An Investigation of the Factor Structure and Convergent and Discriminant Validity of the Five-Factor Model Rating Form

Douglas B. Samuel, Stephanie N. Mullins-Sweatt and Thomas A. Widiger

Assessment published online 7 August 2012
DOI: 10.1177/1073191112455455

Assessment
XX(X) 1–12

DOI: 10.1177/1073191112455455
http://asm.sagepub.com
SAGE

An Investigation of the Factor Structure and Convergent and Discriminant Validity of the Five-Factor Model Rating Form

Douglas B. Samuel[1], Stephanie N. Mullins-Sweatt[2], and Thomas A. Widiger[3]

Abstract

The Five-Factor Model Rating Form (FFMRF) is a one-page measure designed to provide an efficient assessment of the higher order domains of the Five Factor Model (FFM) as well as the more specific, lower order facets proposed by McCrae and Costa. Although previous research has suggested that the FFMRF's assessment of the lower order facets converge reasonably with other FFM measures, the structural validity of the domain-level assessment has not yet been evaluated. The current study employed an exploratory structural equation modeling framework to investigate the fit of a five-factor solution within a combined sample of 757 participants. This was a novel analysis using a combined sample drawn from three previously published studies and was composed primarily of undergraduates but also included a smaller clinical subsample. Results indicated that the FFMRF is well accommodated within a five-factor solution. Furthermore, the FFMRF domain scores evinced large correlations with domain scores from the NEO Personality Inventory–Revised. The results suggest that the FFMRF might hold promise as a choice for those seeking a brief measure that provides a valid assessment of both the broad and specific traits of the FFM.

Keywords

five-factor model, short form, abbreviated, personality, exploratory structural equation modeling

The Five-Factor Model of personality (FFM; McCrae & Costa, 2008) stems from the lexical tradition and comprises five bipolar domains that have been labeled surgency or extraversion (vs. introversion), agreeableness (vs. antagonism), conscientiousness (vs. disinhibition), neuroticism (vs. emotional stability), and intellect or openness (vs. closedness to experience). Although alternative models of general personality with between three and seven domains also have been studied extensively (e.g., the HEXACO model of Ashton and Lee [2007]; Tellegen's [1985] three temperaments, and Cloninger's [2008] psychobiological theory), the FFM has proven useful for integrating diverse personality models within a commonly understood framework (e.g., Caspi, Roberts, & Shiner, 2005; Deary, Weiss, & Batty, 2011; Feingold, 1994; Ozer & Benet-Martinez, 2006). Indeed, each of the models can be included within a common hierarchical structure, with the five-factor model derivable from the three-factor model of Tellegen (Markon, Krueger, & Watson, 2005). However, Markon et al. (2005) do suggest that the FFM represents the "crucial level of analysis for normal personality research" (p. 154). Because of this success some have argued that "the field has now achieved an initial consensus on a general taxonomy of personality traits" consisting of five personality dimensions (John, Naumann, & Soto, 2008, p. 116).

The FFM does have an extensive research literature supporting its validity including evidence regarding heritability (Krueger & Johnson, 2008; Yamagata et al., 2006), developmental antecedents (Caspi et al., 2005; Widiger, De Clercq, & De Fruyt, 2009), universality across cultures (Allik, 2005; McCrae et al., 2005), temporal stability (Roberts & DelVecchio, 2000), and preliminary research suggesting ties with brain structure (DeYoung et al., 2010). In addition, the FFM has evinced meaningful relationships with important life outcomes (Hopwood et al., 2009; Mullins-Sweatt & Widiger, 2010; Ozer & Benet-Martinez, 2006) and psychiatric diagnoses (e.g., Samuel & Widiger, 2008a).

[1]Purdue University, West Lafayette, IN, USA
[2]Oklahoma State University, Stillwater, OK, USA
[3]University of Kentucky, Lexington, KY, USA

Corresponding Author:
Douglas B. Samuel, Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907, USA
Email: dbsamuel@purdue.edu

Although the five higher order domains are the primary units of the FFM, researchers have further subdivided the domains into lower order facets that provide a fine-grained representation of more specific traits (e.g., Costa & McCrae, 1995; DeYoung, Quilty, & Peterson, 2007; Goldberg, 1999). Facet-level assessments have demonstrated incremental validity beyond the FFM domains for predicting specific behaviors including grade point average attainment and dating frequency (Paunonen & Ashton, 2001; Paunonen, Haddock, Fosterling, & Keinonen, 2003) and are useful for differentiating among personality disorder constructs (e.g., Axelrod, Widiger, Trull, & Corbitt, 1997; Reynolds & Clark, 2001). For these reasons, the lower order facets, in addition to the higher order domains, are particularly useful in many contexts.

Because of the utility offered by the facets, there are a number of measures that provide an explicit assessment of lower order constructs (e.g., Ashton & Lee, 2007; Goldberg, 1999). Perhaps the most notable facet level measure is the NEO Personality Inventory–Revised (NEO PI-R; Costa & McCrae, 1992). Based on their work with this instrument, Costa and McCrae (1995) divided each domain into six underlying facets on the basis of factor analytic and theoretical considerations. For instance, they indicated that extraversion could be differentiated into the facets of warmth, gregariousness, assertiveness, activity, excitement seeking, and positive emotions. In choosing the facets, Costa and McCrae (1995) aimed for traits that were comparable in scope, consistent with the existing personality literature, and that represented "the more closely covarying elements within the domain, not arbitrary combinations of elements" (p. 25) to maximize their discriminant validity within the domain.

There are certainly other ways to divide the five domains into their component parts (e.g., DeYoung et al., 2007; Lee & Ashton, 2004; Saucier & Goldberg, 2002) and the specific facets delineated by Costa and McCrae (1995) have received some criticism. For example, because the FFM domains themselves are not entirely orthogonal, some facets relate to more than a single domain (e.g., impulsiveness within neuroticism typically correlates with conscientiousness and angry hostility often relates with antagonism; Costa & McCrae, 1992). In addition, some have lamented that their development occurred outside of the lexical tradition.

An obstacle to the widespread use of the NEO PI-R or most of these facet-level measures is that they include 200 or more items and require at least 20 to 30 minutes to complete. Such a detailed assessment is advantageous in many respects, but there are situations where this time burden is prohibitive. For example, more abbreviated measures are typically preferred for purposes such as (a) prescreening, (b) when an individual assesses a number of different targets, or (c) for large-scale surveys that cover a variety of topics. The NEO PI-R does have an accompanying abbreviated version, the NEO Five-Factor Inventory (NEO-FFI). With only 60 items the NEO-FFI is sometimes used by those seeking a more efficient assessment of the FFM. However, its utility is curtailed by the fact that it does not include a facet-level assessment.

In addition to these issues, abbreviated measures are often the only realistic means by which to collect descriptions from busy professionals, such as therapists or physicians within their clinical practice. This latter point was precisely the motivation behind the development by Lynam and Widiger (2001) of a brief rating form that allowed them to collect facet-level FFM descriptions of personality disorders from expert researchers and clinicians. In doing so, they included the identifying label for each of the NEO PI-R facet scales, as well as a few adjective descriptors to describe each of the poles. These adjectives were drawn from the NEO PI-R manual as well as other FFM adjective checklists (e.g., Goldberg, 1992). This rating form, which serves as a brief measure of the NEO PI-R model, was subsequently termed the *Five-Factor Model Rating Form* (FFMRF) and has been used in a number of studies to collect FFM descriptions from psychologists and other mental health professionals (e.g., Lowe & Widiger, 2009; Mullins-Sweatt & Widiger, 2011; Samuel & Widiger, 2004, 2006, 2009, 2010a). In these studies, the FFMRF has produced descriptions that have been reliable across raters and evinced temporal consistency comparable to other FFM measures (i.e., median 6-month stability for the domains was .54; Samuel & Widiger, 2011a).

Recognizing its potential utility as an abbreviated measure of the FFM facets, Mullins-Sweatt, Jamerson, Samuel, Olson, and Widiger (2006) used the FFMRF to collect self-report ratings. Through a series of studies they made minor revisions to the adjective descriptors (e.g., *sensitive* and *responsive* were replaced by *self-aware* for high openness to feelings) and investigated the psychometric properties of the FFMRF. They found that self-report ratings on the FFMRF displayed reasonable internal consistency with a median alpha value for the domains of .69 across five samples. The FFMRF domain scores also displayed median convergent validity values with the NEO PI-R that ranged from .57 (openness) to .68 (extraversion). At the facet level, the average convergence of the FFMRF items with the respective NEO PI-R facets within each domain ranged from .37 (openness) to .50 (extraversion) and were larger than the mean discriminant validity correlations with facets within (Mdn = .24) and outside (Mdn = .03) the same domain. Finally, the FFMRF scores manifested predictable relationships with personality pathology, supporting the construct validity of the measure. The FFMRF has since been used in other studies as a self-report measure (e.g., Wirth, Lynam, & Williams, 2010).

Despite this evidence supporting the validity of the FFMRF, there are limitations to the existing literature that should be addressed. Further examination is particularly warranted in the case of the FFMRF as abbreviated measures must demonstrate validity comparable to their parent measures rather than simply time savings (Smith, McCarthy, & Anderson, 2000). More specifically, Smith et al. (2000) identified a series of potential concerns that should be explored in the validation process of short form measures. For example, they indicated that abbreviated measures should demonstrate empirically that they reproduce the factor structure of lengthier inventories. This is particularly relevant for the FFMRF as the higher order factor structure is the defining feature of the FFM. However, there has not yet been an examination of the FFMRF to determine how well its facet ratings cohere within the FFM framework. We address this by examining the fit of a five-factor solution to FFMRF data drawn from 757 participants combined across three previously published data sets. We then investigated the convergent and discriminant validity of the FFMRF scores with those from the NEO PI-R.

Method

Samples and Participants

The data used in the present study were compiled from three previously published studies in which the self-report FFMRF was administered. One hundred and thirty-eight undergraduate participants were drawn from Samuel and Widiger (2008b). An additional 536 undergraduate participants completed the FFMRF as part of a study concerned with measures of obsessive compulsive personality disorder (OCPD) in which a portion of this subsample included individuals prescreened for elevated levels of OCPD pathology (Samuel & Widiger, 2010b). Finally, 83 participants were drawn from a clinical sample (Samuel & Widiger, 2010a), yielding a total combined sample of 757 individuals. However, the current results represent a novel use of the data. The clinical sample has been used in existing publications, but those previous reports did not concern any data from the self-report FFMRF. The sample of 536 undergraduates completed the full FFMRF, but only the conscientiousness domain score was used in a previous publication (Samuel & Widiger, 2011b), such that all facet-level information is novel. Finally, although no FFMRF data were reported in Samuel and Widiger (2008b), a preliminary subset of these participants (n = 75) was reported as Study 5 within the original Mullins-Sweatt et al.'s (2006) article. Thus, to avoid potential overlap we use this particular subsample only for the factor analytic procedures and not the convergent and discriminant validity analyses, which are confined to the remaining 619 participants.

The demographics for the three subsamples have been reported previously in the articles cited above, but we note that demographics of the first two were comparable with typical university students (i.e., young, primarily White, and slightly more females than males). The clinical subsample was recruited from various treatment clinics within a moderately sized city in the Southeastern United States, but a majority of participants were drawn from a therapeutic community for women with substance use disorders. Demographically, the clinical subsample was older (M = 34.8 years), more diverse (24% African American), less educated, and was confined to women. Substance use diagnoses predominated the subsample, with cocaine abuse/dependence (59%) the most common. Other prominent diagnoses in this subsample included antisocial (36%), borderline (23%), and avoidant (21%) PDs, as well as major depressive (11%) and bipolar (10%) mood disorders. The pooled demographics for the combined sample was 68% female with a mean age of 20.9 years (SD = 6.4). Eighty-nine percent were White and 7% were African American.

Measures

Five-Factor Model Rating Form. The FFMRF is a one-page rating form consisting of an item representing each of the 30 facets of the FFM. The 30 items are organized with respect to the five domains. For example, under the heading Neuroticism are six items. Each item is rated on a 1 to 5 scale where 1 is *extremely low*, 2 is *low*, 3 is *neither high nor low*, 4 is *high*, and 5 is *extremely high*. For example, the neuroticism facet of anxiousness was assessed with the descriptors "fearful, apprehensive versus relaxed, unconcerned, cool" and the openness facet of ideas was assessed with the descriptors "strange, odd, peculiar, creative versus pragmatic, rigid." Cronbach's α values for the FFMRF domains within the combined sample ranged from .63 (openness to experience) to .80 (conscientiousness), with a median of .72.

NEO Personality Inventory–Revised. The NEO PI-R (Costa & McCrae, 1992) was also administered to the participants in each sample and contains 240 statements to which the individual responds *strongly disagree, disagree, neutral, agree*, or *strongly agree* (0-4 Likert-type scale). Eight items assess each of the 30 facets, which are in turn summed to score the five domains. For the combined sample Cronbach's α values ranged from .73 (openness) to .85 (conscientiousness), with a median of .81 across the five domains.

Factor Analytic Procedures

Because the FFMRF implies a clear five-factor structure we sought to investigate how well it would fit within a traditional confirmatory factor analytic (CFA) framework using M*plus 6.11* (Muthen & Muthen, 2011), employing the

maximum likelihood estimator (MLR). Nonetheless, we recognized that the strict assumptions of CFA might provide an unreasonable standard for personality measures (e.g., Hopwood & Donnellan, 2010). Indeed, even the most robust FFM inventories, such as the NEO PI-R, have failed to reproduce the expected five-factor framework (McCrae, Zonderman, Costa, Bond, & Paunonen, 1996; Vassend & Skrondal, 1997). Thus, we also used exploratory structural equation modeling (ESEM), which blends elements of CFA with those from exploratory factor analysis. Recent advances in this methodology have demonstrated that ESEM can adequately model data from FFM inventories, where CFA methods fail (Marsh et al., 2010; Rosellini & Brown, 2011). For the ESEM analyses, we used the MLR estimation procedure. Although the FFM domains are theorized to be orthogonal, we used the geomin rotation method recommended by Browne (2001) as this oblique method allows factors to emerge as orthogonal if this reflects the data (e.g., Fabrigar, Wegener, MacCallum, & Strahan, 1999). Echoing previous reports (i.e., Marsh et al., 2010), we used several fit indices. These included the comparative fit index (CFI) and Tucker–Lewis index (TLI), with values above .90 and .95 indicating acceptable and excellent fit, respectively (Hu & Bentler, 1999). We also used the root-mean-square error of approximation (RMSEA) with values lower than .080 and .050 indicating close and reasonable fit, respectively, and the standardized root mean square residual where values below .050 indicating good fit (Marsh, Hua, & Wen, 2004).

Results

Descriptive Statistics

Means and standard deviations for the FFMRF facets and domains are provided in Table 1. These values are largely comparable with those reported by Mullins-Sweatt et al. (2006). In most cases, the facet scores for extraversion, agreeableness, and conscientiousness were above the scale's midpoint (i.e., 3.0), whereas those for neuroticism tended to be below.

Table 1. Descriptive Statistics for FFMRF Items

FFMRF item	*M*	*SD*
n1 Anxiousness	2.87	1.03
n2 Angry Hostility	2.20	0.93
n3 Depressiveness	2.47	1.09
n4 Self-consciousness	2.94	1.04
n5 Impulsivity	2.81	0.99
n6 Vulnerability	2.47	0.96
e1 Warmth	4.00	0.87
e2 Gregariousness	3.69	0.99
e3 Assertiveness	3.11	0.95
e4 Activity	3.67	0.93
e5 Excitement Seeking	3.33	0.95
e6 Positive Emotions	3.80	0.84
o1 Fantasy	3.39	1.06
o2 Aesthetics	3.38	0.82
o3 Feelings	3.85	0.81
o4 Actions	3.15	0.95
o5 Ideas	3.51	0.88
o6 Values	3.39	1.03
a1 Trust	3.47	1.06
a2 Straightforwardness	3.79	0.87
a3 Altruism	3.73	0.82
a4 Compliance	3.66	0.83
a5 Modesty	3.40	0.95
a6 Tender-mindedness	3.72	0.91
c1 Competence	3.76	0.89
c2 Order	3.50	1.00
c3 Dutifulness	3.77	0.89
c4 Achievement	3.57	0.86
c5 Self-Discipline	3.59	0.87
c6 Deliberation	3.57	0.86

Note. FFMRF = Five-Factor Model Rating Form. $N = 757$.

Factor Analytic Procedures

We first employed a traditional CFA using the procedures described earlier. As with prior research using FFM instruments, a five-factor solution provided an inadequate fit, CFI = .753, TLI = .728, RMSEA = .067, and standardized root mean square residual (SRMR) = .080. We then proceeded with the ESEM analysis and this revealed a much better fit. The ESEM results indicated at least acceptable fit according to all the metrics, CFI = .943, TLI = .915, RMSEA = .033, SRMR = .027. To provide a clearer test of the differences between these two models, we examined the sample size adjusted Bayesian information criteria (BIC) for each model. The BIC for the CFA model was 55,427, and the value for the ESEM model was 54,668. This difference of 759 is quite a lot larger than the BIC difference of 10 that indicates very strong evidence (Raftery, 1995) for improved fit.

Table 2 provides the parameter estimates based on the ESEM solution. These results suggest that not only does a five-factor ESEM solution fit the data adequately, but that the specific loadings align closely with the FFMRF's a priori structure. Parameter estimates greater than |.30| are in boldface type, consistent with previous approaches to identify items with meaningful loadings on a factor. The standard errors for each of these loadings appear in parentheses. The first factor was defined primarily by the neuroticism facets, which obtained loadings ranging from .38 (impulsivity) to .69 (depressiveness). In contrast, the largest loading

Table 2. Exploratory Structural Equation Model of a Five-Factor Solution for the 30 Facets of the FFMRF

	Factor 1		Factor 2		Factor 3		Factor 4		Factor 5	
FFMRF facet	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Anxiousness (n1)	**.68**	(.04)	−.02	(.04)	.06	(.04)	.08	(.11)	−.06	(.05)
Angry Hostility (n2)	**.41**	(.06)	.03	(.05)	.08	(.05)	**−.39**	(.09)	.04	(.04)
Depressiveness (n3)	**.69**	(.05)	−.17	(.06)	.00	(.03)	−.08	(.10)	.08	(.05)
Self-Consciousness (n4)	**.64**	(.04)	−.07	(.05)	−.05	(.04)	.14	(.11)	−.06	(.05)
Impulsivity (n5)	**.38**	(.06)	**.33**	(.08)	**−.30**	(.08)	−.14	(.15)	.04	(.05)
Vulnerability (n6)	**.50**	(.05)	.02	(.04)	−.14	(.04)	.09	(.10)	−.07	(.05)
Warmth (e1)	.09	(.05)	**.41**	(.08)	.02	(.03)	**.34**	(.10)	−.02	(.04)
Gregariousness (e2)	−.04	(.04)	**.60**	(.08)	−.05	(.05)	.11	(.12)	−.07	(.06)
Assertiveness (e3)	.07	(.06)	**.39**	(.07)	.20	(.07)	**−.39**	(.09)	.01	(.03)
Activity (e4)	−.12	(.06)	**.48**	(.07)	.20	(.07)	−.01	(.06)	.02	(.05)
Excitement Seeking (e5)	−.05	(.05)	**.36**	(.08)	−.07	(.06)	−.19	(.08)	**.30**	(.07)
Positive Emotions (e6)	−.15	(.05)	**.35**	(.06)	.13	(.05)	.17	(.06)	.10	(.04)
Fantasy (o1)	.07	(.05)	.03	(.05)	−.09	(.05)	.05	(.05)	**.50**	(.06)
Aesthetics (o2)	−.02	(.04)	.01	(.05)	.13	(.05)	.02	(.04)	**.32**	(.06)
Feelings (o3)	.04	(.05)	.15	(.07)	.10	(.04)	.16	(.06)	.16	(.06)
Actions (o4)	.02	(.04)	.09	(.06)	−.09	(.05)	−.09	(.05)	**.41**	(.05)
Ideas (o5)	−.01	(.02)	−.03	(.04)	.07	(.04)	.03	(.03)	**.55**	(.04)
Values (o6)	.00	(.04)	.05	(.05)	−.08	(.05)	.10	(.05)	**.38**	(.06)
Trust (a1)	−.10	(.08)	.19	(.08)	−.11	(.05)	**.49**	(.06)	.03	(.06)
Straightforwardness (a2)	−.08	(.05)	.16	(.07)	.21	(.05)	.26	(.06)	−.03	(.05)
Altruism (a3)	.02	(.03)	.13	(.06)	.13	(.04)	**.39**	(.05)	.04	(.04)
Compliance (a4)	−.02	(.04)	.00	(.04)	.17	(.05)	**.39**	(.05)	.01	(.04)
Modesty (a5)	.11	(.06)	-.07	(.06)	.08	(.05)	**.44**	(.05)	.06	(.04)
Tender-mindedness (a6)	.14	(.08)	.02	(.08)	.01	(.03)	**.59**	(.05)	.03	(.05)
Competence (c1)	.08	(.04)	.02	(.04)	**.52**	(.04)	−.01	(.03)	.09	(.06)
Order (c2)	−.01	(.03)	.07	(.05)	**.61**	(.04)	.00	(.03)	−.13	(.06)
Dutifulness (c3)	.04	(.03)	.06	(.05)	**.58**	(.04)	.08	(.04)	−.05	(.05)
Achievement (c4)	−.01	(.04)	.07	(.05)	**.54**	(.04)	−.03	(.03)	−.01	(.04)
Self-discipline (c5)	−.09	(.04)	−.05	(.05)	**.61**	(.04)	−.03	(.06)	−.03	(.03)
Deliberation (c6)	.00	(.03)	−.11	(.06)	**.46**	(.06)	.18	(.07)	.07	(.05)

Note. FFMRF = Five-Factor Model Rating Form; N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness. Factor loadings ≥|.30| are in boldface. Exploratory structural equation modeling was conducted with robust maximum likelihood estimation and geomin rotation. $N = 757$.

of a facet from another domain was −.15 for the positive emotions facet of extraversion. The second factor was defined by loadings of the six extraversion facets ranging from .35 (positive emotions) to .60 (gregariousness). The loading of the impulsivity facet from neuroticism (.33) was the only other that was larger than .30. The third factor was defined by loadings for the six facets of conscientiousness ranging from .46 (deliberation) to .61 (competence). The impulsivity facet from neuroticism also displayed a sizeable negative loading (−.30) on this factor. The fourth factor was more mixed as it included sizeable loadings of five of the agreeableness facets. Although the facet of straightforwardness (.26) failed to reach this threshold, the other agreeableness facets ranged from .39 (altruism) to .59 (tendermindedness). This factor also evinced loadings from the facets of warmth (.34), assertiveness (−.39), and angry hostility (−.39). The fifth and final factor was defined by five facets of openness, as well as the excitement seeking facet from extraversion (.30). The facet of openness to feelings failed to load appreciably on this factor (.16).

Convergent and Discriminant Validity

Table 3 presents the correlation matrix between the domain scores from FFMRF and the NEO PI-R. The convergent coefficients for the domains ranged from .50 (openness) to .68 (neuroticism), with a mean of .58. The mean discriminant correlations for the FFMRF domains were much lower and ranged from .09 (openness) to .20 (neuroticism), with a mean value across the domains of .13. The largest discriminant correlation between any two domains was −.30

Table 3. Correlation Matrix Between FFMRF and NEO PI-R Domain Scores

	NEO PI-R domains				
FFMRF domains	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
Neuroticism	.68	−.29	.02	−.18	−.30
Extraversion	−.19	.55	.10	.02	.09
Openness	.02	.16	.50	.05	−.14
Agreeableness	−.09	.19	.06	.51	.16
Conscientiousness	−.21	.17	−.05	.13	.64

Note. FFMRF = Five-Factor Model Rating Form; NEO PI-R = NEO Personality Inventory–Revised. Minimum pairwise $n = 563$. All correlations >|.11| are significant at $p < .01$.

between FFMRF neuroticism and NEO PI-R conscientiousness.

Table 4 presents the convergent correlations and summarizes the discriminant values for the FFMRF facets. Discriminant values for the facets were calculated in two ways: first the average correlation with other facets from the same domain and then the mean of the absolute values for correlations with all facets from other domains. The use of absolute values provides a more conservative test in that large discriminant correlations in opposite directions cannot cancel each other out. In all cases, the mean values were calculated by first transforming the correlations to *z*-scores using Fisher's method, then averaging and converting back to correlations. In each case, the minimum and maximum of the discriminant correlations also are provided.

The first column of Table 4 provides the convergent correlations, indicating the degree of one-to-one correspondence between the measures. The convergent values for the facets were considerably more variable than for the domains, ranging from a low of .17 (tendermindedness) to .64 (order), with a mean value across of .41 across the 30 facets. In all cases the convergent correlations were larger than the average discriminant correlations with facets from other domains. These mean discriminant values with facets from other domains ranged from .05 (openness to aesthetics) to .15 (depressiveness), with a mean value of .10. When examining the range of discriminant correlations of each individual FFMRF facet, it is apparent that there was only one that was larger than the convergent correlation. This was for the FFMRF facet of tendermindedness (from agreeableness) that correlated more highly with the NEO PI-R extraversion facet of warmth ($r = .21$) than it did with NEO PI-R tendermindedness ($r = .17$).

The discriminant values with other NEO PI-R facets from the same domain were understandably larger given they assess related, but independent traits. Nonetheless, all but one of the mean discriminant values inside the home domain were lower than the convergent values. The one exception was again tendermindedness, which obtained a mean value of .23 with other facets from its home domain of agreeableness. However, there were three additional facets for which the FFMRF correlated more highly with a discriminant NEO PI-R facet within the same domain than with the target scale. FFMRF openness to ideas correlated more highly with NEO PI-R openness to fantasy (.33) than the convergent correlation (.28). Similarly, the FFMRF agreeableness facet of compliance related more strongly to NEO PI-R altruism (.28) than its home facet (.27), and the FFMRF conscientiousness facet of competence correlated more strongly with NEO PI-R achievement striving (.31) than its convergent value (.29). Tendermindedness, though, was the most problematic as this FFMRF facet correlated more highly with four NEO PI-R facets (straightforwardness, altruism, modesty, and compliance) than the NEO PI-R facet of the same name.

Discussion

An important validity test for any abbreviated measure is that it demonstrates a factor structure that is consistent with that of the parent instrument (Smith et al., 2000). This is even more crucial in the case of an instrument designed to assess a model such as the FFM, where the higher order factor structure is a hallmark. The current study demonstrated that the FFMRF displayed an adequate fit within a five-factor structure using ESEM and that the loadings of the individual facets within this solution were largely consistent with their theoretical placements. In fact, the fit indices for the FFMRF in the current study were even higher than those obtained for the more established NEO-FFI in two published studies (i.e., Marsh et al., 2010; Rosellini & Brown, 2011). In addition to the hierarchical structure, the FFMRF displayed convergent validity correlations that were large at the domain level (e.g., >.50; Cohen, 1992). Although, there were exceptions, the convergent validity of the FFMRF facets was also quite reasonable (e.g., 24 of the 30 facets obtained values >.30 with an overall mean of .41). The discriminant validity, evidenced by the mean absolute value of correlations outside the home domain, was also quite good, with a value of .13 for the domains and .10 for the facets. Only a single FFMRF facet (tendermindedness) correlated more highly with a NEO PI-R facet outside its home domain, than it did with the convergent scale. These findings build on existing evidence

Table 4. Convergent and Discriminant Correlations Between FFMRF and NEO PI-R

		Discriminant same domain			Discriminant other domains		
FFMRF facets	Conv.	*M*	Minimum	Maximum	*M*	Minimum	Maximum
Anxiousness (n1)	.62	.39	.19	.50	.10	−.21	.15
Angry Hostility (n2)	.52	.15	.08	.20	.13	−.45	.05
Depressiveness (n3)	.58	.36	.21	.44	.15	−.34	.17
Self-Consciousness (n4)	.50	.36	.18	.49	.13	−.31	.20
Impulsivity (n5)	.36	.12	.05	.22	.14	−.42	.21
Vulnerability (n6)	.41	.30	.19	.38	.10	−.28	.16
Warmth (e1)	.38	.16	.08	.25	.09	−.12	.28
Gregariousness (e2)	.49	.28	.25	.39	.07	−.17	.23
Assertiveness (e3)	.39	.11	−.02	.21	.12	−.42	.22
Activity (e4)	.38	.30	.29	.35	.12	-.26	.28
Excitement Seeking (e5)	.44	.20	.13	.28	.10	−.22	.22
Positive Emotions (e6)	.37	.23	.18	.31	.13	−.29	.25
Fantasy (o1)	.52	.18	.15	.23	.08	−.19	.18
Aesthetics (o2)	.36	.23	.14	.31	.05	−.06	.12
Feelings (o3)	.25	.09	.05	.15	.07	−.05	.21
Actions (o4)	.24	.13	.04	.22	.09	−.27	.23
Ideas (o5)	.28	.25	.19	.33	.06	−.13	.15
Values (o6)	.36	.08	−.03	.16	.07	−.17	.09
Trust (a1)	.50	.19	.05	.28	.09	−.35	.32
Straightforwardness (a2)	.31	.14	.05	.23	.12	−.14	.24
Altruism (a3)	.39	.24	.17	.29	.13	−.30	.32
Compliance (a4)	.27	.21	.13	.28	.10	−.27	.23
Modesty (a5)	.31	.17	.07	.25	.08	−.20	.15
Tender-mindedness (a6)	.17	.23	.11	.31	.10	−.18	.21
Competence (C1)	.29	.22	.13	.31	.07	−.15	.19
Order (c2)	.64	.34	.29	.43	.09	−.18	.14
Dutifulness (c3)	.41	.34	.26	.40	.10	−.21	.18
Achievement (c4)	.51	.34	.21	.41	.11	−.22	.29
Self-discipline (c5)	.47	.38	.32	.47	.11	−.23	.17
Deliberation (c6)	.37	.21	.18	.27	.09	−.16	.20
Mean[a]	.41	.23			.10		

Note. FFMRF = Five-Factor Model Rating Form; NEO PI-R = NEO Personality Inventory–Revised; Conv. = convergent correlations; Discriminant same domain = mean absolute value of discriminant correlations with other facets from the same domain; Discriminant other domain = the mean absolute value of discriminant correlations with all facets from other domains. Minimum pairwise $N = 598$. All correlations $>|.11|$ are significant at $p < .01$. All convergent values were significant at $p < .001$.

a. Mean values of correlations were first transformed via Fisher's *r* to *z* method, then averaged and returned to *r* values.

to indicate that the FFMRF provides a concise and efficient assessment of the FFM's higher order domains as well as the more specific facets proposed by Costa and McCrae (1992). Taken together, this suggests the FFMRF should be considered when one is interested in assessing the higher and lower order traits of the FFM, but constraints prohibit the administration of a lengthier measure.

Hierarchical Structure

As indicated above, the FFMRF demonstrated an adequate fit with the five-factor solution and the specific loadings were largely consistent with theoretical placements. Nonetheless, there were instances in which specific facets evinced significant cross-loadings on additional factors and at least one case where a facet failed to load appreciably on the predicted factor. There were six facets that obtained a sizeable (≥.30) loading on a factor other than that predicted. However, in most of these instances the cross-loading of certain facets was understandable given their content and was consistent with prior research on FFM instruments. For example, the facets of warmth and assertiveness from extraversion and angry hostility from neuroticism also loaded on agreeableness. The finding that warmth would load highly on a factor defined by agreeableness is hardly surprising or without precedent in the literature. Research has routinely demonstrated a "fuzzy boundary" (John et al., 2008, p. 136) between agreeableness and extraversion

and warmth has often been at the center of this, both literally and figuratively.

When these two domains are placed on orthogonal axes, one can examine where individual traits fall, and this type of research has demonstrated that the term warmth falls within the interstitial space between the two domains (Hofstee, de Raad, & Goldberg, 1992; John et al., 2008). The most well-known of these circles is the interpersonal circumplex (Wiggins, 1979), which is a 90-degree rotation of agreeableness and extraversion, and explicitly labels the upper pole warmth. Thus, although Costa and McCrae (1992) placed the facet of warmth within the domain of extraversion, they also obtained notable cross-loadings with agreeableness. Other measures have placed it within agreeableness (e.g., Saucier & Ostendorf, 1999), and research from the lexical tradition has generally indicated that warmth is closer to agreeableness than extraversion (Goldberg, 1992; John, 1990). The current results support the primary placement of warmth within extraversion, but again indicate that it relates nearly as highly with agreeableness.

Similarly, angry hostility is placed within the domain of neuroticism on the FFMRF and the current results support this placement. However, it also loaded strongly (−.39) on the factor defined by agreeableness. Likewise, the facet of assertiveness, although assigned to extraversion, also loaded negatively on agreeableness (−.39). Again it is not particularly surprising as factor analyses of other FFM instruments also show these cross-loadings (e.g., Costa & McCrae, 1992; Trull et al., 1998) and these facets often correlate highly negatively with measures of agreeableness (John et al., 2008).

The current factor analysis also suggested meaningful secondary cross-loadings of other facets including impulsivity and excitement-seeking. The facet of impulsivity loaded primarily on its intended domain of neuroticism (.37), but evinced secondary loadings with conscientiousness (−. 30) and extraversion (.33). This is again consistent with previously published findings as the impulsivity facet from the Structured Interview for the FFM (SIFFM; Trull & Widiger, 1997) actually obtained a primary negative loading with conscientiousness (Trull et al., 1998). The difficulty of placing the facet labeled impulsivity may stem from the fact that the construct of impulsivity is quite broad. Whiteside and Lynam (2001) identified four facets from the NEO PI-R that assessed distinct aspects of what others have termed impulsivity. These included the facets of impulsivity (neuroticism), as well as excitement-seeking (extraversion), deliberation (conscientiousness), and self-discipline (conscientiousness). Thus, it not particularly surprising that the FFMRF facet of impulsivity would also cross-load with the domains of conscientiousness and extraversion.

Continued research on the FFMRF's assessment of these particular facets will be helpful in determining if they achieve expected correlations with measures beyond the NEO PI-R and/or other outcomes. In any event, the current study suggested that the FFMRF evinces a five-factor structure that is largely consistent with its conceptual organization. In fact, the structural validity of the FFMRF appears to be as high as other existing measures of the FFM. As indicated previously, the ESEM fit indices for the five-factor solution obtained in the present study were higher than those obtained for the NEO-FFI in two previous publications (e.g., Marsh et al., 2010; Rosellini & Brown, 2011). Although we are not aware of any published studies that used ESEM to examine the fit for the NEO PI-R or the SIFFM, exploratory factor analysis of these instruments display structures similar to what we found for the FFMRF.

Nearly all of the FFMRF facets achieved a robust, primary loading on their home domain. The primary exception was the facet of openness to feelings, which loaded only .16. Nonetheless, this particular facet appeared not to fit well with the FFMRF at all, as it also did not load any higher on another factor. This particular facet has been problematic for the FFMRF in prior research (see Mullins-Sweatt et al., 2006) and has been revised previously to address this issue. However, in the current study it did not load cleanly on the openness factor and achieved among the lowest convergent relationships with the NEO PI-R. Although its discriminant validity was adequate, this might reflect that it does not correspond well with any of the other content on the NEO PI-R, whether within openness or across the instrument. Thus, the performance of this item suggests that results concerning openness to feelings on the FFMRF should be interpreted carefully and that additional revisions might be considered.

Convergent and Discriminant Validity

The current study suggested that the FFMRF evinced large convergent correlations at the domain level that were consistent with the values reported by Mullins-Sweatt et al. (2006). The mean convergent value of .58 for the domains suggests that the content assessed by the FFMRF is comparable to that assessed by the NEO PI-R. The convergent values for the facets were also quite large in many instances (i.e., r = .64 for the conscientiousness facet of order). However, there were exceptions. For example, the FFMRF facet of tendermindedness achieved relatively poor convergence with the NEO PI-R (r = .17). The findings for the openness facets of feelings, actions, and ideas as well as the agreeableness facet of compliance were also lower than might be expected. Although single items are undoubtedly handicapped in this regard, the adequate convergence for several other facets raises the bar and suggests that the FFMRF's assessment of these traits might differ in important ways from the NEO PI-R.

It is notable in this regard that three facets of openness were among the most problematic. Not all of openness had this issue as, for example, the facet of openness to fantasy obtained a convergent value of .52. Nonetheless, there may be something specific that limits the convergence between

the NEO PI-R and FFMRF assessments of openness. One possible explanation for the limited convergence is that the FFMRF was originally developed as a measure to be used by clinicians. This might serve to decrease the convergent values in two ways. First, the FFMRF includes some terms that connote specific meanings to professionals but that may be more difficult for laypersons to comprehend and respond to effectively. This is particularly salient for a few facets of openness (e.g., *alexythymic*). An investigation of the reading level required for comprehension of the FFMRF would be useful to address this possibility. Second, the adjectives included on the FFMRF often reflect more pathological terms that are hypothesized to be maladaptive, extreme variants of the same traits assessed by the NEO PI-R (Widiger & Trull, 2007). Although a body of research has provided relatively consistent support for this hypothesis (e.g., Samuel & Widiger, 2008a), the findings for certain aspects have been varied (e.g., the relationship of high openness with oddity/eccentricity; Watson, Clark, & Chmielewski, 2008). Thus, some FFMRF adjectives might not have perfect one-to-one correspondences with the NEO PI-R's assessment of these traits (Haigler & Widiger, 2001).

It is important to note in this regard that even if this were the case it might not necessarily reflect a limitation of the FFMRF. The NEO PI-R is only a single representation of the FFM and there are other measures that also provide an assessment of these same traits. Thus, research that examines the correlation of the FFMRF with other FFM measures beyond the NEO PI-R (and those assessing similar constructs, such as the HEXACO PI-R) would be useful in determining the extent to which the FFMRF representation of these traits is unique from those designed to assess the more normative aspects.

The discriminant values for the FFMRF also were mostly in line with expectations. There was appropriate convergence among different facets within domains ($M = .23$) and more limited relationships to facets from other domains ($M = .10$). Although three facets (compliance, competence, and ideas) obtained a discriminant correlation within the same domain that was larger than the convergent value, these differences were slight (i.e., .28 vs. .27). The only FFMRF facet that can be said to have truly problematic discriminant validity is tendermindedness. Not only was the mean correlation with other NEO PI-R agreeableness facets higher than the convergent value, but FFMRF tendermindedness even correlated more highly with the NEO PI-R facet of warmth (from extraversion) than with its own companion facet. As discussed previously, the similarity between the facet of warmth and aspects of agreeableness (particularly tendermindedness) is not unexpected. Nonetheless, this convergent value from the current study (.17) is notably lower than what was obtained by Mullins-Sweatt et al. (2006), so future research is needed to determine if this limited convergence will replicate in additional samples.

Limitations and Future Directions

Taken together, the current findings concerning the hierarchical structure and convergence with an established measure provided additional support for the validity of the FFMRF. This suggests that the FFMRF is an abbreviated measure that maintains the higher order structure of the FFM while also providing an effective assessment of the lower order facets. The inclusion of the lower order facets is crucial for the utility of the FFMRF, as they have been shown to be useful for differentiating constructs (Reynolds & Clark, 2001) and predicting specific behaviors (Paunonen et al., 2003). Nonetheless, the current article was limited by the fact that a majority of the participants were undergraduate students. Although a clinical group did comprise a portion of our combined sample, it would be useful to replicate and extend these findings using a community sample with a greater range of ages and other demographic variables as well as a variety of reading levels.

An additional property of the FFMRF that could have affected the current findings was the fact that the FFMRF items are grouped according the domains to which they are assigned. It is possible that such a format could increase the cohesion of the factor structure. This should be investigated in future research that uses alternative ordering of the FFMRF items. Nonetheless, the fact that a number of FFMRF facets did display secondary loadings that are consistent with previous factor analytic research does suggest that participants rated facets at least somewhat independently.

The FFMRF is emerging as a potentially useful choice for individuals wishing to obtain a valid and efficient assessment of the FFM. Currently, the 44-item Big Five Inventory (BFI; John, Donahue, & Kentle, 1991) is the most commonly employed short-form measure for assessing personality within psychology research (John et al., 2008). This use is warranted as the BFI provides a well-validated, freely available, and brief measure of the FFM domains. In fact, the convergence between the BFI and NEO-FFI in prior research ($M = .77$; John & Soto, 2007) was higher than that obtained for the FFMRF and NEO PI-R in the current study ($M = .58$). Thus, it would be unreasonable to suggest the FFMRF should replace the BFI in instances where only domain scores are necessary. However, a possible limitation of the BFI is that it does not allow the assessment of lower order traits or facets. Thus, it is important that the current results suggest the FFMRF provides a valid assessment of the five higher order factors but also yields information about lower order traits. Given the increased predictive validity demonstrated for facet-level measures in general, future research that directly compares the BFI and the FFMRF in terms of their respective abilities to predict important behavioral outcomes associated with personality would be quite helpful.

Similarly, the NEO-FFI is also an established measure of the FFM domains that is used quite frequently. It has the advantage of being explicitly tied to the NEO PI-R (indeed, the NEO-FFI is simply a subset of the NEO PI-R items), but like the BFI is limited by only providing an assessment of the FFM domains. Thus, it would also be useful for future studies to compare the NEO-FFI (along with the BFI) with the FFMRF in terms of comparative validity.

It would also be useful for future research to investigate other important properties of the FFMRF. For example, one potential concern with single-item assessments is their temporal consistency, so studies that report test–retest values for the FFMRF are necessary. This should include both short- (i.e., a week) and long-term (i.e., 6 months or more) retest intervals to gather information about both the dependability (Watson, 2004) and consistency (Roberts, Wood, & Caspi, 2008) of the measure. Another potentially useful analysis for short-form measures is to quantify the time-savings versus the validity cost relative to a longer measure (Smith et al., 2000). In other words, research should investigate the time it takes individuals to complete the FFMRF relative to a longer inventory, such as the NEO PI-R, and then compare the two in terms of their validity. This would provide information so that assessors could make informed decisions about the trade-off between efficiency and validity when considering a short form.

Conclusions

The current study builds on previous research to demonstrate that the FFMRF has a higher order structure that corresponds well with its prescribed FFM framework and evinces large convergent correlations with the NEO PI-R domains. Although some FFMRF facets deserve further scrutiny, most demonstrated defensible convergent and discriminant validity. These findings indicate that the FFMRF can be considered a robust measure of the FFM domains that also includes the added benefit of assessing specific, lower order traits. Further research is needed, but our findings suggest the FFMRF is a reasonable alternative for those seeking an abbreviated measure of the FFM that includes both broad and specific traits. Nonetheless, we emphasize that any abbreviated measure offers psychometric trade-offs in exchange for time savings, and the complete NEO PI-R is recommended when resources permit. However, when barriers prevent the administration of a lengthy instrument and lower order traits are desired, the one-page FFMRF appears to be a practical option for assessing personality.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article:

This research was partially supported by a fellowship (MH074245) from the National Institute of Mental Health, awarded to the first author.

References

Allik, J. (2005). Personality dimensions across cultures. *Journal of Personality Disorders, 19*, 212-232.

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review, 11*, 150-166.

Axelrod, S. R., Widiger, T. A., Trull, T. J., & Corbitt, E. M. (1997). Relation of five-factor model antagonism facets with personality disorder symptomatology. *Journal of Personality Assessment, 69*, 297-313.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research, 36,* 111-150.

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology, 56*, 453-484.

Cloninger, C. R. (2008). The psychobiological theory of temperament and character: Comment on Farmer and Goldberg (2008). *Psychological Assessment, 20*, 292-299.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment, 64*, 21-50.

Deary, I. J., Weiss, A., & Batty, G. D. (2011). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest, 11*(2), 53-79.

DeYoung, C. G., Hirsh, J. B., Shane, S., Papademetris, X., Rajeevan, N., & Gray, J. R. (2010). Testing predictions from personality neuroscience: Brain structure and the Big Five. *Psychological Science, 21*, 820-828.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93*, 880-896.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116*, 429-456.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*, 26-42.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7-28). Tilburg, Netherlands: Tilburg University Press.

Haigler, E. D., & Widiger, T. A. (2001). Experimental manipulation of NEO PI-R items. *Journal of Personality Assessment, 77*, 339-358.

Hofstee, W. K. B., de Raad, B., & Goldberg, L. R. (1992). Integration of the Big-Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology, 63*, 146-163.

Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review, 14*, 332-346.

Hopwood, C. J., Morey, L. C., Ansell, E. B., Grilo, C. M., Sanislow, C. A., McGlashan, T. H., . . . Skodal, A. E. (2009). The convergent and discriminant validity of five-factor traits: Current and prospective social, work, and recreational dysfunction. *Journal of Personality Disorders, 23*, 466-476.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and Research* (pp. 66-100). New York, NY: Guilford Press.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 54.* Berkeley: University of California, Berkeley, Institute of Personality and Social Research.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality* (3rd ed., pp. 114-158). New York, NY: Guilford Press.

John, O. P., & Soto, C. J. (2007). The importance of being valid: Reliability and the process of construct validation. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 461-494). New York, NY: Guilford Press.

Krueger, R. F., & Johnson, W. (2008). Behavioral genetics and personality: A new look at the integration of nature and nurture. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality* (3rd ed., pp. 287-310). New York, NY: Guilford Press.

Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research, 39*, 329-358.

Lowe, J. R., & Widiger, T. A. (2009). Clinicians' judgments of clinical utility: A comparison of the DSM-IV with dimensional models of general personality. *Journal of Personality Disorders, 23*, 211-229.

Lynam, D. R., & Widiger, T. A. (2001). Using the five-factor model to represent the DSM-IV personality disorders: An expert consensus approach. *Journal of Abnormal Psychology, 110*, 401-412.

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology, 88*, 139-157.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralising Hu & Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320-341.

Marsh, H. W., Ludtke, O., Muthen, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22*, 471-491.

McCrae, R. R., & Costa, P. T. (2008). The five factor theory of personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality* (3rd ed., pp. 159-181). New York, NY: Guilford.

McCrae, R. R., Terracciano, A., & 78 Members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology, 88*, 547-561.

McCrae, R. R., Zonderman, A. B., Costa, P. T., Jr., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the revised NEO Personality Inventory: Confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology, 70*, 552-566.

Mullins-Sweatt, S. N., Jamerson, J. E., Samuel, D. B., Olson, D. R., & Widiger, T. A. (2006). Psychometric properties of an abbreviated instrument of the Five-Factor Model. *Assessment, 13*, 119-137.

Mullins-Sweatt, S. N., & Widiger, T. A. (2011). Clinicians' judgments of the utility of the DSM-IV and Five-Factor models for personality disordered patients. *Journal of Personality Disorders, 25*, 463-477.

Mullins-Sweatt, S. N., & Widiger, T. A. (2010). Personality-related problems in living: An empirical approach. *Personality Disorders: Theory, Research, & Treatment, 1*, 230-238.

Muthen, L. K., & Muthen, B. O. (2011). *Mplus user's guide.* Los Angeles, CA: Muthen & Muthen.

Ozer, D. J., & Benet-Martinez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology, 57*, 401-421.

Paunonen, S. V., & Ashton, M. C. (2001). Big five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology, 81*, 524-539.

Paunonen, S. V., Haddock, G., Fosterling, F., & Keinonen, M. (2003). Broad versus narrow personality measures and the prediction of behaviour across cultures. *European Journal of Personality, 17*, 413-433.

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

Reynolds, S. K., & Clark, L. A. (2001). Predicting dimensions of personality disorder from domains and facets of the Five-Factor Model. *Journal of Personality, 69*, 199-222.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*, 3-25.

Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality* (3rd ed., pp. 375-398). New York, NY: Guilford Press.

Rosellini, A. J., & Brown, T. A. (2011). The NEO Five-Factor Inventory: Latent structure and relationships with dimensions of anxiety and depressive disorders in a large clinical sample. *Assessment, 18*, 27-38.

Samuel, D. B., & Widiger, T. A. (2004). Clinicians' personality descriptions of prototypic personality disorders. *Journal of Personality Disorders, 18*, 286-308.

Samuel, D. B., & Widiger, T. A. (2006). Clinicians' ratings of clinical utility: A comparison of the DSM-IV and Five Factor Models. *Journal of Abnormal Psychology, 115*, 298-308.

Samuel, D. B., & Widiger, T. A. (2008a). 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, 28*, 1326-1342.

Samuel, D. B., & Widiger, T. A. (2008b). Convergence of narcissism measures from the perspective of general personality functioning. *Assessment, 15*, 364-374.

Samuel, D. B., & Widiger, T. A. (2009). Comparative gender biases in models of personality disorder. *Personality and Mental Health, 3*, 12-25.

Samuel, D. B., & Widiger, T. A. (2010a). Comparing personality disorder models: Cross-method assessment of the FFM and DSM-IV-TR. *Journal of Personality Disorders, 24*, 721-745.

Samuel, D. B., & Widiger, T. A. (2010b). A comparison of obsessive-compulsive personality disorder scales. *Journal of Personality Assessment, 92*, 232-240.

Samuel, D. B., & Widiger, T. A. (2011a). Clinicians' use of personality disorder models within a particular treatment setting: A longitudinal comparison of temporal consistency and clinical utility. *Personality and Mental Health, 5*, 12-28.

Samuel, D. B., & Widiger, T. A. (2011b). Conscientiousness and obsessive-compulsive personality disorder. *Personality Disorders: Theory, Research, and Treatment, 2*, 161-174.

Saucier, G., & Goldberg, L. R. (2002). Assessing the Big Five: applications of 10 psychometric criteria to the development of marker scales. In B. De Raad & M. Perugini (Eds.), *Big five assessment* (pp. 29-58). Bern, Switzerland: Hogrefe & Huber.

Saucier, G., & Ostendorf, F. (1999). Hierarchical subcomponents of the Big Five personality factors: A cross-language replication. *Journal of Personality and Social Psychology, 76*, 613-627.

Smith, G. S., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment, 12*, 102-111.

Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681-706). Hillsdale, NJ: Lawrence Erlbaum.

Trull, T. J., & Widiger, T. A. (1997). *Structured Interview for the Five-Factor Model of Personality.* Odessa, FL: Psychological Assessment Resources.

Trull, T. J., Widiger, T. A., Useda, J. D., Holcomb, J., Doan, B.-T., . . . Gershuny, B. S. (1998). A structured interview for the assessment of the five-factor model of personality. *Psychological Assessment, 10*, 229-240.

Vassend, O., & Skrondal, A. (1997). Validation of the NEO Personality Inventory and the five-factor model: Can findings from exploratory and confirmatory factor analysis be reconciled? *European Journal of Personality, 11*, 147-166.

Watson, D. (2004). Stability versus change, dependability versus error: Issues in the assessment of personality over time. *Journal of Research in Personality, 38*, 319-350.

Watson, D., Clark, L. A., & Chmielewski, M. (2008). Structures of personality and their relevance to psychopathology: II. Further articulation of a comprehensive unified trait structure. *Journal of Personality, 76*, 1545-1585.

Whiteside, S. P., & Lynam, D. R. (2001). The five factor model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences, 30*, 669-689.

Widiger, T. A., De Clercq, B., & De Fruyt, F. (2009). Childhood antecedents of personality disorder: an alternative perspective. *Development and Psychopathology, 21*, 771-791.

Widiger, T. A., & Trull, T. J. (2007). Plate tectonics in the classification of personality disorder: Shifting to a dimensional model. *American Psychologist, 62*, 71-83.

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37*, 395-412.

Wirth, J. H., Lynam, D. R., & Williams, K. D. (2010). When social pain is not automatic: Personality disorder traits buffer ostracism's immediate negative impact. *Journal of Research in Personality, 44*, 397-401.

Yamagata, S., Suzuki, A., Ando, J., One, Y., Kijima, N., Yoshimura, K., . . . Spinath, F. M. (2006). Is the genetic structure of human personality universal? A cross-cultural twin study from North America, Europe, and Asia. *Journal of Personality and Social Psychology, 90*, 987-998.