The Reliability and Validity of Big Five Inventory Scores With African American College Students

Frank C. Worrell and William E. Cross Jr.

The reliability and validity of scores on the Big Five Inventory (BFI; O. P. John, E. M. Donahue, & R. L. Kentle, 1991) were examined in a sample of 336 African American college students. Results indicated moderate reliability and structural validity for BFI scores. Additionally, BFI subscales had few meaningful correlations with self-esteem and social desirability.

La fiabilidad y la validez de los resultados del Inventario "Big Five" (BFI; O. P. John, E. M. Donahue, y R. L. Kentle, 1991) se examinan en una muestra de 336 estudiantes universitarios afro-americanos. Los resultados indican una fiabilidad moderada y una validez estructural para los resultados del BFI. Adicionalmente, las subescalas del BFI tuvieron pocas correlaciones importantes con el auto-estima y el deseo social.

The Big Five personality factors have been accepted widely in the litera-ture on personality for a number of years (John, 1989, 1990), and many researchers have argued that no assessment of personality is complete without measuring these five basic factors (Aguilar, Kaiser, Murray, & Ozer, 1998). The five-factor model (FFM) of personality has had an interesting his-tory in the research literature. When the Big Five were first reported by Tupes and Christal (1961/1992) and replicated by Norman (1963), they were ignored by many personality researchers, in part because there were too many other personality traits competing to be designated as the basic personality ones (McCrae & John, 1992). However, since the reemergence of the FFM in the 1980s, evi-dence for the model has been strong and convincing, and McCrae and John argued, on the basis of current evidence in the literature, that researchers should accept the FFM as an accurate depiction of the personality traits.

Each of the five factors can be identified by more than one name, but they are most commonly referred to as: Extraversion, Agreeableness, Conscien-tiousness, Neuroticism, and Openness. John and Srivastava (1999, p. 121) de-scribed Extraversion as follows: This factor "implies an *energetic approach* to the social and material world and includes traits such as sociability, activity,

Frank C. Worrell, Cognition and Development, University of California, Berkeley; William E. Cross Jr., Department of Psychology, City University of New York. Correspondence concerning this article should be addressed to Frank C. Worrell, Cognition and Development, 4511 Tolman Hall, University of California, Berkeley, CA 94720-1670 (e-mail: frankc@uclink.berkeley.edu).

assertiveness, and positive emotionality." In spite of the debate about where Extraversion falls on the interpersonal circumplex, an emerging consensus suggests that the factor has a relatively broad content (McCrae & John, 1992). Agreeableness, the second factor named, involves characteristics related to the prosocial and caring side of humanity, such as altruism, affection, and nurturance.

Conscientiousness carries all of the weight of that word in English and involves characteristics related to behavior that is task- or goal-directed (John & Srivastava, 1999), such as impulse control, organization, and delay of gratification. Neuroticism is the only Big Five factor associated with nondesirable behaviors. This factor is related to anxiety and negative emotions. McCrae and John (1992) suggested that there is more definitional consensus about Neuroticism than there is about any of the other factors. On the other hand, Openness is the factor about which there has been the most controversy (McCrae & John, 1992). In natural language studies, the Openness factor consists of words such as *intelligent* and *perceptive*, whereas questionnaire studies have used these same descriptors as well as descriptors related to unconventionality, sensitivity to aesthetics, and the need for variety (McCrae & John, 1992). Raters tend to slant Openness toward the intellect, whereas questionnaire studies include a lot more than an intellectual component. Readers can obtain a more comprehensive review of the FFM by referring to the work of John and Srivastava and of McCrae and John.

The high level of interest in the FFM has led to the development of a number of instruments measuring the Big Five. Some of the more prominent of these include the NEO Personality Inventories (the revised NEO–PI [Personality Inventory] and the NEO–FFI [Five-Factor Index]; Costa & McCrae, 1992), the Big Five Inventory (BFI; John, Donahue, & Kentle, 1991), and the 100-item Trait Descriptive Adjectives (TDA; Goldberg, 1992). Although proponents of the model concede that "the five factors do not exhaust the description of personality," they do believe that the "five factors . . . represent the highest hierarchical level of trait description" (McCrae & John, 1992, p. 190). In keeping with Cross and Markus's (1999) claim that any comprehensive theory of personality must be applicable across cultures, many validation studies have been conducted on the FFM. The structural validity of scores on instruments measuring the FFM have been examined in the United States (McCrae & Costa, 1990); studies have also been conducted to assess the structural validity of these instruments when translated into other languages, including Japanese (Bond, Nakazato, & Shiraishi, 1975), Chinese (Yang & Bond, 1990), and Spanish (Benet-Martinez & John, 1998).

However, despite its importance to personality research and the potential utility of these factors in counseling settings (McCrae & Costa, 1991), few Big Five studies have used African American samples. Aguilar et al. (1998) identified only two studies that examined the FFM in African Americans (Robins, John, & Caspi, 1994; Robins, John, Caspi, & Moffitt, 1996), and a recent PsycINFO search on the five-factor personality model and African Americans or Blacks identified six additional studies (Chan, 1997; Collins & Gleaves, 1998;

Goldberg, Sweeney, Merenda, & Hughes, 1998; Heuchert, Parker, Stumpf, & Myburgh, 2000; Vaughn, 1998; Victor, Dent, Carter, Halverson, & Havill, 1998).

The nine studies of African American and other Black subsamples examined a range of topics. The two earliest studies (Robins et al., 1994; Robins et al., 1996) examined personality in boys during early adolescence. Using the California Child Q-Set (Caspi et al., 1992) to measure the FFM, these researchers identified three personality types with distinct personality profiles in both the African American and Caucasian participants: a well-adjusted type; an introverted, highly agreeable type; and an extraverted, disagreeable type. One other study (Victor et al., 1998) focused on children by interviewing African American parents about their children. These researchers were interested in identifying developmental precursors to the Big Five factors.

Three of the studies used college students as participants. In a study with undergraduate students, Aguilar et al. (1998) demonstrated that reliable measures of the Big Five factors could be created from the Adjective Q-sort, and they replicated the study with 10th graders. Chan (1997), using the NEO–FFI, found that there were no perceived predictive validity differences between Caucasian and African American undergraduates, although Black students did perceive a cognitive test as less valid than did White students. In the third study, Heuchert et al. (2000) examined the structural validity of the NEO–PI–R in a sample of White, Black, Indian, and mixed-race college students in South Africa. They found that the factor structure was similar across races but that Whites obtained high scores on Openness to Feelings, Blacks obtained low scores, and Indians scored in the intermediate range, between Whites and Blacks.

The final three studies of the FFM in African American samples focused on adult populations. Vaughn (1998) administered the NEO–PI–R to 76 HIV-infected women and found that Agreeableness was a significant predictor of health outcome in its own right, with Conscientiousness being a predictor of health outcome when other personality variables were included in the analyses. Collins and Gleaves (1998) found that the FFM, as measured by the Bipolar Adjective Checklist, was moderately supported in both African American and Caucasian job applicants, but they also reported that the factors were highly intercorrelated. Finally, Goldberg et al. (1998) examined demographic differences (i.e., gender, age, education, and ethnic/racial status) on personality factors using the Adjective Checklist. These researchers reported average correlations (.08) between the Big Five and the demographic variables, indicating that the FFM does not seem to differ across demographic groups.

None of the studies with African American samples used the BFI, although the BFI is shorter than many other instruments measuring the FFM and can be completed in a matter of minutes. John and Srivastava (1999) indicated that the BFI was developed specifically as a brief and psychometrically sound measure of the FFM. Moreover, although a substantial amount of evidence has been reported on the structural and criterion-related validity of the BFI (e.g., Benet-

Martinez & John, 1998; John et al., 1991; John & Srivastava, 1999), there have been no studies of the psychometric properties of BFI scores for African American samples. Goodwin and Goodwin (1999, p. 409) pointed out that reliability and validity are properties of the scores on instruments "under certain conditions and with a particular group of participants." The use of the instrument with a number of African American samples indicates that an analysis of the validity and reliability of BFI scores in this population is overdue.

The present study examined the internal consistency, structural validity, and discriminant validity of BFI scores in a sample of African American college students. To examine discriminant validity, we correlated BFI scores with measures of self-esteem, social desirability, and optimism. Social desirability was chosen because this variable is a potential bane for all self-report measures. Self-esteem has traditionally been used as a global, positive, personal identity factor and needs to be ruled out when measuring attitudinal variables. Finally, optimism and other future-oriented variables have been found to be related to a variety of important factors, including school risk status (Worrell & Hale, 2001; Worrell, Latto, & Perlinski, 1999), achievement at school and on the job (Seligman, 1991, 1995), involvement in risky behaviors and delinquency (Oyserman & Markus, 1990a, 1990b; Stein, Roeser, & Markus, 1998), and resilience (Wyman, Cowen, Work, & Kerley, 1993; Wyman et al., 1992). Thus, optimism is a potentially important factor to differentiate from the Big Five factors in the social context of the United States.

We hypothesized that Neuroticism scores would be negatively correlated with social desirability, self-esteem, and optimism and that the other subscales would be positively correlated with these variables. Although no correlations were expected to be substantial (i.e., > .6), we predicted that Neuroticism would have a moderate negative relationship with self-esteem.

method

PARTICIPANTS

Participants were 329 African American students (34% male) attending a predominantly White university in the Northeast. Participants' ages ranged from 17 to 59 years ($M = 20.56$, $SD = 3.61$), and they reported a mean grade point average of 2.76 ($SD = .56$) on a 4.00 scale. Undergraduates made up 93% of the sample, and 90% of the students characterized the communities in which they were raised as either working or middle class. Sixty-six percent of the participants reported that both of their parents had at least a high school diploma.

MEASURES

Participants completed five measures for this study: a demographic sheet, the BFI (John et al., 1991), the Rosenberg Self-Esteem Scale (RSES; Rosenberg,

1965), the Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1984, 1991), and the Measure of Perceived Life Chances (MPLC; Jessor, Donovon, & Costa, 1990). Instruments were counterbalanced in the packets to control for response sets.

BFI. The BFI (John et al., 1991) consists of 44 items that measure the FFM. The five subscales are Extraversion (8 items), Agreeableness (9 items), Conscientiousness (9 items), Neuroticism (8 items), and Openness (10 items). All items consist of short phrases (e.g., is talkative; is depressed, blue; tends to be lazy) that are based on prototypical trait adjectives related to each construct (John & Srivastava, 1999) and are rated on a 5-point scale (1 = *disagree strongly*, 5 = *agree strongly*). Subscale scores are created by reverse scoring specified items, summing the ratings for the items on each subscale, and dividing by the total number of items to obtain a mean score. John and Srivastava reported alpha reliabilities from .75 to .80 for subscales and 3-month test–retest reliabilities from .80 to .90. Validity coefficients with the NEO–FFI (Costa & McCrae, 1992) and the TDA (Goldberg, 1992) that were corrected for attenuation averaged .91 for Extraversion, Agreeableness, and Conscientiousness, .88 for Neuroticism, and .83 for Openness (John & Srivastava, 1999).

RSES. The RSES is one of the most frequently used measures of global self-esteem (Vandiver & Fhagen-Smith, 1998). This scale consists of 10 items, 5 of which are negatively worded to avoid response sets. Responses are made on a 4-point Likert-type scale (1 = *strongly disagree*, 4 = *strongly agree*), with higher scores indicating a more positive self-concept. Because some studies with African Americans have yielded mixed results with regard to the factor structure of the scores, Vandiver and Fhagen-Smith suggested that when using the RSES with African Americans, researchers should examine its factor structure. An exploratory factor analysis supported a single factor in this sample, and the total score was used in these analyses. The reliability of RSES scores in this sample was adequate (see Table 1).

BIDR. The BIDR (Paulhus, 1984, 1991) is a 40-item inventory measuring two aspects of social desirability: Impression Management (IM) and Self-Deceptive Enhancement (SDE). Each subscale consists of 20 items, and items are rated on a 7-point scale (1 = *not true,* 4 = *somewhat true,* and 7 = *very true*). Half of the items are negatively worded and must be reversed scored, with higher scores reflecting greater levels of social desirability. Internal consistency estimates have ranged from .68 to .80 for SDE scores and from .75 to .86 for IM scores in samples of adults and college students (Paulhus, 1991). Factor analyses of the scores have yielded two factors with intercorrelations of .40 or less. As the name suggests, SDE is internally focused and was found to have moderate, positive correlations with the Defense Mechanism Inventory and the Ways of Coping Scale. IM, which is focused on self-presentation, was found to be correlated with the Eysenck's Lie Scale and the Minnesota Multiphasic Inventory's Lie scale (Paulhus, 1991).

TABLE 1

Descriptive Statistics of Major Variables in the Study and Correlation of the Big Five Inventory Subscales

Subscale	M	SD	Skew	Kurtosis	α
Rosenberg Self-Esteem Scale	3.50	0.44	−1.49	4.25	.83
Measure of Perceived Life					
Chances	4.17	0.53	−0.79	2.33	.84
Social Desirability subscales of					
BIDR					
Impression Management	3.34	0.86	0.24	0.04	.60
Self-Deceptive Enhancement	4.36	0.59	−0.07	0.45	.80
Big Five Inventory subscales					
Extraversion (E)	3.01	0.69	−0.17	−0.48	.83
Agreeableness (A)	3.70	0.60	−0.18	−0.37	.70
Conscientiousness (C)	3.69	0.68	−0.29	−0.25	.79
Neuroticism (N)	2.93	0.76	0.10	−0.55	.77
Openness (O)	3.92	0.57	−0.47	−0.33	.72

Correlation of Big Five Inventory Subscales					
	E	A	C	N	O
Extraversion	—	.17	.32	−.31	.35
Agreeableness	.13	—	.34	−.53	.17
Conscientiousness	.26	.25	—	−.42	.23
Neuroticism	−.25	−.39	−.33	—	−.16
Openness	.27	.12	.17	−.12	—

Note. N = 329. BIDR = Balanced Inventory of Desirable Responding. Correlations above the diagonal are corrected for attenuation.

MPLC. The MPLC (Jessor et al., 1990) is a 10-item instrument measuring an individual's belief that life in the future will turn out well. Questions are rated on a 5-point Likert-type scale, with respondents rating the chances of each statement coming to pass (1 = *very high,* 2 = *high,* 3 = *about 50/50,* 4 = *low,* and 5 = *very low*). Items are reversed scored on completion, with higher scores indicating a more positive outlook. The first two questions on the MPLC, which was originally designed for use with a high school sample, asked about the chances of graduating from high school and going to college.

With college populations, the first two questions are reworded to ask about the chances of graduating from college and going to graduate school. Jessor et al. (1990) reported reliability estimates for scores in the .8 range and indicated that MPLC scores were related to healthy behaviors (e.g., not smoking) in adolescents. Worrell et al. (1999) reported that MPLC scores had low to moderate correlations with self-esteem and had reliability estimates in the .8 range. They also found that academically talented students and at-risk students attending their home high school rated themselves significantly higher on the MPLC than at-risk students who had been transferred from their home high school to a continuation school.

PROCEDURE

Students were recruited in several sites on campus, including residence halls, the cultural center, student organizations, and the student union building. A request announcing the study and soliciting participation was also sent out on a listserv for minority students. Instruments were administered as part of a larger packet of measures to 336 African American students. The 329 participants (98%) in this study completed all the measures that were used. Administrators consisted of Black graduate students and faculty, and administration of the packet was standardized with each administrator receiving a list of instructions to follow. Participants were paid $10 on completion of the packet.

results

DESCRIPTIVE STATISTICS

Means, standard deviations, skews, kurtoses, and internal consistency estimates of scores for the major variables, as well as BFI subscale intercorrelations, are presented in Table 1. As can be seen, the distributions were relatively normal except for the RSES and MPLC scores, which had elevated kurtosis. Internal consistency estimates for the BFI subscales ranged from .70 to .83 ($Mdn = .77$), indicating moderate reliability for subscale scores in this sample. Internal consistency estimates for the other variables were also in the moderate range, with one estimate (for IM) falling below .70. Intercorrelations among the summed BFI subscale scores ranged from |.13| to |.38| ($Mdn\ r = |.25|$), indicating that the subscales as described by John et al. (1991) are relatively independent. Corrections for attenuation increased correlations, but only two of them were larger than |.40|.

STRUCTURAL VALIDITY ANALYSES

Because the BFI structure had not been examined with an African American sample prior to this study, exploratory factor analysis was used to examine the factor structure of the scores (Bryant & Yarnold, 1995; Goodwin & Goodwin, 1999; Stevens, 1996; Tabachnick & Fidell, 2001). A principal-axis procedure was used to eliminate the influence of error variance (Thompson & Daniel, 1996). Both varimax and oblique rotations were examined, with the latter being used to examine the level of intercorrelations among the extracted factors. The Bartlett test of sphericity was significant (4,132.28, $p < .001$), and the Kaiser-Meyer-Olkin measure of sampling adequacy was adequate (.81). Multiple methods were used to determine how many factors to extract (Floyd & Widaman, 1995; Thompson & Daniel, 1996). The initial extraction resulted in 11 factors with eigenvalues greater than 1. However, both the scree test (Cattell, 1966) and parallel analysis (Lautenschlager, 1989; Watkins, 2000) suggested six factors,

and the scale is based on a five-factor model. With a wide range of commu-
nalities $(Mdn = .48)$ and an approximately 1:10 variable/factor ratio, a sample
size of 300 was adequate for an acceptable solution (Gorsuch, 1983; MacCallum,
Widaman, Zhang, & Hong, 1999).

A six-factor extraction with varimax rotation produced five viable factors
and one factor with a single nonsalient loading. A five-factor extraction pro-
duced the hypothesized factors with only four cross-loadings above .35, which
was used as the minimum loading for assignment to a factor (Floyd & Widaman,
1995). This solution, presented in Table 2 , accounted for 35% of the variance
in the scores. Factor 1 (Extraversion) consisted of the 8 items assigned to this
factor, with loadings ranging from .50 to .76. Factor 2 (Openness) consisted of
10 items with loadings of |.38| to |.66|. Of these, 7 items were originally as-
signed to Openness, 2 Extraversion items cross-loaded positively, and 1 Neu-
roticism item (#34) loaded negatively. Factor 3 (Conscientiousness) consisted
of the 9 items assigned to this factor with loadings of .40 to .66. Neuroticism
(Factor 4) had 9 items with loadings from |.37| to |.65|. Seven of the items were
assigned to Neuroticism, and 1 item each from Extraversion and Conscien-
tiousness cross-loaded negatively on this factor. Finally, Factor 5 (Agreeable-
ness) consisted of 6 of the 9 items assigned to this factor, with loadings ranging
from .36 to .64.

In sum, four items loaded on two factors and one item loaded on a factor to
which it was not assigned. Two Extraversion items cross-loaded positively on
Openness, and one Extraversion item cross-loaded negatively on the Neuroti-
cism, but all three had higher loadings on Extraversion. One Conscientious-
ness item cross-loaded negatively and with a lower loading on Neuroticism,
and one Neuroticism item loaded inversely on Openness and did not load on
Neuroticism. Additionally, three Openness items and three Agreeableness items
did not load on any factor above .35, although in all cases, they had their
highest loadings on the factor to which they were assigned. Construct reliabil-
ity estimates for BFI factors based on salient loadings were comparable with
the ones calculated using Cronbach's alpha, ranging from .65 to .81 $(Mdn =$
.79). The results produced from an oblique rotation were quite similar, with
differences in the order of factor extraction. Factor intercorrelations from the
oblique rotation ranged from |.04| to |.24|.

DISCRIMINANT VALIDITY

Correlations between the five BFI subscales and the RSES, SDE, IM, and
MPLC scales are presented in Table 3, as are correlations corrected for attenu-
ation. Fifteen of the 20 observed correlations were significant $(p < .002)$, but
only 2 of these accounted for more than 16% of the variance in scores. Addi-
tionally, relationships were in keeping with theory. Only correlations greater
than .30 were accepted as meaningful. Conscientiousness was positively and

TABLE 2

Five-Factor Solution From Principal Axis Varimax Rotation of Big Five Inventory (BFI) Scores (*N* = 329)

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Communality
BFI 21	*.76*	.03	.07	−.10	−.08	.68
BFI 1	*.70*	.06	.04	.13	.08	.53
BFI 36	*.67*	.16	.08	−.10	.22	.55
BFI 31	*.55*	.08	.10	*−.37*	−.13	.52
BFI 11	*.53*	.17	.26	−.11	.25	.54
BFI 26	*.52*	*.38*	.22	−.12	−.00	.59
BFI 16	*.50*	*.38*	.18	.02	.26	.54
BFI 6	*.50*	−.08	−.03	−.09	−.06	.50
BFI 25	.16	*.66*	.08	−.17	−.01	.57
BFI 5	.19	*.62*	.15	−.19	.03	.47
BFI 44	.08	*.58*	−.01	−.07	.00	.54
BFI 40	.08	*.57*	.02	−.04	.15	.44
BFI 20	.07	*.53*	−.05	.15	.02	.48
BFI 30	.10	*.51*	.06	.10	.11	.49
BFI 15	−.04	*.50*	.18	.12	.00	.52
BFI 34	.19	*−.39*	−.18	.31	−.12	.43
BFI 33	.04	.16	*.66*	.02	.26	.57
BFI 3	.07	.05	*.62*	.03	.15	.55
BFI 28	−.02	.13	*.60*	.05	.20	.52
BFI 38	.09	.08	*.58*	−.10	.03	.40
BFI 18	.04	−.06	*.55*	−.19	−.16	.53
BFI 43	−.03	.07	*.51*	*−.38*	−.11	.44
BFI 23	.08	.14	*.50*	−.23	.02	.42
BFI 13	.15	.12	*.50*	.03	.17	.37
BFI 8	.06	−.03	*.40*	−.23	−.05	.34
BFI 14	−.05	.17	−.04	*.65*	−.07	.45
BFI 19	−.14	.05	−.14	*.58*	.03	.46
BFI 9	−.00	−.14	−.07	*.56*	−.16	.55
BFI 39	−.11	−.20	−.15	*.50*	−.00	.41
BFI 4	−.20	−.04	−.17	*.50*	−.24	.46
BFI 24	−.02	−.15	−.08	*.45*	−.28	.51
BFI 29	.00	.04	.10	*.39*	−.27	.34
BFI 32	.08	.17	.04	.04	*.64*	.48
BFI 7	.01	.12	.07	−.00	*.51*	.45
BFI 42	.16	.11	−.00	−.05	*.50*	.30
BFI 12	−.25	−.01	.10	−.22	*.47*	.54
BFI 37	−.13	−.06	.05	−.34	*.41*	.56
BFI 17	.18	.19	−.00	−.16	*.36*	.38
BFI 2	−.08	−.06	.19	−.29	.33	.32
BFI 27	.08	−.13	.04	−.30	.33	.31
BFI 22	.02	−.05	.08	−.07	.30	.20
BFI 10	.21	.33	.03	.09	.23	.38
BFI 41	−.12	.24	.03	−.04	−.01	.22
BFI 35	.06	.12	.05	−.10	−.08	.25
Eigenvalue	6.3	3.0	2.3	2.1	1.6	
% variance	14.3	6.8	5.3	4.7	3.7	
Alpha	.81	.78	.79	.74	.65	

Note. Factor 1 = Extraversion; Factor 2 = Openness; Factor 3 = Conscientiousness; Factor 4 = Neuroticism; Factor 5 = Agreeableness; % variance = percent of variance accounted for. Alpha estimates are construct reliabilities based on salient loadings. Boldface italic values indicate salient loadings.

TABLE 3

**Correlations Between the Big Five Inventory Subscales and Other
Major Variables**

Variable	Self-Esteem		Self-Deceptive Enhancement		Impression Management		Perceived Life Chances	
Extraversion	.31	(.37)	.27	(.33)	.05	(.07)	.25	(.30)
Agreeableness	.20	(.26)	.16	(.21)	.36	(.56)	.17	(.22)
Conscientiousness	.31	(.38)	.45	(.57)	.26	(.38)	.24	(.29)
Neuroticism	−.39	(−.49)	−.47	(−.60)	−.17	(−.25)	−.13	(−.16)
Openness	.13	(.18)	.09	(.12)	.01	(.02)	.15	(.19)

Note. N = 329. All correlations above .15 are significant (*p* < .002). Correlations in parentheses
are corrected for attenuation.

meaningfully correlated with both RSES and SDE. Agreeableness was posi-
tively correlated with all four variables, but only the correlation with IM was
meaningful. As predicted, Neuroticism had negative correlations with all four
variables, including meaningful correlations with RSES and SDE. Extraver-
sion was positively related to RSES, and Openness did not have meaningful
correlations with any of the variables. The MPLC score had its strongest cor-
relations with Conscientiousness and Extraversion but was not meaningfully
correlated to any of the BFI subscales.

discussion

The results reported in this study provide preliminary evidence on the reli-
ability and validity of BFI scores for African Americans. Scores on the BFI
subscales produced moderate reliability estimates based on both Cronbach's
alpha and salient loadings on factors. Additionally, the factor structure reported
by John et al. (1991) was essentially replicated, albeit with lower factor load-
ings. BFI scores were relatively distinct from measures of self-esteem, social
desirability, and optimism, and the pattern of correlations made theoretical
sense. In general, the results suggest that the FFM as measured by the BFI can
be applied to African American college students.

DESCRIPTIVE ANALYSES

An examination of the descriptive statistics indicated that they were consistent
with the sample's characteristics. This college-educated group had high global
self-concept scores as well as a high level of perceived life chances. Means on
the social desirability and personality scales were average with minimal skew
and kurtosis. Intercorrelations among the BFI subscales were also low to
moderate and in keeping with previous studies of this instrument (John &

Srivastava, 1999), and in contrast to the study using the Bipolar Adjective Checklist (Collins & Gleaves, 1998). With the exception of IM, reliability estimates of scores on the scales were in the acceptable range.

FACTOR-ANALYTIC RESULTS

The factor analyses identified the five factors associated with the FFM, with Extraversion emerging first and accounting for the largest proportion of variance as in many studies (John & Srivastava, 1999). In fact, with the exception of Agreeableness and Openness, the factors emerged in the order that they typically do. Openness, which is typically the last factor to be extracted, emerged second where Agreeableness often falls, and Agreeableness was extracted fifth. This difference in the typical order may reflect a larger problem with the Openness and Agreeableness factors in this sample. Three items that typically load on Agreeableness did not have meaningful coefficients on that factor, and the factor that was extracted in this sample's scores is made up of six rather than nine items. Similarly, three Openness items did not have meaningful loadings on the Openness factor. However, Openness was supplemented by three cross-loadings, two of which came from the strongest factor, Extraversion.

These differences strengthened the Openness factor and weakened the Agreeableness factor. Because this was a first test of the BFI structure in this population, it is not possible to interpret these findings. The results of subsequent studies will indicate whether the items that did not attain meaningful factor loadings in this sample are problematic for African Americans or are anomalous.

DISCRIMINANT ANALYSES

The relationships between BFI scores and self-esteem, social desirability, and optimism were generally in the low to moderate range, as expected. Despite this, the pattern of correlations was interesting. Neuroticism had a larger correlation with self-esteem than any of the other factors, sharing 15% of the variance with that variable. These findings suggest that although individuals who are high in neuroticism tend to have a lower general self-concept, the Big Five personality factors contribute little to general self-concept. Conscientiousness, typically associated with control and self-constraint (John & Srivastava, 1999), had a moderate association with SDE, the more internally focused measure of social desirability. Agreeableness, which reflects a "prosocial and communal orientation" (John & Srivastava, 1999, p. 121), had a moderate relationship with IM, the social desirability measure that measures how one attempts to present to others. Neuroticism also had a moderate negative relationship to SDE, but no relationship to IM, suggesting that people who are neurotic are not overlooking their anxieties nor are they interested in misrepresenting themselves to others. An examination of the corrected correlations does not change

the essence of the findings, which, taken together, present a compelling argument for the validity of the various BFI subscale scores.

IMPLICATIONS FOR COUNSELING

Counselors who work with clients on a variety of life and personal issues sometimes wish to measure the FFM to get some idea of the global traits of their clients. BFI scores provide a reliable and valid measure of the Big Five personality factors. Moreover, like the NEO–PI, the instrument can be administered and scored quickly, and its items are not slanted toward pathology (McCrae & Costa, 1991). Although the BFI does not have multiple measures of any Big Five factor, unlike the NEO–PI, there are many instances in which a counselor or clinician will simply want a global view of the client on the FFM. In cases like this, with a Flesch-Kincaid readability index of Grade 4.1 and an administration time of approximately 10–15 minutes, the BFI is probably the fastest instrument to administer and score. Additionally, the pattern of relationships reported in this study provide some hypotheses for a clinician to examine: High scores on Agreeableness and Conscientiousness that do not seem to reflect a client's actual behavior may suggest impression management or self-deception, respectively. Similarly, when a client has an overly high score on Neuroticism, a clinician may wish to follow up with an assessment of global self-esteem or signs of psychopathology. Of course, the utility of BFI scores in practice settings–both clinical and vocational–needs to be examined empirically before definitive statements can be made.

LIMITATIONS

This study had a number of limitations that should be mentioned. First, the sample came from one university setting and may not generalize to other African American samples. A second concern about generalization is related to educational achievement. All of the participants were college students and, thus, represented the educational elite of the African American community. The validity of BFI scores in this group may not generalize to groups of African Americans who are less well educated. Third, although it was not done in this study, it will be important to establish the convergent validity of BFI scores in African American samples with longer measures of the FFM, if we are to be confident in using BFI scores in making decisions in clinical and other settings.

CONCLUSION

This study examined the reliability and validity of BFI scores in a sample of African American university students. The results indicated that the structure of the BFI was substantially replicated, although 6 of the 44 items did not achieve meaningful loadings on any factor. Further, reliability estimates of BFI

scores were in the moderate range and comparable with estimates reported in previous studies. BFI scores were found to be independent and were distinct from measures of self-esteem, social desirability, and optimism. In general, the results of this study support the use of the BFI with African Americans. The study adds to the growing body of evidence in personality psychology that the FFM "is a replicable phenomenon" across groups and cultures that "can be profitably used in applied settings" (McCrae & John, 1992, p. 206). However, this study was the first to look at the psychometric properties of BFI scores with this population. Other studies using a variety of techniques and drawing from other samples of African Americans (e.g., noncollege populations) as well as convergent and predictive validity studies must be conducted to "build a body of evidence" (Benson, 1998, p. 10) that supports the use of BFI scores with this population.

references

Aguilar, M. L., Kaiser, R. T., Murray, C. B., & Ozer, D. J. (1998). Validation of an Adjective Q-sort as a measure of the Big Five personality structure. *Journal of Black Psychology, 24*, 145–163.

Benet-Martinez, 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.

Benson, J. (1998). Developing a strong program of construct validation: A test anxiety example. *Educational Measurement: Issues and Practice, 17*, 10–22.

Bond, M. H., Nakazato, H., & Shiraishi, D. (1975). Universality and distinctiveness in dimensions of Japanese person perception. *Journal of Cross-Cultural Psychology, 6*, 346–357.

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99–136). Washington, DC: American Psychological Association.

Caspi, A., Block, J., Block, J. H., Klopp, B., Lynam, D., Moffitt, T. E., & Stouthamer-Loeber, M. (1992). A "common-language" version of the California Child Q-set for personality assessment. *Psychological Assessment, 4*, 512–523.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245–276.

Chan, D. (1997). Racial subgroup differences in predictive validity perceptions on personality and cognitive ability tests. *Journal of Applied Psychology, 82*, 311–320.

Collins, J. M., & Gleaves, D. H. (1998). Race, job applicants, and the five-factor model of personality: Implications for Black psychology, industrial/organizational psychology, and the five-factor theory. *Journal of Applied Psychology, 83*, 531–534.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.

Cross, S. E., & Markus, H. R. (1999). The cultural constitution of personality. In L. A. Previn & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 378–396). New York: Guilford Press.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*, 286–299.

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

Goldberg, L. R., Sweeney, D., Merenda, P. H., & Hughes, J. E., Jr. (1998). Demographic variables and personality: The effects of gender, age, education, and ethic/racial status on self-descriptions of personality attributes. *Personality and Individual Differences, 24*, 393–403.

Goodwin, L. D., & Goodwin, W. L. (1999). Measurement myths and misconceptions. *School Psychology Quarterly, 14,* 408–427.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Heuchert, J. W. P., Parker, W. D., Stumpf, H., & Myburgh, C. P. H. (2000). The five-factor model of personality in South African college students. *American Behavioral Scientist, 44,* 112–125.

Jessor, R., Donovan, J. E., & Costa, F. (1990). Personality, perceived life chances, and adolescent health behavior. In K. Hurrelmann & F. Lösel (Eds.), *Health hazards in adolescence* (pp. 25–42). New York: Walter de Gruyter.

John, O. P. (1989). Towards a taxonomy of personality descriptors. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 261–271). New York: Springer.

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

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

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Previn & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). New York: Guilford Press.

Lautenschlager, G. J. (1989). A comparison of alternatives to conducting Monte Carlo analyses for determining parallel analysis criteria. *Multivariate Behavioral Research, 24,* 365–395.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4,* 84–99.

McCrae, R. R., & Costa, P. T. (1990). *Personality in adulthood.* New York: Guilford Press.

McCrae, R. R., & Costa, P. T. (1991). The NEO Personality Inventory: Using the five-factor model in counseling. *Journal of Counseling & Development, 69,* 367–372.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60,* 175–215.

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,* 574–683.

Oyserman, D., & Markus, H. R. (1990a). Possible selves and delinquency. *Journal of Personality and Social Psychology, 59,* 112–125.

Oyserman, D., & Markus, H. R. (1990b). Possible selves in balance: Implications for delinquency. *Journal of Social Issues, 46,* 141–157.

Paulhus, D. L. (1984). Two-component model of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598–609.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press.

Robins, R. W., John, O. P., & Caspi, A. (1994). Major dimensions of personality in early adolescence: The Big Five and beyond. In C. F. Halverson, Jr., G. A. Kohnstramm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 267–291). Hillsdale, NJ: Erlbaum.

Robins, R. W., John, O. P., Caspi, A., & Moffitt, T. E. (1996). Resilient, overcontrolled, and undercontrolled boys: Three predictable personality types. *Journal of Personality and Social Psychology, 70,* 157–171.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Seligman, M. E. (1991). *Learned optimism.* New York: Knopf.

Seligman, M. E. (1995). *The optimistic child.* New York: Houghton Mifflin.

Stein, K. F., Roeser, R., & Markus, H. R. (1998). Self-schemas and possible selves as predictors and outcomes of risky behaviors in adolescents. *Nursing Research, 47,* 96–106.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (3rd ed.). Needham Heights, MA: Allyn & Bacon.

Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement, 56,* 197–208.

Tupes, E. C., & Christal, R. E. (1992). Recurrent personality factors based on trait ratings. *Journal of Personality, 60,* 225–251. (Original work published 1961)

Vandiver, B. J., & Fhagen-Smith, P. (1998). *A validation of the Rosenberg and Collective Self-Esteem Scales with African American students.* Unpublished manuscript, The Pennsylvania State University.

Vaughn, V. S. (1998). Personality domains as predictors of health outcome in HIV-infected African American females (Doctoral dissertation, Auburn University, 1998). *Dissertation Abstracts International, 59,* 451.

Victor, J. B., Dent, H. E., Carter, B., Halverson, C. F., Jr., & Havill, V. L. (1998). How African American parents describe their children. In G. A. Kohnstamm & C. F. Halverson, Jr. (Eds.), *Parental descriptions of child personality: Developmental antecedents of the Big Five* (pp. 167–187). Mahwah, NJ: Erlbaum.

Watkins, M. W. (2000). Mac Parallel Analysis [Computer software]. State College, PA: Author.

Worrell, F. C., & Hale, R. L. (2001). The relationship of hope in the future and perceived school climate to school completion. *School Psychology Quarterly, 16,* 370–388.

Worrell, F. C., Latto, I. K., & Perlinski, M. A. (1999). The relationship of risk status to self-esteem and perceived life chances. *The Journal of At-Risk Issues, 5*(2), 33–38.

Wyman, P. A., Cowen, E. L., Work, W. C., & Kerley, J. H. (1993). The role of children's future expectation is self-system functioning and adjustment to life stress: A prospective study of urban at-risk children. *Development and Psychopathology, 5,* 649–661.

Wyman, P. A., Cowen, E. L., Work, W. C., Raoof, A., Gribble, P. A., Parker, G. R., & Wannon, M. (1992). Interviews with children who experience major life stress: Family and child attributes that predict resilient outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 904–910.

Yang, K., & Bond, M. H. (1990). Exploring implicit personality theories with indigenous or imported constructs: The Chinese case. *Journal of Personality and Social Psychology, 58,* 1087–1095.