

## **Measuring Personality in Business: The General Personality Factor in the Mini IPIP**

**Thomas R. Carretta**  
**United States Air Force Research Laboratory**

**Malcolm James Ree**  
**Our Lady of the Lake University**

**Mark S. Teachout**  
**University of the Incarnate Word**

*The Mini-IPIP is a 20 item measure of the Big Five personality domains, used in business for pre-employment selection. A study was conducted to determine if this short version of the IPIP had a general personality factor as found in longer Big Five measures. A sample of 269 adult males and females was administered the Mini-IPIP. Confirmatory factor analyses examined several models. A single general personality factor (GFP) model fit the data well as did a two factor model (DeYoung, Peterson, & Higgins, 2002; Digman, 1997), and a higher-order general personality factor (GFP) model (Musek, 2007; Rushton & Irwing, 2008). The current analyses confirmed the general factor. Implications for research and practice are discussed, including a call for investigating the practical application in business and theoretical consistency of the general personality factor.*

### **INTRODUCTION**

The contemporary study of personality has arrived at a consensus that traits can be reliably organized into five broad domains called the Big Five (Costa & McCrae, 1992). The Big Five domains are: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Neuroticism is sometimes called by its positive counterpart, Emotional Stability. Uses of the Big Five range from clinical evaluations (Nigg, Blaskey, Huang-Pollock, Hinshaw, John, Willcutt, & Pennington, 2002; Watson, Clark, & Harkness, 1994) to modeling or predicting job performance (Barrick & Mount, 1991; Ones, Mount, Barrick, & Hunter, 1994; Tett, Jackson, & Rothstein, 1991). The use of personality measures in the prediction of occupational outcomes has been well established. Barrick and Mount (1991) conducted a meta-analysis of the predictive efficiency of the Big Five personality domains for training and job performance across 117 studies. They found that conscientiousness was consistently related to training proficiency and job performance in every criterion and occupational category measured and estimated the meta-analytic correlation between conscientiousness and job performance at  $\rho = .22$ . Tett et al. (1991) refined the Barrick and Mount estimate to  $\rho = .26$ . Barrick, Mount, and Judge (2001) noted “results support the previous findings that conscientiousness is a valid predictor across performance measures in all occupations studied” (p. 9).

Goldberg (1981) pioneered lexical theory which became a foundation for the *NEO Personality Inventory* (*NEO-PI*; Costa & McCrae, 1985) that measured the domains of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. Using the *NEO-PI-R* (*NEO-PI –Revised*; Costa & McCrae, 1992) as an example, each of the five domains has six facets that are subscales measuring a particular construct. The sum of the six facets yields the domain score. There are 240 items in the *NEO-PI-R* and administration takes about one hour. The *NEO Five-Factor Inventory* (*NEO-FFI*) is a 60 item version of the *NEO-PI-R*. It assesses only the domains, but can be administered in about 15 minutes. Non-commercial versions of Big Five inventories also are available. Among the most frequently used for research is the 100 item *International Personality Item Pool*, *IPIP*, (Goldberg, Johnson, Eber, Hogan, Ashton, Cloninger, & Gough, 2006). There is also a 50 item version of the *IPIP* (International Personality Item Pool, 2011), a 20 item *Mini-IPIP* (Donnellan, Oswald, Baird, & Lucas, 2006) and a *Ten Item Personality Inventory* (*TIPI*; Gosling, Rentfrow, & Swann, 2003). Business reasons for shortening the Big Five instrument include saving administration time, reducing fatigue, and allowing other instruments to be administered in a specific time period.

Digman (1997) suggested two hierarchical factors for the Big Five in which Alpha was defined by Neuroticism, Agreeableness, and Conscientiousness, and Beta was defined by Extraversion and Openness. DeYoung, Peterson, and Higgins (2002) in their study relabeled these factors as Stability and Plasticity. Musek (2007) confirmed a general factor of personality (*GFP*) that was hierarchically superior to Alpha/Stability and Beta/Plasticity in the Big Five model in three samples. Rushton and Irwing (2008), using meta-analytic correlations from 16 separate studies totaling 2,916 participants, reported that measures of the Big Five have a hierarchical structure with two second-level factors identical to Digman's and a single third-level general personality factor, *GFP*.

Ashton, Lee, Goldberg, and de Vries (2009) demonstrated that hierarchical factors can be found from uncorrelated broad factors if arranged optimally among observed scores. They did not offer a practical empirical example from the literature.

Findings by van der Linden, Scholte, Cillessen, te Nijenhuis, and Seger (2010) supported the existence of a general personality factor including "likeability" and "popularity" in a sample of adolescents. de Vries (2011) reanalyzed their data and concluded that support for a general personality factor was unwarranted due to analytic issues such as Heywood Cases and better fit of non-hierarchical models. Revelle and Wilt (2010) argued that the *GFP* accounted for much less common variance than the general factor in cognitive ability (*g*) and is difficult to interpret. Just (2011) reviewed the literature regarding the *GFP* and concluded that "Data from a number of highly diverse measures of personality have revealed the position of the *GFP* atop the hierarchy of personality structure..." (p. 770).

While there has been substantial research on the longer versions of Big Five measures, there has been less research on shorter versions. Cooper, Smillie, and Corr (2010) conducted a confirmatory factor analysis of the *Mini-IPIP* items and concluded that a five factor model reflecting the Big Five domains was the most plausible model. However, there has not been an investigation of the general personality factor of the *Mini-IPIP*. If the longer and shorter versions are measuring the same personality constructs, then there should be a *GFP*. If the longer and shorter versions are not the same, the test scores cannot be interpreted in the same manner and it is likely that the tests are not measuring the same construct. Hence, if the tests are not comparable, comparisons, accumulation of results, and inferences, will be misleading.

The purpose of the current study was to determine if the *Mini-IPIP* revealed a general personality factor. Additionally, we sought to determine whether a hierarchical structure, commensurate with longer measures of the Big Five, similar to Digman, (1997), DeYoung et al. (2002), Musek (2007), and Rushton and Irwing (2008) was plausible in this short version. This will also help determine if the longer and shorter measures assess the same constructs and permit meaningful comparisons and inferences.

## METHOD

### Participants

The participants were 269 adult males (38%) and females (62%) from metropolitan San Antonio, Texas ranging in age from 18 to 81 with an average of 44 years. Their average job experience was 23 years. Educational attainment was 1.1% high school graduate, 8.4% some college, 24.9% college graduate, and 65.6% post-graduate. By ethnicity, the participants were 47.6% Hispanic, 39.6% white, 9.5% African American, and 3.3% other. All were employed full time in a broad array of jobs.

### Instruments

The 100 item *IPIP* is a public domain instrument that produces scores for the Big Five (Goldberg et al., 2006). The *Mini-IPIP* is a 20 item version of a Big Five instrument (Donnellan et al., 2006). It was developed because of the need for a measure of the Big Five domains that took less time to administer. The developers recognized that unlike the 100 item *IPIP* or the *NEO-PI-R*, 20 items would not allow for the replication of the 30 facet scores. Because reduced administration time was an important goal, only the five domain scores were included in the *Mini-IPIP*.

Correlations of the domain score of the *IPIP* and the *Mini-IPIP* were used as a developmental criterion (Donnellan et al., 2006). The correlations between the *IPIP* and the *Mini-IPIP* for like-named domain scores from a development study of 2,663 participants are: Conscientiousness ( $r = .90$ ), Agreeableness ( $r = .89$ ), Neuroticism ( $r = .92$ ), Openness to Experience ( $r = .85$ ), and Extraversion ( $r = .93$ ).

### Analyses

Multiple confirmatory factor analyses (CFAs) were conducted on the five domain scores of the *Mini-IPIP* using LISREL 8.7 (Jöreskog & Sörbom, 2004). The CFAs used maximum likelihood estimation. Four CFAs were performed. The first CFA sought to determine if the data could be well represented by a single factor. The second applied Digman's (1997) two factor (Stability and Plasticity) model. The third model applied Rushton and Irwing's (2008) hierarchical model, with intermediate factors representing Stability and Plasticity and a higher-order *GFP*.

The fourth model employed the unequivocal test of Erdle, Irwing, Rushton, and Park (2010). It tested for the existence of a general personality factor by specifying the two factors of Stability and Plasticity as uncorrelated. This test first estimates the model with Digman's two factors uncorrelated and then estimates it with the factors correlated. A difference in fit indicates the most likely structure. A better fit for the uncorrelated factors removes the possibility of the hierarchical *GFP*. A better fit for the correlated factors provides support for the hierarchical *GFP*.

Several fit statistics were used:  $\chi^2$ , Root Mean Square Error of Approximation, (RMSEA), Comparative Fit Index (CFI), Goodness of Fit Index (GFI), Standardized Root Mean Square Residual (SRMR), and Critical N (CN). We applied the values for determining quality of fit as suggested by the literature (Hu & Bentler, 1999; Lance & Vandenberg, 2009).

## RESULTS

Table 1 summarizes the means and standard deviations of the *Mini-IPIP* domain scores. Table 2 presents the correlations of the scores. After reflecting the negative correlations with Neuroticism, the correlations between the domain scores ranged from .007 (Neuroticism and Extraversion) to .293 (Agreeableness and Openness) with a mean correlation of .141.

**TABLE 1**  
**MEANS AND STANDARD DEVIATIONS FOR THE MINI-IPIP DOMAIN SCORES**

Score	Mean	SD
Agreeableness (A)	16.42	2.33
Openness (O)	16.10	2.54
Extraversion (E)	15.04	3.01
Conscientiousness (C)	17.00	2.27
Neuroticism (N)	11.87	2.54

**TABLE 2**  
**CORRELATIONS AMONG THE MINI-IPIP DOMAIN SCORES**

Domain	A	O	E	C	N
Agreeableness (A)	1.000				
Openness (O)	0.293	1.000			
Extraversion (E)	0.261	0.171	1.000		
Conscientiousness (C)	0.178	0.133	0.048	1.000	
Neuroticism (N)	-0.186	-0.099	0.007	-0.036	1.000

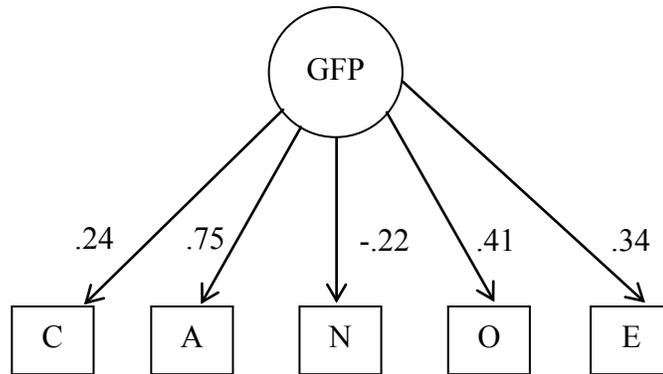
**CFAs**

Results of the CFAs are presented in Table 3. The first CFA, which sought to determine if the data could be represented by a single general personality factor, indicated a good fit of the model to the data. LISREL modification indexes suggested that the error terms of the Openness and Agreeableness observed variables should be permitted to correlate. This was done, but gave unacceptable results with a negative error variance for Agreeableness and a strong negative correlation ( $r = -.87$ ) between Agreeableness and Openness. These two results were deemed unacceptable and work on this correlated-errors model was discontinued. The single general personality factor CFA that accounted for 32% of the common variance was retained. This model is presented in Figure 1

**TABLE 3**  
**FIT INDEXES FOR THE CFA MODELS**

	One Factor	Digman (1997)	Rushton & Irwing (2008)	Non-Hierarchical
$X^2$ df p	3.84 5 p<.57	3.20 4 p<.52	3.20 3 p<.36	41.53 6 p<.00
RMSEA ( lower-upper)	0.00 (0 -.074)	0.00 (0 -.080)	0.01 (0 -.110)	0.14 (.10-.190)
SRMR	.027	.025	.025	.11
CFI	1.00	1.00	1.00	.46
GFI	1.00	1.00	1.00	.95
CN	1111.73	1020.85	950.28	109.49

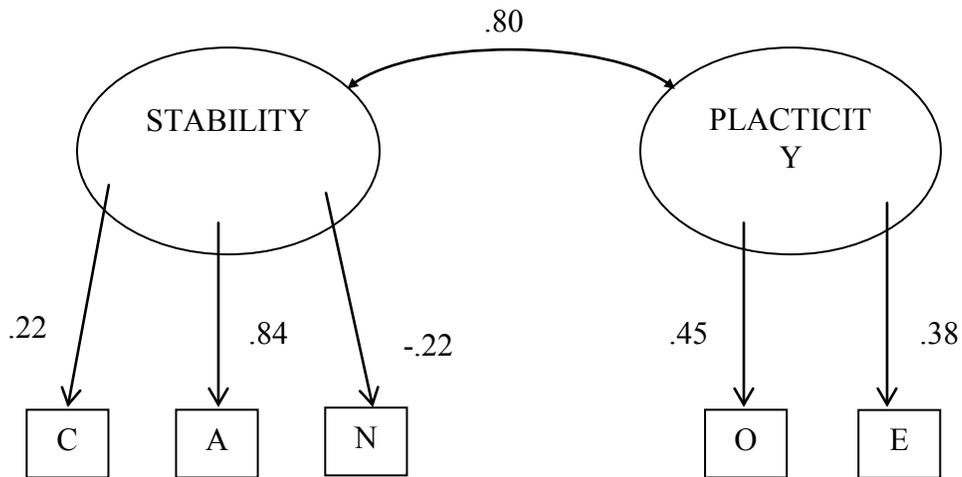
**FIGURE 1**  
**FACTOR LOADINGS FOR A GENERAL PERSONALITY FACTOR**



Note: The factors are the general factor of Personality (GFP), conscientiousness (C), agreeableness (A), neuroticism (N), openness (O), and extraversion (E).

The second CFA applied Digman's (1997) two factor model (i.e., Alpha and Beta, latter called Stability and Plasticity) to the data. This model also fit the data quite well. See Figure 2.

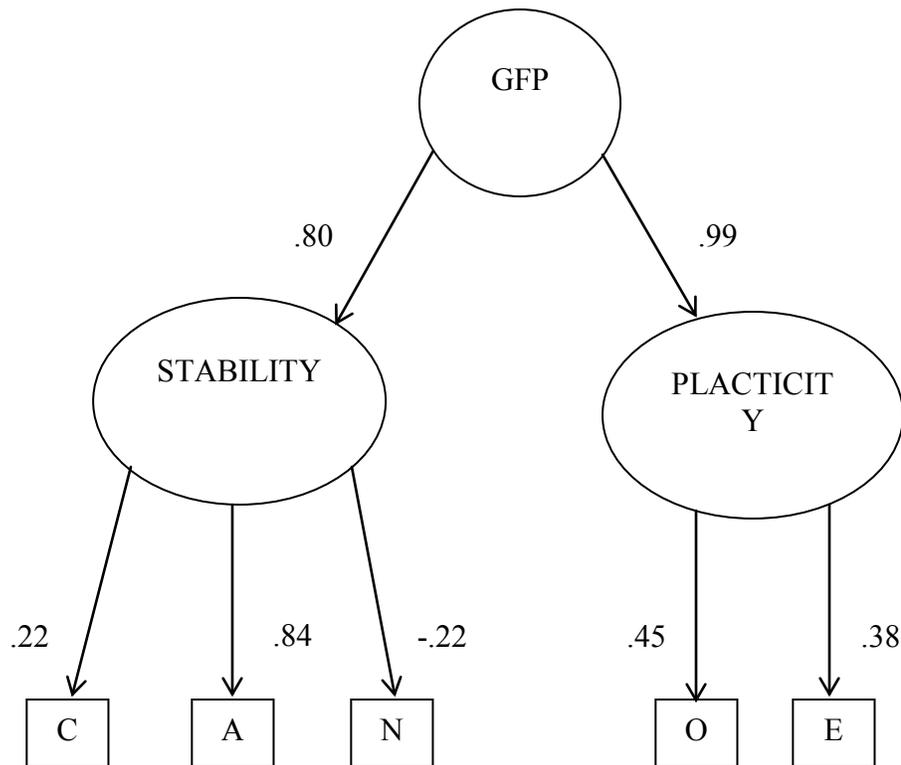
**FIGURE 2**  
**FACTOR LOADINGS FOR THE MODEL WITH TWO FIRST-ORDER FACTORS**



Note: The factors are Stability, Plasticity, conscientiousness (C), agreeableness (A), neuroticism (N), openness (O), and extraversion (E).

A third CFA evaluated a higher level general factor of personality (*GFP*) as suggested by Musek (2007) and Ruston and Irwing (2008). Both lower-order factors were influenced by the higher-order factor. The fit was very good. Figure 3 illustrates this model and factor loadings.

**FIGURE 3**  
**FACTOR LOADINGS FOR THE HIERARCHICAL MODEL**



Note: The factors are general factor of personality (*GFP*), Stability, Plasticity, conscientiousness (C), agreeableness (A), neuroticism (N), openness (O), and extraversion (E).

Finally, following the unequivocal test of Erdle et al. (2010), a CFA was conducted in which the two first-order factors were specified to be uncorrelated. Consistent with Erdle et al. the fit was very poor, further supporting a general personality factor.

## DISCUSSION

Cooper, Smillie, and Corr (2010) conducted a confirmatory factor analysis of the *Mini-IPIP* items and concluded that a five factor model reflecting the Big Five domains was the most plausible model. The current study is the first investigation of the general factor and hierarchical structure of the *Mini-IPIP*. Results confirmed the existence of a general personality factor in both a one factor model and a hierarchical factor model with Stability and Plasticity similar to that of longer Big Five measures (DeYoung et al., 2002; Digman, 1997; Musek, 2007; Rushton & Irwing, 2008). Because the *GFP* and the higher order factor structures of the longer and shorter versions are the same, they are generally measuring the same personality constructs, and can be interpreted in much the same manner. Hence,

comparisons are likely to be more meaningful. Inferences are likely to be consistent with studies using the longer versions once results are corrected for unreliability. In addition, a better understanding of the similarity could be made by meta-analyses including both the long and short versions and conducting moderator analyses of long versus short versions on the corrected correlations. Finally, using the shorter Big Five instrument makes business sense, saving administration time, reducing applicant or employee fatigue, and allowing other instruments to be administered in a finite time period.

Given the good fit of the one factor model and the results found by Cooper et al. (2010), the *Mini-IPIP* seems to be more homogenous than expected. One reason may be due to the process for developing the *Mini-IPIP*. The developers of the *Mini-IPIP* followed the best professional practices. They used four good items from each domain and maximized the inter-item consistency via factor analytic results. Reducing the *IPIP* to 20 items necessarily left out content found in some of the facets for each domain.

Erdle et al. (2010) observed “There was no plausible alternative to a model without a general factor” (p. 344). We observed much the same. The existence of hierarchical factors has implications for practice and research. The fundamental issue is determining the appropriate level of score to use from the *Mini-IPIP*. Is it the five domains, the two hierarchical factors, or the single general factor? The two hierarchical factors and the general factor are expected to be more reliable than the domains, but the domains are more familiar to practitioners and researchers. If the purpose is to maximize the prediction of important outcome variables such as job performance, then the hierarchical factors or the general factor may be more useful. If the purpose is to understand the factors that make up personality then the five domains may be more useful.

Several questions must be answered. First, is the general factor or are the hierarchical factors the sources of valid variance analogous to the factor of general mental ability (*g*) in cognitive tests? If so, then development of future Big Five measures can concentrate on establishing the hierarchical factors rather than trying to measure the specific sources of variance. This would facilitate construction of future Big Five measures.

Determination of the sources of validity, general factor, hierarchical factors, or the domain specific factors, can be found by comparing the validity of the hierarchical factors to the validity of the hierarchical factors and the specific factors. An increase in the  $R^2$  would indicate a contribution to validity by the specific factors. See Ree and Earles (1991) for an example.

Is the general factor the source of validity? The answer to this question has theoretical as well as practical implications. The interpretation of this factor would need to be developed and tested across a range of situations. If the general factor is the source of validity it suggests that *theories* of personality would need to be adapted to reflect this result.

It is also important to estimate the mean and variance differences on all the factors between men and women and among ethnic groups. This would establish normative data for these groups on each factor to facilitate the ongoing understanding of expected results. This also would permit the assessment of adverse impact to see if it is more or less likely with the factors compared to the five domain scores.

Further, it would be important to know how much of the differences between domain scores were due to hierarchical factors and how much were due to lower-order factors. This would be especially important to interpretation of profiles that portray differences in a graphic display to highlight high and low scores. Profile misinterpretation could occur if differences due to hierarchical factors were interpreted as differences due to lower-order factors.

A highly focused and rigorous stream of research investigating the practical application and theoretical consistency of personality factors, and factor structure is needed. We join Just (2011) in calling for multiple routes for future *GFP* research.

## NOTE

<sup>1</sup>The opinions expressed are those of the authors and are not necessarily those of the United States Government, the Department of Defense, or the United States Air Force.