*]
Extraversion	−.28[*]	−.14	.10	−.21	−.44[*]	.02
Openness	.06	.12	.13	.09	−.09	.16
Agreeableness	−.45[*]	−.26	−.45[*]	−.41[*]	−.37[*]	−.52[*]
Conscientiousness	−.51[*]	−.31[*]	−.49[*]	−.62[*]	−.39[*]	−.56[*]
R^2 domains	.56[*]	.23[*]	.37[*]	.52[*]	.50[*]	.60[*]

Note: N = 59 for parental impairment; N = 130 for all other impairment variables; Overall = overall impairment; Romantic = romantic impairment; Parental = parental impairment; Occup = occupational impairment; Social = social impairment.

[*]
p .001.