

# Towards Understanding Assessments of the Big Five: Multitrait-Multimethod Analyses of Convergent and Discriminant Validity Across Measurement Occasion and Type of Observer

Jeremy C. Biesanz

University of Wisconsin—Madison

Stephen G. West

Arizona State University

**ABSTRACT** Multitrait-multimethod analyses were used to examine the degree of convergent and discriminant validity of the Big Five. Phase 1 examined self-reports of the Big Five across three measurement occasions. Self-reports of the Big Five traits were stable, but were moderately intercorrelated. Phase 2 examined assessments of the Big Five across different types of informants (self, peer, and parent). Assessments converged across types of informants and, importantly, there was no evidence of correlation between the Big Five traits across the perspectives of different types of informants. The present results suggest that the degree of orthogonality of the Big Five traits depends on the source of the data. A single informant produces Big Five traits that are intercorrelated, whereas diverse informants tend to produce a much more orthogonal structure. Discussion focuses on methodological considerations in examining levels of convergent and discriminant validity and the theoretical implications

Jeremy C. Biesanz, Department of Psychology, University of Wisconsin—Madison and Stephen G. West, Department of Psychology, Arizona State University. This work was partially supported by the National Institute of Child Health and Human Development (HD007376).

We thank William Graziano, William Hoyt, and Abigail Panter for their valuable comments on previous versions of this paper. Please address correspondence to Jeremy C. Biesanz, who is now at the Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, B.C., Canada V6T 1Z4, or [jbiesanz@psych.ubc.ca](mailto:jbiesanz@psych.ubc.ca).

*Journal of Personality* 72:4, August 2004.  
Blackwell Publishing 2004

for understanding personality assessments and the relationship between three-and five-factor models of personality.

One of the primary tasks within personality research is to identify the important dimensions along which people differ. In pursuit of this endeavor, considerable evidence suggests that natural language descriptors of personality comprise five broad traits, called the Big Five, or the five factors (for historical reviews and theoretical perspectives, see Digman, 1996; John, 1990; McCrae & Costa, 1999; Saucier & Goldberg, 1996; Wiggins & Trapnell, 1996, 1997; for empirical evidence see Costa & McCrae, 1988a; Fiske, 1949; Goldberg, 1992; Norman, 1963; Tupes & Christal, 1961/1992). These five broad traits—Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Openness to Experience—have provided a framework that has helped spark a resurgence of interest and research in personality.

The vast array of research provided in support of the Big Five is quite impressive. These same five factors emerge from factor analyses of self-reports as well as reports of knowledgeable informants (e.g., Goldberg 1990). The longitudinal stability of these broad traits has been well documented and is quite substantial (Costa & McCrae, 1988a; Roberts & DelVecchio, 2000). Self-reports converge with reports from knowledgeable informants for both measures of relative consistency (e.g., McCrae & Costa, 1987) as well as coherence (Biesanz & West, 2000). Finally, assessments of Big Five traits predict important outcomes and behaviors (e.g., Barrick & Mount, 1991; Friedman et al., 1995).

One of the major benefits of the Big Five framework has been to bring clarity to the diverse array of proposed personality measures. During the decades after World War II, measures of individual personality traits and broad inventories of personality proliferated. Commentators such as Sechrest (1986) called for the adoption of a benchmark against which other measures could be evaluated. The Big Five has served this purpose, providing a common yardstick that has permitted comparison of different inventories and their predictive correlates. Ostensibly different personality inventories really were assessing the same underlying broad constructs (for reviews, see John, 1990; John & Srivastava, 1999). The cacophony of different inventories jangling their benefits has been reduced into a common rubric that holds forth the promise of integration and

consolidation of knowledge within one of the levels and domains of personality (McAdams, 1995).

The Five-Factor Model, however, is not without its critics (e.g., Block, 1995; Eysenck, 1992). One oft-leveled criticism centers on the interpretation of these broad five factors: Does the Five-Factor Model reflect the highest order constructs underlying individual differences as its proponents maintain (e.g., McCrae & Costa, 1999) or are these five factors simply facets of even broader constructs (e.g., Digman, 1997; Eysenck, 1992, 1994)? In other words, how are we to understand these dimensions on which people differ from each other? Whether the Big Five are independent, orthogonal dimensions of personality or whether they are instead related in a manner that reflects other important, even broader dimensions is one focus of the present article. We begin by briefly reviewing theoretical and methodological issues related to understanding potential relationships among the Big Five, evidence suggesting higher-order factors above the Big Five, and then methods for understanding relationships among these broad traits.

### Are the Big Five Orthogonal?

Both Goldberg (1993a, 1993b) and Costa and McCrae (1995) conceptualize the Big Five as being essentially orthogonal. The work establishing five factors has generally used orthogonal factor rotation procedures such as varimax in exploring the factor structure in personality ratings (e.g., Tupes & Christal, 1961/1992) and this practice has continued (e.g., Benet-Martínez & John, 1998; Goldberg, 1992; McCrae & Costa, 1985). Orthogonal rotations by definition guarantee that the resulting factors will be uncorrelated. They provide only very indirect evidence about the degree of relationship between factors. Non-orthogonality can potentially appear in the form of substantial loadings on secondary factors. Although orthogonal rotations of Big Five inventories generally do not seem to produce obvious problems with large secondary loadings (e.g., Benet-Martínez & John, 1998; Goldberg, 1992), there are nontrivial cross-loadings for some of the facets of the revised NEO Personality Inventory (Costa & McCrae, 1992a, 1995). Fabrigar, Wegener, MacCallum, and Strahan (1999) have suggested that far more direct information is provided through confirmatory factor analysis procedures and exploratory factor analysis procedures that use an initial oblique

rotation. Researchers can then examine the factor intercorrelations and determine the degree to which orthogonality has or has not been achieved.

At the scale level, however, the Big Five do not appear to be orthogonal and independent dimensions. Small to moderate cross loadings can lead to correlations among the scales constructed for each factor on the basis of exploratory factor analysis. Different instruments such as Goldberg's (1992) trait adjectives and the NEO-PI-R (Costa & McCrae, 1992a) demonstrate moderate to substantial intercorrelations between factor scales (see Digman, 1997). It is quite clear that the Big Five, *as currently assessed*, are not even close to orthogonal (Panter, Tanaka, & Hoyle, 1994). Given substantial correlations between the Big Five factors (e.g., median = .30, range = .23 to .62, Graziano & Ward, 1992, Table 1), it is not surprising that criticism has been leveled at the Big Five that the promised parsimony is a promise yet unfulfilled (e.g., Bandura, 1999; Block, 1995).

Researchers have consequently examined the intercorrelations among the Big Five to propose even broader constructs – higher-order factors that reside above the Big Five. For example, Digman (1997) factor-analyzed correlation matrices among the Big Five and proposed that Agreeableness, Conscientiousness, and Neuroticism are facets of a broader construct, labeled  $\alpha$ , and Extraversion and Openness represent facets of the broader construct, labeled  $\beta$  (see also Costa & McCrae, 1992b). Eysenck's P-E-N theory posits three broad factors and Eysenck (1992, 1994) has argued that Agreeableness and Conscientiousness are facets of the broader construct Psychoticism. In advocating a Big Three model, Clark and Watson (1999) have recently argued that the construct of Disinhibition versus Constraint is a complex combination of Agreeableness and Conscientiousness. Each of these competing models is based on the critical premise that the Big Five traits are correlated so that more general personality dimensions can be built from combinations of the Big Five.

### Multitrait-Multimethod (MTMM) Analyses and the Big Five

Questions concerning the orthogonality of each pair of the Big Five's broad traits—or meaningful lack thereof – reflect nuanced questions of discriminant validity. Campbell and Fiske's (1959) development

**Table 1**  
**Examining Stability within Self-Reports: Multitrait-Multioccasion Correlation Matrix (N = 339)**

Time	Self-Report (Time 1)					Self-Report (Time 2)					Self-Report (Time 3)				
	AG	CO	EX	NE	OP	AG	CO	EX	NE	OP	AG	CO	EX	NE	OP
Time 1															
AG	(.90)														
CO	.40	(.89)													
EX	.22	.21	(.89)												
NE	-.33	-.20	-.20	(.84)											
OP	.33	.44	.42	-.17	(.84)										
Time 2															
AG	.72	.33	.21	-.25	.31	(.93)									
CO	.32	.77	.19	-.14	.40	.48	(.90)								
EX	.20	.20	.80	-.20	.40	.35	.28	(.91)							
NE	-.26	-.14	-.16	.68	-.15	-.47	-.27	-.29	(.88)						
OP	.28	.40	.38	-.10	.80	.43	.51	.46	-.16	(.86)					
Time 3															
AG	.64	.30	.21	-.14	.32	.73	.41	.26	-.28	.34	(.92)				
CO	.24	.69	.13	-.10	.35	.34	.76	.16	-.16	.43	.44	(.91)			
EX	.20	.21	.75	-.13	.38	.28	.26	.80	-.20	.41	.34	.26	(.90)		
NE	-.23	-.15	-.13	.66	-.13	-.26	-.16	-.16	.70	-.06	-.33	-.24	-.26	(.87)	
OP	.25	.34	.38	-.11	.78	.34	.41	.41	-.14	.82	.41	.46	.47	-.15	(.86)
SD	1.02	1.05	1.05	1.06	0.83	1.06	1.07	1.10	1.16	0.85	1.00	1.11	1.04	1.11	0.88
Mean	5.70	5.21	4.79	4.22	5.28	5.71	5.28	4.91	4.08	5.28	5.73	5.26	5.00	3.98	5.30

*Note:* Cronbach's alphas are printed on the main diagonal in parentheses. Correlations between different traits across different measurement occasions are printed in gray. Stability coefficients — correlations between the same trait assessed at different measurement occasions — are represented by the black diagonals within the gray blocks.

of the multitrait-multimethod matrix (MTMM) provided a framework for a rigorous examination of questions concerning convergent and discriminant validity. By examining relationships among different traits across different measures, trait effects can be disentangled from potential method effects. The convergence of a trait across different assessment methods helps strengthen the interpretation of that trait as representing the underlying construct of interest. Similarly, the divergence of different traits, assessed with the same or different methods, is used to demonstrate that they can be discriminated and that those traits represent separate constructs.

Campbell and Fiske's (1959) original approach to convergent and discriminant validity and their analysis of correlation matrices focused on providing binary yes/no answers to two questions: Is there convergence among different measures of the same trait? Can ostensibly different traits be discriminated? Early work on the Big Five applied Campbell and Fiske's original rules for the examination of correlation matrices (see West & Finch, 1997) and repeatedly found clear support for convergent validity of these five factors (McCrae & Costa, 1987; see also Kenrick & Funder, 1988). Similarly, it is quite clear that the Big Five represent separate, distinguishable dimensions. However, whether the Big Five represent orthogonal dimensions or are meaningfully related through higher-order factors is not a question that can be answered using the traditional Campbell and Fiske approach. Rather, we must turn to other statistical methods of investigating personality structure to address this question (Finch & West, 1997).

Developments in confirmatory factor analysis have allowed researchers to apply these techniques to address questions concerning *the degree* of convergent and discriminant validity in MTMM matrices in a unified and coherent framework (Marsh & Grayson, 1995; Reichardt & Coleman, 1995). A review of MTMM confirmatory factor analyses of the Big Five, however, suggests that intercorrelations among the Big Five factors remain even after removing method effects (e.g., Benet-Martinez & John, 1998; Borkenau & Ostendorf, 1990; John and Srivastava, 1999; see also Finch, Panter, & Caskie, 1999; Panter et al., 1994). These studies clearly demonstrate convergence among different inventories of the Big Five. These studies also suggest, at first glance, that the Big Five are not orthogonal, independent dimensions: Moderate factor intercorrelations remain even after removing method effects.

However, understanding these observed relationships among the Big Five remains difficult as each study includes multiple self-reports using different inventories. Under these measurement designs, method effects associated with self-assessment cannot be disentangled from the true scores on the specific traits. Self-reports, as Kenny (1994) notes, “carry a great deal of excess baggage” (p. 202). Consequently, it is possible that self-assessments of the Big Five are intercorrelated because of the use of a common self-report methodology rather than the existence of higher-order factors.

Furthermore, interpreting these observed relationships among the Big Five is difficult as different inventories were used. Block (1995) and John and Srivastava (1999) have noted that the major Big Five inventories each assess a slightly different conception of each of the Big Five traits. Such inventory-specific differences in the conceptions of the Big Five traits can attenuate the strength of relationships *between* traits compared to multiple self-assessments using the same inventory.

How then should the degree of discriminant validity be examined? In personality research, convergence among the perspectives of different reporters is frequently used to help validate constructs. The degree to which bias is shared across reporters is largely eliminated so that an improved estimate of convergent validity can be obtained. Similarly, relationships between different reporters using the same inventories can be used to establish levels of discriminant validity among different dimensions. To examine the degree of discriminant validity among the Big Five dimensions of personality, assessments from different knowledgeable informants (i.e., self, peer, and parent) need to be examined *using the same inventory*. If the Big Five are meaningfully related in a manner reflecting higher-order factors, then these relationships will continue to emerge across assessments from different informants. On the other hand, if the Big Five are truly orthogonal dimensions, then the Big Five will not intercorrelate across different reporters once potential bias from the use of the same informant is eliminated.

### Overview

The present investigation examines the degree of convergent and discriminant validity for the Big Five using confirmatory factor analyses of multitrait-multimethod matrices. For ease of presentation, we have conceptually divided the present data into two phases

even though they are based on the same large sample of participants. In Phase 1, measurement occasions are examined as the “method.” Three separate self-assessments of the Big Five using the same inventory are examined to determine the degree of temporal convergence and the degree of discriminant validity for reports of the Big Five based on data from a single informant. Phase 1 thus represents a multitrait, multioccasion investigation (Backteman & Magnusson, 1981; Conley, 1985; Steyer, Schmitt, & Eid, 1999) of the Big Five. Under this approach, each latent trait represents the stable aspect of self-reports across measurement occasions, and the reliability of each assessment is expected to be quite high.

In Phase 2, informant is examined as the method. Self-, peer, and parental assessments of the Big Five, using the same inventory, are examined to determine the degree of convergent and discriminant validity across different knowledgeable informants. Under this approach, each latent trait now represents the shared common perspective across these three different informants, and the reliability of each assessment is expected to be lower than in Phase 1. Phase 2 represents a multitrait, multi-informant investigation of the Big Five.

## Phase 1

### METHOD

#### Participants

Introductory psychology students ( $N = 387$ ) were recruited to participate in return for partial fulfillment of their introductory psychology class requirements. A total of 339 participants completed the basic study requirements of attending three measurement sessions (226 women and 113 men; mean age = 19.48 years,  $SD = 3.05$ ). Note that a subset (69%) of these participants were analyzed in a different context (Biesanz & West 2000, Study 2). MTMM confirmatory factor analysis models have frequently failed to converge to proper solutions in the literature, in part due to inadequate sample sizes (Kenny & Kashy, 1992). Consequently, we have aggregated all available data accrued over 3 separate years of data collection to achieve the recommended sample size of at least 250 (Marsh & Bailey, 1991; Marsh & Grayson, 1995).

#### Materials, Design, and Procedure

Participants rated themselves on 97 unipolar trait adjectives — 20 for Agreeableness, 19 for Conscientiousness, 20 on Extraversion, 18 on Neu-



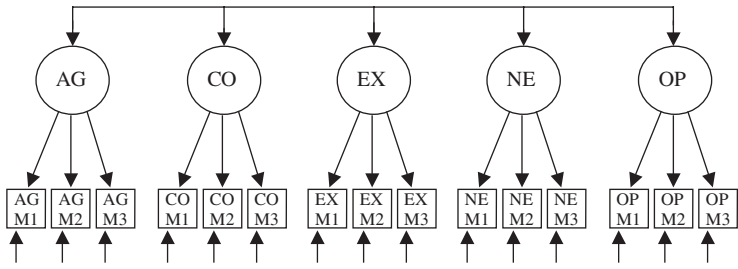
roticism and 20 for Openness to Experience (Goldberg, 1992). Three trait adjectives proposed by Goldberg (1992)—*imperturbable*, *haphazard*, and *unexcitable*—were not included because they were unfamiliar or confusing to a large proportion of the respondents (cf. Biesanz & West, 2000; Graziano, Jensen-Campbell, Steele & Hair, 1998). All ratings were on a 9-point scale ranging from (0) extremely inaccurate to (8) extremely accurate. Participant's self-rating instructions were modified from Goldberg (1992) to limit self-assessments of behavior to the previous week (see Biesanz, West, & Graziano, 1998, for the specific instructions). Participants completed the self-report inventory three times, at no less than 1-week intervals, in a lecture hall reserved for that purpose.

### Analytic Strategy

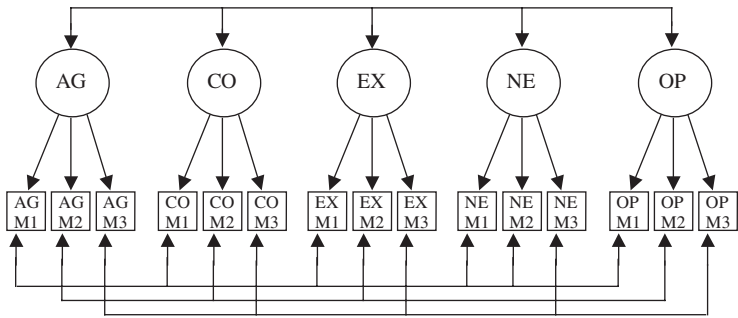
When using multitrait-multimethod, confirmatory factor analysis models, the factor loadings, variances, and relationships among the latent traits are used to interpret the level or degree of convergent and discriminant validity. The degree of convergent validity is evidenced through the size of the trait factor loadings as well as the amount of variance present in the latent trait factors. The degree of discriminant validity present is determined through the estimated relationships among the latent traits.

We present the overall results from seven different multitrait-multimethod, confirmatory factor analytic models and include the four models advocated by Marsh and Grayson (1995). Some of these analyses represent various baseline models with which the models of central interest are compared.

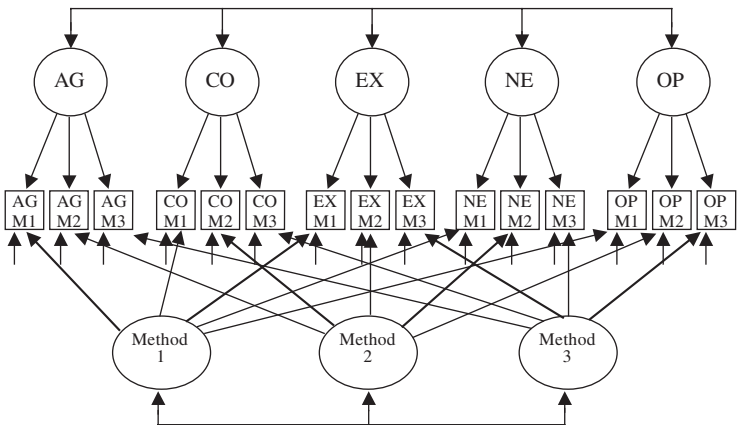
1. The correlated traits, no methods model (CTNM) depicted in Figure 1A presumes five latent trait factors corresponding to each of the Big Five, but makes no assumptions about and does not model the structure of the method effects. The trait factors are permitted to be correlated.
2. The correlated traits, correlated uniquenesses model (CTCU) depicted in Figure 1B represents method effects as correlations between the uniquenesses of variables measured with a common method (in Phase 1, at the same measurement occasion). No assumptions are made concerning the dimensionality or structure of the method effects. The trait factors are permitted to be correlated. The CTCU model does assume that method effects are uncorrelated with each other.
3. The orthogonal traits, correlated uniquenesses model (OTCU) assumes that the five trait factors are independent. This model is nested under the CTCU model and consequently the difference in fit



A. Correlated Traits, No Methods (CTNM) Model.



B. Correlated Traits, Correlated Uniquenesses (CTCU) Model.



C. Correlated Traits, Correlated Methods (CTCM) Model.

**Figure 1**

Three multitrait-multimethod confirmatory factor analytic models examining the Big Five across three different assessment methods (M1, M2, and M3).

between these two models presents a test of the orthogonality of the Big Five.

4. The correlated traits, correlated methods model (CTCM) depicted in Figure 1C represents an alternative to the structure assumed by the CTCU model presented in Figure 1B. The CTCM model assumes that a *single* latent factor underlies each method and estimates the relationships among the method factors. Note that the CTCM model is *not* nested under and thus cannot be directly compared to either the CTCU or the OTCU models.
5. The correlated traits, orthogonal methods model (CTOM) also assumes that a single latent factor underlies each method (measurement occasion), but – in contrast to the CTCM model – presumes that the three method factors are uncorrelated. The CTOM model is nested under the CTCM as well as the CTCU model (but not the OTCU). Consequently, the comparison between the CTOM and the CTCU presents a test of whether the method effects are unidimensional. In addition, the comparison between the CTOM and the CTCM presents a test of whether method effects are uncorrelated with each other.
6. The orthogonal traits, correlated methods model (OTCM) presumes that there is a single latent method factor and estimates the relationships among the method factors. In addition, this model constrains the traits to be orthogonal and thus is nested under the CTCM model.
7. The orthogonal traits, orthogonal methods model (OTOM) presumes that there is a single latent methods factor and that the methods are unrelated. This model is nested under the CTCU, OTCU, CTCM, CTOM, and OTCM models.

All models presented were analyzed using the full covariance matrix and estimated using EQS v. 5.6 under maximum likelihood (Bentler, 1995) and verified with LISREL v. 8.3 (Jöreskog & Sörbom, 1996) with no material differences in the solutions. Before presenting the MTMM confirmatory factor analyses, we first briefly examine the MTMM correlation matrix using Campbell and Fiske's (1959) approach, recognizing the limitations of this approach (Marsh & Grayson, 1995; West & Finch, 1997).

## RESULTS

### Traditional MTMM Analysis: Time as a Method Factor

The MTMM correlation matrix of self-reported Big Five is presented in Table 1. The test-retest correlations (same trait, different times)

of the traits were substantial (mean  $r = .74$ ; range = .64 to .82). Within a given assessment, there are moderate correlations among the Big Five (different traits, same time; mean absolute  $r = .34$ ; range of absolute  $r$ s = .15 to .48). Correlations among the Big Five were lower when examined across different assessments (different traits, different times; mean absolute  $r = .25$ ; range of absolute  $r$ s = .06 to .43). Thus, there is clear evidence of convergent and discriminant validity in this multitrait, multioccasion investigation.

### Confirmatory Factor MTMM Analyses: Multiple Times as $\alpha$ Methods

Table 2 presents the results of the series of multitrait, multioccasion confirmatory factor analyses for each of the models described above. We present the  $\chi^2$ , df, CFI, RMSEA, and SRMSR to provide information about the fit of each model (see Hu & Bentler, 1998). We focus below on the correlated traits correlated uniquenesses model (CTCU) that best reproduces the analyzed covariance matrix. The CTCU model provides an excellent fit to the data,  $\chi^2(50, N = 339) = 67.11$ , *ns*, CFI = .996, RMSEA = .032, SRMSR = .030. The parameter estimates for the CTCU model (illustrated in Figure 1B) are presented in Table 3. The factor loadings for the first assessment are fixed at 1.00 to establish the scale of measurement for the latent factors.

*Convergent Validity.* In the CTCU model, where measurement occasion serves the role of method, the latent Big Five factors represent self-reports that are unassociated with time-specific errors of measurement. Two clear findings emerge from examination of the latent trait factors. First, the factor loadings for the second and third assessments are close to 1.00, substantial, and statistically significant, and thus indicate that the measurement structure did not change appreciably over time. Second, the variances of the Big Five latent factors are all significant, indicating that individual differences in self-reports are stable over time, replicating substantial previous research.

*Discriminant Validity.* In the CTCU model, the Big Five latent factors are all significantly intercorrelated. The Orthogonal Traits Correlated Uniquenesses (OTCU; Model 3 in Table 2) is nested within the CTCU model, and therefore the difference between these models is an omnibus test of the orthogonality of self-reports of the

**Table 2**  
 Confirmatory Factor Analysis Model Fit Indices of the Multitrait-Multioccasion Matrix ( $N = 339$ )

Model	$\chi^2$	df	CFI	RMSEA	SRMSR
1. CTNM: Correlated Traits No Method Factors	629.99**	80	.863	.143	.040
2. CTCU: Correlated Traits Correlated Uniquenesses	67.11	50	.996	.032	.030
3. OTCU: Orthogonal Traits Correlated Uniquenesses	327.06**	60	.933	.115	.235
4. CTCM: Correlated Traits Correlated Methods	103.73**	62	.990	.045	.024
5. CTOM: Correlated Traits Orthogonal Methods	112.61**	65	.988	.047	.030
6. OTCM: Orthogonal Traits Correlated Methods	187.18**	72	.971	.069	.091
7. OTOM: Orthogonal Traits Orthogonal Methods	377.50**	75	.924	.109	.231

Note: CFI: Comparative Fit Index. RMSEA: Root Mean Square Error of Approximation. SRMSR: Standardized Root Mean Square Residual.

\*\* $p < .01$ .

Big Five. The chi-square difference test comparing models 2 and 3 indicates that the hypothesis of orthogonal factors can be clearly rejected,  $\chi^2(10, N = 339) = 259.95, p < .0001$ . Thus, even when time-specific method effects are removed, there are moderate to strong relationships among self-report measures of the Big Five (mean absolute  $r = .34$ ).

*Uniquenesses.* The uniquenesses in the CTCU model represent a combination of pure measurement error variance and a reliable variance component of each trait assessment that is specific to the measurement occasion. The variance of each uniqueness is significant, and within each measurement occasion these uniquenesses are

**Table 3**  
 Examining Time as a Method Factor: Correlated Traits Correlated  
 Uniquenesses (CTCU) Parameter Estimates of the Multitrait-  
 Multioccasion Matrix

Parameter	Latent Trait				
	AG	CO	EX	NE	OP
Factor Loadings					
Time 1 (fixed)	1.00	1.00	1.00	1.00	1.00
<i>standardized</i>	.82	.86	.87	.81	.88
Time 2	1.07*	1.05*	1.10*	1.09*	1.05*
<i>standardized</i>	.88	.91	.92	.83	.92
Time 3	0.96*	0.99*	0.98*	1.09*	1.04*
<i>standardized</i>	.82	.82	.87	.84	.89
Latent Factor Correlations					
Agreeableness	1.00				
Conscientiousness	.45*	1.00			
Extraversion	.31*	.27*	1.00		
Neuroticism	-.35*	-.19*	-.23*	1.00	
Openness	.41*	.51*	.49*	-.14*	1.00
Variance	.72*	.84*	.84*	.75*	.55*
Uniqueness Correlations					
Time 1					
Agreeableness	1.00				
Conscientiousness	.46*	1.00			
Extraversion	.20*	.21*	1.00		
Neuroticism	-.35*	-.24*	-.18*	1.00	
Openness	.27*	.37*	.24*	-.12	1.00
Variance	.36*	.31*	.27*	.40*	.17*
Time 2					
Agreeableness	1.00				
Conscientiousness	.38*	1.00			
Extraversion	.45*	.25*	1.00		
Neuroticism	-.61*	-.48*	-.39*	1.00	
Openness	.46*	.29*	.31*	-.27*	1.00
Variance	.25*	.19*	.17*	.42*	.11*
Time 3					
Agreeableness	1.00				
Conscientiousness	.39*	1.00			
Extraversion	.27*	.21*	1.00		
Neuroticism	-.40*	-.30*	-.43*	1.00	
Openness	.34*	.37*	.33*	-.19*	1.00
Variance	.33*	.40*	.27*	.37*	.16*

\* $p < .05$ .

intercorrelated, indicating that the time-specific method effects within self-reports are associated with each other across the Big Five. Note that Neuroticism is scored so that positive scores are associated with less socially desirable values, accounting for the negative correlation between the uniqueness for Neuroticism and uniqueness for the other Big Five traits. The Correlated Traits Orthogonal Methods (CTOM; Model 5 in Table 2) is nested within the CTCU model, and comparison of these two models reveals a significant difference,  $\chi^2(15, N = 339) = 45.50, p < .001$ . This suggests that, although uniquenesses at each assessment are intercorrelated and the absolute fit for the CTOM model appears adequate, a single method factor does not fully capture the relationship of the method effects within each measurement occasion.<sup>1</sup>

*Exploratory Factor Analyses.* To better understand this pattern of results, we conducted separate *exploratory* factor analyses of the correlation matrix of the latent factors (see Table 3). Following the recommendations of Fabrigar et al. (1999), all exploratory factor analyses were conducted using maximum likelihood when possible with oblique rotation. Maximum likelihood factor analysis produces significance tests that provide one approach to determining the number of factors that are necessary to adequately reproduce the observed correlation matrix.

Analyses of the correlations among the latent factor correlations (Table 3, Panel 2) suggested that two factors were insufficient to account for this pattern of correlations,  $\chi^2(1, N = 339) = 5.99, p < .02$ , RMSEA = .12. Although the two-factor model did not fit well, the general pattern of the two-factor solution corresponded generally with Digman's (1997) results. Consequently, we extracted three moderately to strongly intercorrelated factors, using Principal Axis Factoring, whose loading matrix approached simple structure. Agreeableness, Conscientiousness, and Openness all loaded on the first factor and Extraversion and Neuroticism were each associated with separate factors. Thus, a second-order factor, comprised of three, broad, intercorrelated factors with Agreeableness, Conscientiousness, and Openness as indicators of a single higher-order factor

1. Separate maximum likelihood exploratory factor analyses of the uniqueness correlations provided further evidence that a one factor model did not fit the data for any of the measurement occasions,  $\chi^2(5) = 16.21, 29.97, \text{ and } 39.29, p < .01$ , all RMSEAs > .10, respectively for assessments 1, 2, and 3.

accounted for most of the correlations among the self-reported Big Five factors across time.

## DISCUSSION

Short term self-assessments of the Big Five demonstrate substantial correlations between the same traits over measurement occasions (convergent validity). With respect to the degree of discriminant validity, self-reports of the Big Five demonstrate moderate to substantial correlations between the Big Five factors. It is clear that self-assessments of the Big Five do not yield orthogonal, independent dimensions. The pattern of relationships among the Big Five for these stable self-assessments appear to closely match the general pattern of relationships found in previous research using single assessments (Digman, 1997). This pattern of results also mirrors that of both previous confirmatory factor analytic MTMM research (Borkenau & Ostendorf, 1990; Benet-Martinez & John, 1998; John and Srivastava, 1999) as well as interbattery factor analytic research using college student samples (Finch et al., 1999), all of which were based on different inventories of the Big Five.

One potential limitation of our findings is that they are limited to very short 1-week lags between adjacent assessments of the Big Five traits. A general concern may be raised – if the lag between assessments were longer, the results would have been different in important ways. As a means of exploring the effect of a longer delay between measurements, we artificially increased the variance associated with each measured variable while keeping the covariance among assessments unchanged. Our procedure had the effect of reducing the correlations among measures by approximately 50%. This strategy is consistent with data and models that indicate that correlations between measurements of the same trait decrease as the time between measurements increases (e.g., Conley, 1984; Costa & McCrae, 1988b). When we re-ran the Correlated Traits, Correlated Uniqueness model on the modified data set, we found that no appreciable changes occurred in central results. The only change in the results was an increase in the magnitude of the errors of measurement. Given that within reporter (self-report) analyses failed to yield orthogonal Big Five factors, we now examine the degree of convergent and discriminant validity when different knowledgeable informants provide the same data.



## Phase 2

### METHOD

#### Participants

A subset of participants in Phase 1 provided consent for obtaining a parental rating via mail and brought an acquaintance into the laboratory in exchange for additional credit toward fulfillment of their course requirements. The analyses in Phase 2 are based on all participants ( $n = 256$ ) for whom a complete set of self-, peer, and parent reports were available.<sup>2</sup>

#### Materials

Peers and parents rated the participant on the same 97 unipolar trait adjectives described in Phase 1 using the same 9-point rating scale. Peers and parents received Goldberg's (1992) standard rating instructions with the participant's name embedded within the instructions.

#### Design and Procedure

Each peer completed the self-report inventory in a lecture hall reserved for that purpose and was separated from the participant he or she rated. Questionnaires were mailed to the parent designated by participants. Peer and parent questionnaires were prefaced with an explanatory cover letter. Note that each peer and parent provided only one assessment of the participant; the three self-assessments were aggregated into a single composite for these analyses to provide a trait-like assessment that was more comparable to the peer and parent ratings.<sup>3</sup>

## RESULTS

### Traditional MTMM Analysis: Informant Perspective as a Method Factor

The MTMM correlation matrix between the Big Five traits within and between self-, peer, and parental reports are presented in Table 4. All three self-assessments were included in the present analysis and averaged across the three assessments to produce a single, more

2. Note that the results of Phase 1 do not change materially when the sample is reduced to those with a peer and parental rating.

3. Note that the results do not change appreciably if only the data from the first self-assessment are used in the analyses presented in Phase 2.

stable self-assessment. The average convergent validity coefficient across informant perspective was moderate (same trait, different informants, mean  $r = .30$ ). These coefficients varied from a low of .20, .22, and .18, respectively, for self-peer, self-parent, and peer-parent agreement for Agreeableness to a high of .43, .41, and .43, respectively, for self-peer, self-parent, and peer-parent agreement for Extraversion. Discriminant validity correlations between the Big Five were much lower when examined across different perspectives (different informant, different trait, mean absolute  $r = .05$ ). Finally, the discriminant validity comparison (same informant, different traits) also yields moderate correlations between the Big Five traits (same informant, different trait, mean  $r = .30$ ). Campbell and O'Connell (1982) have pointed out that when there are appreciable method effects, this criterion may be too stringent as the convergent validity coefficients are reduced because of nonshared method effects. Thus, evidence for convergent and discriminant validity exists, although the claims based on these traditional MTMM analyses must be interpreted with some caution in light of the failure to satisfy both forms of discriminant validity (see West & Finch, 1997). The moderate magnitude of the convergent validity coefficients is consistent with other research in which college students are rated on broad personality traits by different types of informants (Funder & West, 1993; John & Srivastava, 1999).

#### **Confirmatory Factor MTMM Analyses: Informant Perspective as a Method Factor**

Table 5 presents the results of the series of multitrait-multimethod confirmatory factor analyses. To maintain a parallel presentation to Phase 1, we again focus on the correlated traits, correlated uniquenesses model (CTCU). In absolute terms, the CTCU model again provides an excellent fit to the data,  $\chi^2(50, N = 256) = 68.41, p = .04$ , CFI = .985, RMSEA = .038, SRMSR = .042. The parameter estimates for the CTCU model (illustrated in Figure 1B) are presented in Table 6. The factor loadings for self-reports are fixed at 1.00 to establish the scale of measurement for the latent variables and to estimate the variances of the latent trait factors.

*Convergent Validity.* In the CTCU model with type of informant as the method factor, the latent Big Five factors represent the convergence across self-, peer, and parental ratings, as opposed to

**Table 4**  
**Examining Convergence among Observer's Perspectives: Multitrait-Multi-Informant Correlation Matrix (N = 256)**

Rating Source	Self-Report					Peer-Report					Parental-Report				
	AG	CO	EX	NE	OP	AG	CO	EX	NE	OP	AG	CO	EX	NE	OP
Self															
AG	(.90)														
CO	.47	(.89)													
EX	.29	.19	(.89)												
NE	-.36	-.18	-.24	(.85)											
OP	.42	.44	.43	-.16	(.84)										
Peer															
AG	.20	.08	.02	-.07	.11	(.93)									
CO	.01	.39	-.11	-.06	-.00	.40	(.89)								
EX	.03	.04	.43	.03	.07	.23	.08	(.89)							
NE	.03	.09	.04	.33	-.02	-.39	-.19	-.16	(.88)						
OP	.03	.07	.13	-.08	.23	.54	.42	.42	-.27	(.85)					
Parent															
AG	.22	.07	.03	-.03	.04	.18	.08	.05	-.05	.16	(.93)				
CO	.04	.35	-.06	.00	-.04	-.05	.27	-.12	.07	-.02	.48	(.92)			
EX	.02	.04	.41	.13	.13	.02	-.03	.43	.01	.11	.20	.07	(.92)		
NE	.02	.06	.06	.26	.03	-.05	-.08	.03	.31	-.11	-.52	-.36	-.05	(.88)	
OP	-.02	-.01	.02	-.03	.28	-.03	-.01	.03	.09	.21	.38	.29	.26	-.22	(.88)
SD	0.90	0.99	1.01	1.02	0.77	1.16	1.10	1.11	1.17	0.90	1.06	1.18	1.20	1.14	0.92
Mean	5.76	5.27	4.91	4.09	5.29	5.91	5.38	5.11	3.91	5.48	6.20	5.67	5.24	3.60	6.01

*Note:* Cronbach's alphas are printed on the main diagonal in parentheses. Correlations between different traits reported by different observers are printed in gray. Validity coefficients—correlations between the same trait as reported by different observers—are represented by the black diagonals within the gray blocks. Self-reports are averaged across the three self-assessments; Reliability is presented only for the first self-assessment.

the stability of self-reports as in Phase 1. Again, two clear findings emerge from examination of the latent factors. First, the factor loadings for peer and parental reports are generally close to 1.00 and significant, indicating that the scale of measurement (and relative importance) of each different type of informant is comparable. Second, the variances of the Big Five latent factors are all moderate and statistically significant, indicating reliable convergence among self-, peer, and parental assessments for each of the Big Five. The magnitude of the variances of the latent traits is smaller in Phase 2

**Table 5**  
**Confirmatory Factor Analysis Model Fit Indices of the**  
**Multitrait-Multi-Informant Matrix ( $N = 256$ )**

<i>Model</i>	$\chi^2$	df	CFI	RMSEA	SRMSR
1. CTNM: Correlated Traits No Method Factors	—	—	—	—	—
2. CTCU: Correlated Traits Correlated Uniquenesses	68.41*	50	.985	.038	.042
3. OTCU: Orthogonal Traits Correlated Uniquenesses	79.04	60	.985	.035	.052
4. CTCM: Correlated Traits Correlated Methods	166.41**	62	.916	.081	.044
5. CTOM: Correlated Traits Orthogonal Methods	172.29**	65	.913	.081	.047
6. OTCM: Orthogonal Traits Correlated Methods	186.83**	72	.907	.079	.058
7. OTOM: Orthogonal Traits Orthogonal Methods	192.98**	75	.905	.079	.060

*Note.* The Correlated Traits No Method Factors (CTNM) model could not be estimated as it did not converge to a proper solution.

CFI: Comparative Fit Index. RMSEA: Root Mean Square Error of Approximation. SRMSR: Standardized Root Mean Square Residual.

\* $p < .05$ . \*\* $p < .01$ .

than in Phase 1, which results in the lower standardized loadings observed in Phase 2. Three observations may be made about this finding : (a) A substantial portion of the variance observed in Phase 1 was variance that is reliable over time, but which is unique to the self-perspective, leading to inflated estimates of true trait variance in Phase 1. (b) The constructs assessed by each type of rater may have been slightly different, but clearly related (Ozer, 1999). (c) The common variance assessed for each trait across raters in Phase 2 may be attenuated by nonshared biases (methods effects) that are not fully captured by the additive MTMM model of methods effects (Campbell & O'Connell, 1982).

*Discriminant Validity.* In contrast to Phase 1, *none* of the Big Five latent factors were significantly correlated. The comparison between Model 2 (CTCU) and Model 3 (OTCU) in Table 5 is an omnibus test of the orthogonality of the Big Five factors *across observer perspectives*. The chi-square difference test provides no evidence that the Big Five factors are correlated,  $\chi^2(10, N = 256) = 10.63, ns$ .<sup>4</sup> Furthermore, Model 5 (CTOM) and Model 4 (CTCM) also did not differ in their fit to the data  $\chi^2(3, N = 256) = 5.88, ns$ , minimizing one potential weakness of the correlated uniqueness model – that it does not allow correlations between method factors.

One possible explanation for the apparent orthogonality of the Big Five factors across informant perspectives is that there is insufficient power to detect relationships. Consequently, we examined the statistical power to detect relationships among the latent factors using  $\alpha = .05$  (MacCallum, Browne, & Sugawara, 1996). At a general level, if the Big Five were, in reality, all correlated at  $\rho = .30$ , the statistical power to detect that in the present framework was .99. At a more focused level, there are 10 potential intercorrelations among the Big Five. The statistical power to detect a *single* correlation of size  $\rho = .30$  averaged .70 across the 10 possible correlations. The corresponding statistical power for  $\rho = .40$  was .90. In sum, there

4. The OTCU model resulted in a slight increase in the SRMSR fit index, relative to the CTCU model, as well as the CTCM and CTOM models, which clearly did not represent the data as well overall. As Hu and Bentler (1998) note, the SRMSR statistic is sensitive to misspecification of the covariances. The OTCU model makes the strong assumption that all cross-method covariances (e.g., self-report of agreeableness with peer-report of conscientiousness) are exactly zero. The average observed cross-method absolute correlation of .05 is thus reflected in the SRMSR index for this model.

**Table 6**  
 Examining Observer as a Method Factor: Correlated Traits  
 Correlated Uniquenesses (CTCU) Parameter Estimates of the  
 Multitrait-Multi-Informant Matrix

Parameter	Latent Trait				
	AG	CO	EX	NE	OP
Factor Loadings					
Self-Report (fixed)	1.00	1.00	1.00	1.00	1.00
<i>standardized</i>	.50	.68	.65	.58	.52
Peer-Report	0.90*	0.89*	1.02*	1.12*	0.72*
<i>standardized</i>	.35	.54	.61	.58	.32
Parent-Report	1.03*	0.89*	1.19*	0.89*	1.36*
<i>standardized</i>	.43	.52	.65	.47	.60
Latent Factor					
Agreeableness	1.00				
Conscientiousness	.11	1.00			
Extraversion	.02	-.10	1.00		
Neuroticism	-.04	.11	.18	1.00	
Openness	.08	-.03	.16	-.03	1.00
Variance	.20*	.45*	.42*	.36*	.16*
Uniqueness					
Self-Report					
Agreeableness	1.00				
Conscientiousness	.65*	1.00			
Extraversion	.40*	.37*	1.00		
Neuroticism	-.52*	-.39*	-.51*	1.00	
Openness	.54*	.74*	.54*	-.20*	1.00
Variance	.59*	.53*	.60*	.72*	.44*
Peer-Report					
Agreeableness	1.00				
Conscientiousness	.51*	1.00			
Extraversion	.31*	.22*	1.00		
Neuroticism	-.47*	-.28*	-.31*	1.00	
Openness	.58*	.53*	.50*	-.34*	1.00
Variance	1.18*	0.85*	0.74*	0.88*	0.73*
Parent-Report					
Agreeableness	1.00				
Conscientiousness	.58*	1.00			
Extraversion	.27*	.17*	1.00		
Neuroticism	-.64*	-.49*	-.16*	1.00	
Openness	.50*	.45*	.32*	-.31*	1.00
Variance	0.94*	1.01*	0.84*	1.02*	0.54*

\* $p < .05$ .

was substantial statistical power within the present context to detect moderate relationships between the Big Five factors across observer perspectives.

*Uniquenesses.* The uniquenesses in the CTCU model represent a combination of pure measurement error variance and a reliable variance component of each trait that is informant type-specific. The variance of each uniqueness is significant and each type of informant's unique perspective was moderately to strongly related across the Big Five. In other words, informant type-specific effects were consistently related within each of the informant types across the Big Five. The relatively high correlations that occur between different Big Five traits based on the reports of the same type of informant relative to those based on the reports of different informant types in Table 4 appear to be due to shared method effects. Note that these correlations are not observed between different Big Five factors in the absence of a common informant type (see Table 6, Panel 2).

In Phase 1, where measurement occasions were examined as a method factor, three unidimensional latent method factors corresponding to each measurement occasion did not adequately account for the pattern of the relationships of the uniquenesses across the Big Five. A similar finding emerged when using a method factor to represent each informant type. A comparison of the Correlated Traits Orthogonal Methods (CTOM; Model 5 in Table 5) and the CTCU model reveals a significant difference,  $\chi^2(15, N = 256) = 103.88, p < .0001$ . Although uniquenesses for each type of informant are intercorrelated, a single method factor does not capture these relationships and suggests that method effects associated with the perspective of each type of informant are multidimensional.

*Exploratory Factor Analyses.* To better understand this pattern of results, we conducted separate *exploratory* factor analyses of their correlation matrices of the uniquenesses for each informant (see Table 6, Panel 3) following the approach outlined in Phase 1. All exploratory factor analyses were conducted using maximum likelihood when possible with oblique rotation.

For both self- and peer reports, two factors were not sufficient to capture the structure of their uniquenesses, both  $\chi^2(1, N = 256) > 11, p < .001$ , although the structure of the two-factor solution generally conformed to Digman's (1997) results. We consequently extracted three moderately to strongly intercorrelated factors using Principal Axis Factoring whose loading matrix approached simple

structure. Agreeableness, Conscientiousness, and Openness all loaded on the first factor and Extraversion and Neuroticism were each associated with separate factors. These results replicate the structure obtained examining the latent factor structure of self-reports across time in Phase 1.

In contrast, the structure of uniquenesses for parental reports was well fit by a 2-factor model,  $\chi^2(1, N = 256) = 1.68, ns$ . Agreeableness, Conscientiousness, and Neuroticism loaded on one factor and Extraversion and Openness loaded on the second, highly related, factor. Note that this structure replicates Digman's (1997) two broad factor solution.

### GENERAL DISCUSSION

The present study both replicates previous findings and provides new insights into the convergent and discriminant validity of the Big Five. Consistent with substantial previous research, self-reports were temporally stable and correlated with reports from other knowledgeable informants (e.g., Costa & McCrae, 1988a). The Big Five also demonstrated substantial discriminant validity, replicating evidence reviewed by John and Srivastava (1999). However, the present study also clearly showed that the *degree of discriminant validity* among the Big Five traits is determined by the lens through which it is measured.

Phase 1 examined the relationship between the Big Five traits in the context of a multitrait-multioccasion matrix. This approach helps minimize potential biases that are associated with a single measurement occasion (e.g., the possibility that the reports of each of the traits are affected by the participant's current level on a temporary state such as mood or fatigue). When self-reports are examined over time, the pattern of relationships among the Big Five that emerges corresponds closely to those from a single assessment that have been found repeatedly in past research (see Digman, 1997; John & Srivastava, 1999). This general pattern of relationships corresponds well to those that emerge from examining self-reports across different assessment inventories based on college and adult samples (see John & Srivastava, 1999). Taken together, these findings make it quite clear that the degree of discriminant validity of the Big Five traits is far from absolute as long as the Big Five are measured using



*self-report* inventories. This pattern of moderate relationships between the Big Five traits gives rise to the possibility that one or more second-order factors could provide a more parsimonious account of personality structure (e.g., Digman, 1997). Indeed, our second-order exploratory factor analysis of the correlations among the latent factors provided results consistent with such a conception.

Phase 2 examined the relationship between the Big Five traits in the context of a multitrait-multi-informant type matrix. This approach helps minimize potential biases that are associated with the use of self-reports in Phase 1. For example, stable individual differences in self-presentational and self-deceptive responding (Paulhus & John, 1998) or in the evaluative components associated with trait terms (Peabody, 1967; Saucier, 1994) would be expected to lead to bias in the positivity of self-reports, resulting in spurious positive relationships between the Big Five traits. Other studies have found similar apparent relationships between the five factors when only ratings of a single knowledgeable informant were used (e.g., teacher ratings, Graziano & Ward, 1992). Biases in responding (e.g., halo effects) across the Big Five traits would also be expected to lead to spurious positive relationships between the five factors. In contrast, the use of different types of informants in Phase 2 minimized the extent to which biases would be shared across different types of informants. When we examined the degree of discriminant validity between the Big Five traits across different informant types, the Big Five traits were *not* significantly related. This contrasting pattern of relationships in the two phases of the study cannot be explained by the higher correlations between measures in Phase 1 (due to shared bias over measurement occasions). When random error was added to reduce the magnitude of the relationships to be comparable to those in Phase 2, the pattern of results did not change.

The pattern of results obtained in Phase 2 is consistent with theorizing by Costa and McCrae (1995) and Goldberg (1993b) that the Big Five in adult populations are best conceptualized as five broad independent traits. It also provides no evidence to support conceptions that propose higher-order factors above the Big Five. Further, the specific results in Phase 2 permit an unusually strong inference. Model 1 (CTNM), which specified no method factors, clearly did not fit the data as it did not even converge. Models 2 (CTCU) and 3 (OTCU), the correlated uniqueness models, did not differ in their fit to the data, supporting the more parsimonious orthogonal traits,

correlated uniquenesses model. Models 5 (CTOM) and 4 (CTCM) did not differ in their fits, indicating that correlations between methods factors were not necessary to account for the present data. Note that this finding minimizes the importance of one of the major limitations of the CTCU model, namely that it implicitly assumes that methods are orthogonal (Marsh, 1989a, 1989b). In addition, inspection of the correlated uniqueness matrix ruled out *unexpected* significant negative correlations between uniquenesses,<sup>5</sup> a condition under which the CTCU can have a spuriously good fit to the data (Millsap, 1990).

Finally, the CTCU Model (Model 2) has a superior fit to the data relative to the CTOM (Model 5) suggesting that the method effects could not be reduced to a single dimension for each type of informant. Indeed, exploratory factor analyses of the correlated uniqueness matrix within informant type found multiple dimensions were necessary to account for these relationships. Interestingly, a two-factor solution conformed generally with Digman's (1997) results, although three factors were necessary to account for the pattern of relationships for self- and peer reports. These results strongly suggest that observed correlations among Big Five traits are the product of informant-specific effects. Given that the same general pattern of relationships emerged from within each informant suggests that future research aimed specifically at disentangling the contributions of different potential processes is necessary to understand these observed relationships among the five factors within a single informant's report.

The strong results in the present study relative to those in many other confirmatory factor analyses of MTMM matrices is noteworthy. Kenny and Kashy (1992) provided evidence that many of the published confirmatory factor analyses of MTMM matrices in the literature up to that time had produced improper solutions. However, as Marsh and Grayson (1995) noted, the prototypical MTMM study reviewed by Kenny and Kashy involved only three traits and three methods with a sample size of about 125. Based on extensive analyses of MTMM matrices by Marsh and Bailey (1991), Marsh and Grayson recommended that a minimum of four traits, three methods, and a sample size of 250 be used if stable solutions were to

5. We obtained negative correlations between the uniquenesses for Neuroticism and the other four Big Five traits. This result was expected, given that Neuroticism was scored in the opposite direction from the other traits.

be attained. With five traits, three methods, and sample sizes of 339 in Phase 1 and 256 in Phase 2, the present studies exceeded the minimal guidelines and encountered no problems in the estimation of these models. In addition, the use of more than three traits permitted comparison of the CTCU and the CTOM models, thereby permitting a formal test of the unidimensionality of the methods factors (Millsap, 1995). The present results illustrate the statistical and conceptual advantages of adhering to the Marsh and Grayson guidelines.

The results of the present studies also have implications for the controversy between the three-factor and five-factor models of personality in adult populations. In these models, the dimensions of Extraversion and Neuroticism (versus Emotional Stability) are similar across formulations of both models (e.g., Clark & Watson, 1999). In addition, the three-factor models do not include Openness to Experience as a broad factor. Consequently, the core of disagreement between the models has focused on understanding the nature of Agreeableness and Conscientiousness (Clark & Watson, 1999; Eysenck, 1992, 1994; Goldberg, 1993b; Goldberg & Rosolack, 1994).

Central to this disagreement are two seemingly well-documented observations: (a) assessments of Agreeableness and Conscientiousness are highly related and (b) measures of the third factor from different versions of the Big Three model (e.g., Psychoticism, Disinhibition, and Norm-Favoring; see Clark & Watson, 1999) are strongly related to both broad traits. With these two observations as premises, disagreement concerning the nature of Agreeableness and Conscientiousness is understandable and seemingly intractable. However, the present results, based on the structure that is common to multiple informant types supports the interpretation of the five factors as orthogonal, independent broad traits. Agreeableness and Conscientiousness are clearly not related in the present study across the perspectives of different types of informants, suggesting that theoretical frameworks that integrate these traits as facets of a broader construct may need to be reexamined. Of course, the findings in the present study are presently limited to a college-student target population, peer and parent informants, and Goldberg's (1992) unipolar trait adjectives as the measure of the Big Five. Further research, particularly with older populations (e.g., Costa & McCrae, 1992), in which personality achieves greater stability and with other measures of the Big Five, are needed to establish the limits of the present findings.

In conclusion, attempts to understand the dimensions along which people differ from each other have focused on the pattern of apparent correlations between the Big Five traits. Assessments of the Big Five are related within self-reports, within parent reports, and within peer reports. Moreover, confirmatory factor analyses of MTMM matrices strongly support the conclusion that the Big Five traits are related within the perspective provided by self-reports collected on multiple occasions. In contrast, confirmatory factor analyses of MTMM matrices, based on the reports of different types of informants, support the conclusion that the Big Five traits are not meaningfully related. These results suggest that within-informant-type influences (e.g., self-presentation; halo effects) may be largely responsible for the correlations observed between the Big Five traits.

## REFERENCES

- Backetman, G., & Magnusson, D. (1981). Longitudinal stability of personality characteristics. *Journal of Personality*, *49*, 148–160.
- Bandura, A. (1999). Social cognitive theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 154–196). New York: Guilford Press.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel psychology*, *44*, 1–26.
- Benet-Martínez, V., & John, O. P. (1998). *Los Cinco Grandes* across cultures and ethnic groups: Multitrait multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology*, *75*, 729–750.
- Bentler, P. M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.
- Biesanz, J. C., & West, S. G. (2000). Personality coherence: Moderating self-other profile agreement and profile consensus. *Journal of Personality and Social Psychology*, *79*, 425–437.
- Biesanz, J. C., West, S. G., & Graziano, W. G. (1998). Moderators of self-other agreement: Reconsidering temporal stability in personality. *Journal of Personality and Social Psychology*, *75*, 467–477.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, *117*, 187–215.
- Borkenau, P., & Ostendorf, F. (1990). Comparing exploratory and confirmatory factor analysis: A study on the 5-factor model of personality. *Personality and Individual Differences*, *11*, 515–524.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.
- Campbell, D. T., & O'Connell, E. J. (1982). Methods as diluting trait relationships rather than adding irrelevant systematic variance. In D. Brinberg &

- L. Kidder (Eds.), *Forms of validity in research* (pp. 93–111). San Francisco: Jossey-Bass.
- Clark, L. A., & Watson, D. (1999). Temperament: A new paradigm for trait psychology. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 399–423). New York: Guilford Press.
- Conley, J. J. (1985). Longitudinal stability of personality traits: A multitrait-multimethod-multioccasion analysis. *Journal of Personality and Social Psychology*, **49**, 1266–1282.
- Conley, J. J. (1984). Longitudinal consistency of adult personality: Self-reported psychological characteristics across 45 years. *Journal of Personality and Social Psychology*, **47**, 1325–1333.
- Costa, P. T., Jr., & McCrae, R. R. (1988a). From catalog to classification: Murray's needs and the five-factor model. *Journal of Personality and Social Psychology*, **55**, 258–265.
- Costa, P. T., Jr., & McCrae, R. R. (1988b). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology*, **54**, 853–863.
- Costa, P. T., Jr., & McCrae, R. R. (1992a). *The NEO-PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., Jr., & McCrae, R. R. (1992b). Reply to Eysenck. *Personality and Individual Differences*, **13**, 861–865.
- Costa, P. T., Jr., & McCrae, R. R. (1994). Set like plaster? Evidence for the stability of adult personality. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change?* (pp. 21–40). Washington, DC: American Psychological Association.
- Costa, P. T., Jr., & McCrae, R. R. (1995). Solid grounds in the wetlands of personality: A reply to Block. *Psychological Bulletin*, **117**, 216–220.
- Digman, J. M. (1996). The curious history of the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality* (pp. 1–20). New York: Guilford Press.
- Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, **73**, 1246–1256.
- Epstein, S. (1983). Aggregation and beyond: Some basic issues of the prediction of behavior. *Journal of Personality*, **51**, 360–392.
- Eysenck, H. J. (1992). Four ways five factors are *not* basic. *Personality and Individual Differences*, **6**, 667–673.
- Eysenck, H. J. (1994). The Big Five or Giant Three: Criteria for a paradigm. In C. F. Halverson Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 37–51). Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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.
- Finch, J. F., Panter, A. T., & Caskie, G. I. L. (1999). Two approaches for identifying shared personality dimensions across methods. *Journal of Personality*, **67**, 407–438.

- Finch, J. F., & West, S. G. (1997). The investigation of personality structure: Statistical models. *Journal of Research in Personality*, **31**, 439–485.
- Fiske, D. W. (1949). Consistency of the factorial structure of personality ratings from different sources. *Journal of Abnormal and Social Psychology*, **44**, 329–344.
- Friedman, H. S., Tucker, J. S., Schwartz, J. E., Martin, L. R., Tomlinson-Keasey, C., Wingard, D. L., & Criqui, M. H. (1995). Childhood conscientiousness and longevity: Health behaviors and the cause of death. *Journal of Personality and Social Psychology*, **68**, 696–703.
- Funder, D. C., & West, S. G. (Eds.) (1993, December). Viewpoints on Personality Consensus, self-other agreement and accuracy in judgements of personality. *Journal of Personality*, **61**, 4.
- Goldberg, L. R. (1990). An alternative “description of personality”: The Big-Five factor structure. *Journal of Personality and Social Psychology*, **59**, 1216–1229.
- Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment*, **4**, 26–42.
- Goldberg, L. R. (1993a). The structure of personality traits: Vertical and horizontal aspects. In D. C. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds.), *Studying lives through time: Personality and development* (pp. 169–188). Washington, DC: American Psychological Association.
- Goldberg, L. R. (1993b). The structure of phenotypic personality traits. *American Psychologist*, **48**, 26–34.
- Goldberg, L. R., & Rosolack, T. K. (1994). The Big Five factor structure as an integrative framework: An empirical comparison with Eysenck’s P-E-N model. In C. F. Halverson Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 7–35). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Graziano, W. G., Jensen-Campbell, L. A., Steele, R. G., & Hair E. C. (1998). Unknown words in self-reported personality: Lethargic and provincial in Texas. *Personality and Social Psychology Bulletin*, **24**, 893–905.
- Graziano, W. G., & Ward, D. (1992). Probing the Big Five in adolescence: Personality and adjustment during a developmental transition. *Journal of Personality*, **60**, 425–439.
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, **3**, 424–453.
- John, O. P. (1990). The “Big Five” factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford Press.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). New York: Guilford Press.
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User’s Reference Guide*. Chicago: Scientific Software International.
- Kenny, D. A. (1994). *Interpersonal Perception*. New York: Guilford Press.

- Kenny, D. A., & Kashy, D. A. (1992). The analysis of the multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin*, **112**, 165–172.
- Kenrick, D. T., & Funder, D. C. (1988). Profiting from controversy: Lessons from the person-situation debate. *American Psychologist*, **43**, 23–34.
- McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality*, **63**, 365–396.
- McCrae, R. R., & Costa, P. T., Jr. (1985). Updating Norman's "adequate taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, **49**, 710–721.
- McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, **52**, 81–90.
- McCrae, R. R., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 139–196). New York: Guilford Press.
- MacCallum, R. C., Browne, M. W., & Sugawara, H. W. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, **1**, 130–149.
- Marsh, H. W. (1989a). Multitrait-multimethod analysis. In J. P. Keeves (Ed.), *Educational research methodology, measurement and evaluation: An international handbook* (pp. 570–580). Oxford: Pergamon.
- Marsh, H. W. (1989b). Confirmatory factor analyses of multitrait-multimethod data: Many problems and a few solutions. *Applied Psychological Measurement*, **13**, 334–361.
- Marsh, H. W., & Bailey, M. (1991). Confirmatory factor analysis of multitrait-multimethod data: A comparison of alternative models. *Applied Psychological Measurement*, **15**, 47–70.
- Marsh, H. W., & Grayson, D. (1995). Latent variable models of multitrait-multimethod data. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 177–198). Thousand Oaks, CA: Sage.
- Millsap, R. E. (1990). A cautionary note on the detection of method variance in multitrait-multimethod data. *Journal of Applied Psychology*, **75**, 350–353.
- Millsap, R. E. (1995). The statistical analysis of method effects in multitrait-multimethod data: A review. In P. E. Shrout & S. T. Fiske (Eds.), *Personality method, methods, and theory: A festschrift honoring Donald W. Fiske* (pp. 93–109). Hillsdale, NJ: Erlbaum.
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, **66**, 574–583.
- Ozer, D. J. (1999). Four principles of personality assessment. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 671–686). New York: Guilford Press.
- Panter, A. T., Tanaka, J. S., & Hoyle, R. H. (1994). Structural models for multimode designs in personality and temperament research. In C. F. Halverson Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 111–138). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, **66**, 1025–1060.
- Peabody, D. (1967). Trait inferences: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology Monograph*, **7** (4, Whole No. 644).
- Reichardt, C. S., & Coleman, S. C. (1995). The criteria for convergent and discriminant validity in a multitrait-multimethod matrix. *Multivariate Behavioral Research*, **30**, 513–538.
- 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.
- Saucier, G. (1994). Separating description and evaluation in the structure of personality attributes. *Journal of Personality and Social Psychology*, **66**, 141–154.
- Saucier, G., & Goldberg, L. R. (1996). The language of personality: Perspectives on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality* (pp. 21–50). New York: Guilford Press.
- Sechrest, L. (1986). Modes and methods of personality research. *Journal of Personality*, **54**, 318–331.
- Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality*, **13**, 389–408.
- Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality*, **60**, 225–251. (Original work published 1961).
- West, S. G., & Finch, J. F. (1997). Personality measurement: Reliability and validity. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of Personality Psychology* (pp. 143–164). San Diego, CA: Academic.
- Wiggins, J. S., & Trapnell, P. D. (1996). A dyadic-interactional perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality* (pp. 88–162). New York: Guilford Press.
- Wiggins, J. S., & Trapnell, P. D. (1997). Personality structure: The return of the big five. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 737–766). San Diego, CA: Academic Press.