Journal of Personality and Social Psychology
1998, Vol. 75, No. 3, 729-750

Copyright 1998 by the American Psychological Association, Inc.
0022-3514/98/$3.00

Los Cinco Grandes Across Cultures and Ethnic Groups: Multitrait Multimethod Analyses of the Big Five in Spanish and English

Verónica Benet-Martínez
University of Michigan

Oliver P. John
University of California, Berkeley

Spanish-language measures of the Big Five personality dimensions are needed for research on Hispanic minority populations. Three studies were conducted to evaluate a Spanish version of the Big Five Inventory (BFI) (O. P. John et al., 1991) and explore the generalizability of the Big Five factor structure in Latin cultural groups. In Study 1, a cross-cultural design was used to compare the Spanish and English BFI in college students from Spain and the United States, to assess factor congruence across languages, and to test convergence with indigenous Spanish Big Five markers. In Study 2, a bilingual design was used to compare the Spanish and English BFI in a college-educated sample of bilingual Hispanics and to test convergent and discriminant validity across the two languages as well as with the NEO Five Factor Inventory in both English and Spanish. Study 3 replicated the BFI findings from Study 2 in a working-class Hispanic bilingual sample. Results show that (a) the Spanish BFI may serve as an efficient, reliable, and factorially valid measure of the Big Five for research on Spanish-speaking individuals and (b) there is little evidence for substantial cultural differences in personality structure at the broad level of abstraction represented by the Big Five dimensions.

Hispanics are the fastest growing minority group in the United States, and within 25 years they will become the nation's largest minority group (U.S. Bureau of the Census, 1995). Yet, there is little personality research on this minority population, and few articles dealing with this ethnic group ever appear in the pages of personality journals. Insofar as research on this group is dependent on the availability of instruments, personality psychologists need to develop appropriate and easily accessible measures in Spanish. The three studies reported in this article are designed to help remedy this situation with respect to the Big Five personality dimensions.

Work in cultural psychology has identified a number of general value differences between Latin (e.g., Spanish, Hispanic) and Anglo American cultures (Hofstede, 1983; Marín & Marín, 1991; Schwartz, 1994; Triandis, 1990; Triandis, Lisansky, Marín, & Betancourt, 1984). Compared with Anglo American culture, Latin cultures are less individualistic and more collectivist. That is, they emphasize interdependence and the goals of the in-group; they value *simpatía,* which may be described as the need for interpersonal behaviors that promote smooth and harmonious relationships, such as expressing positive emotions and avoiding interpersonal conflict; they have a flexible time-orientation, being more present than future oriented and less likely to delay gratification; and they value familialism, that is, they are strongly attached to and identified with the family. The many Latin cultures also differ from each other in important ways. For example, individuals of Latin American background (e.g., Hispanics who live in the United States) speak a variant of Spanish that is different from the Castillian spoken by Spaniards living in Spain, and they seem to show the cultural characteristics of collectivism, *simpatía,* present-time orientation, and familialism to a greater extent (Hofstede, 1983; Marín & Marín, 1991).

At this point, little is known about whether and how these cultural differences at the group level translate into differences in the organization (or structure) of personality characteristics at the individual level. One possibility, as Gergen, Gulerce, Lock, and Misra (1996) suggested, is that each culture shapes a unique personality structure, thus making multiple, culturally specific personality psychologies necessary. Alternatively, as McCrae and Costa (1997a) recently suggested, certain basic aspects of personality structure may prove to be culturally invariant, that is, universal human ways of acting and experiencing. The four samples used in the present research (monolingual college samples from the United States and from Spain and bilingual Hispanic college and working-class samples) allowed us to begin to explore these substantive issues in personality psychology at the broad level of personality description implied by the Big Five dimensions.

The Big Five Dimensions of Personality Description

An important finding from lexical research on the structure of personality traits (Goldberg, 1993; John, 1990; Saucier &

This research was supported by a postdoctoral fellowship from the Ministerio de Educación y Ciencia (Spain) to Verónica Benet-Martínez and Grant MH49255 from the National Institute of Mental Health. We are grateful to Robert R. McCrae for his help in obtaining a license for our use of the English and Spanish NEO Five Factor Inventory items from Psychological Assessment Resources, Inc. We also thank Lewis R. Goldberg, Robert R. McCrae, and Gerard Saucier for their helpful comments on this work. The Spanish data in Study 1 could not have been collected without the generous help of the psychology faculty at the Universitat Autonoma de Barcelona (especially Montse Goma, Maite Martínez, and Jordi Bachs).

Correspondence concerning this article should be addressed to Verónica Benet-Martínez, Department of Psychology, University of Michigan, 525 East University Street, Ann Arbor, Michigan 48109, or to Oliver P. John, Institute of Personality and Social Research #5050, University of California, Berkley, California 94720-5050. Electronic mail may be sent to veronica@umich.edu.

Goldberg, 1996) is that a five-factor structure, the so-called Big Five (Goldberg, 1981), can capture much of the variance in personality trait ratings. Subsequently, evidence for the Big Five has been obtained across data sources, samples, and instruments (see Goldberg, 1993; McCrae & John, 1992), as well as several language families (see Katigbak, Church, & Akamine, 1996; McCrae & Costa, 1997a). The Big Five dimensions also show theoretically meaningful associations with important life outcomes, such as work and school performance (Barrick & Mount, 1991; John, Caspi, Robins, Moffitt, & Stouthamer-Loeber, 1994), well-being (Costa & McCrae, 1980), delinquency (John et al., 1994), and aspects of psychopathology (Widiger & Trull, 1992).

Note that the Big Five structure does not imply that personality differences can be reduced to only five traits. Rather, the Big Five dimensions represent personality at the broadest level of abstraction, and each dimension includes a large number of distinct, more specific personality characteristics (Costa & McCrae, 1995; John, 1990). Unfortunately, short English labels for dimensions as broad as the Big Five are difficult to come by, and the existing labels have numerous shortcomings and are easily misunderstood (Block, 1995; John, 1990, pp. 95–96); thus, we give short definitions of the five dimensions. Briefly, Extraversion summarizes traits related to activity and energy, dominance, sociability, expressiveness, and positive emotions. Agreeableness contrasts a prosocial orientation toward others with antagonism and includes traits such as altruism, tender-mindedness, trust, and modesty. Conscientiousness describes socially prescribed impulse control that facilitates task- and goal-directed behavior. Neuroticism contrasts emotional stability with a broad range of negative affects, including anxiety, sadness, irritability, and nervous tension. Openness describes the breadth, depth, and complexity of an individual's mental and experiential life.

The Big Five structure, however, is not the last word in taxonomies of personality; even its most ardent supporters recognize that the model has limitations (for reviews, see Benet & Waller, 1995; Benet-Martínez, 1997; Block, 1995; John & Robins, 1993, 1994; McAdams, 1992). As McCrae and John (1992) summarized,

> There are disputes among five-factorists about the best interpretation of the factors; there are certainly important distinctions to be made at the level of the more molecular traits that define the factors; and it is possible that there are other basic dimensions of personality. (p. 177)

For example, recent factor analyses of broad sets of personality descriptors suggest two highly evaluative dimensions in addition to the Big Five dimensions (Almagor, Tellegen, & Waller, 1995; Benet-Martínez & Waller, 1997).

Measuring the Big Five Dimensions

In English, a variety of measures are available to assess the Big Five in adults and adolescents (Costa & McCrae, 1992; Goldberg, 1992; John et al., 1994; John, Donahue, & Kentle, 1991; Saucier, 1994; Trapnell & Wiggins, 1990). In Spanish, the only published instrument is a recent translation of the 240-item Revised NEO Personality Inventory (NEO PI-R; Costa & McCrae, 1992).[1] An initial evaluation of this translation was conducted by Marc Gellman in an unpublished study summarized in a NEO PI-R manual supplement (Psychological Assessment Resources, 1994); 74 Hispanic bilingual college students completed both the English and Spanish versions in one testing session. The Spanish NEO PI-R scales had adequate alpha reliabilities and substantial cross-language convergent validities. The English version of the NEO PI-R has a 60-item short version, called the NEO Five Factor Inventory (NEO-FFI); however, to date there has been no research on a Spanish NEO-FFI, nor is there a published version.

One of the strengths of the NEO PI-R is that it permits differentiated measurement of each Big Five dimension in terms of more specific facets (Costa & McCrae, 1995). However, for many research applications, especially with less acculturated and noncollege Hispanic samples, the Spanish NEO PI-R may be rather lengthy and some of the items may be difficult to understand. In fact, as Burisch (1984) showed, "Short scales not only save testing time but also avoid subject boredom and fatigue . . . there are subjects . . . from whom you won't get any response if the test looks too long" (p. 219). Thus, there is a need for a Spanish Big Five instrument that has short and easily understood items and requires no more than 5 min of administration time. The Spanish Big Five Inventory (BFI) examined in this article was designed to fill this need.

The 44-item English BFI (John et al., 1991) was constructed to allow efficient and flexible assessment of the five dimensions when there is no need for more differentiated measurement of individual facets. Items were selected from Big Five prototype definitions (see John, 1990, Table 3.2) that had been developed through expert ratings and subsequent factor analytic verification in observer personality ratings. Because single trait adjectives are answered less consistently than when they are accompanied by definitions or elaborations (Goldberg & Kilkowski, 1984), the BFI does not use single adjectives as items; instead, one or two prototypical trait adjectives served as the item core to which elaborative, clarifying, or contextual information was added. For example, the Conscientiousness adjective *persevering* served as the basis for the BFI item "Perseveres until the task is finished," and the Openness adjective *original* became the BFI item "Is original, comes up with new ideas." Thus, the BFI items are short and avoid complex sentence structures, retaining the advantages of adjectival items (brevity and simplicity) while avoiding some of their pitfalls (ambiguous or multiple meanings and salient desirability). Moreover, whereas it is often difficult to find exact single-word translations for trait adjectives, the meanings of elaborated phrases are more easily translated (Hofstee, 1990; John, Goldberg, & Angleitner, 1984). The BFI is available to interested researchers and has been used in a wide range of studies, including Clark (1992); Neuberg and Newsom (1993); Watson, Clark, and Harkness (1994); Cialdini,

[1] In Spain, a Castillian translation of the earlier NEO-PI (Costa & McCrae, 1985) was developed by Silva et al. (1994). Moreover, in their research on Agreeableness, Jensen-Campbell, Graziano, and Hair (1996) used Spanish translations of Goldberg's (1992) Big Five marker adjectives for a subgroup of Mexican Americans who "had limited fluency in English" (p. 153).

Trost, and Newson (1995); Gross and John (1995); and S. Johnson and Wolfe (1995).

Overview

There are both practical and theoretical reasons for examining the BFI in Spanish-speaking samples. The practical reasons stem from the need for a short and easily understood Big Five measure for the various Spanish-speaking populations in Spain, Latin America, and the United States. From a theoretical point of view, it is important to examine whether the covariation among the specific traits that define the Big Five in English differs in some fundamental way from their covariation in Spanish and Hispanic samples, thus adding information regarding the cross-cultural status of the Big Five. Previous research has shown that the structure of an instrument may change when translated and administered in another language context. For instance, different factor structures have been found for the Spanish versions of well-known psychological instruments such as Rotter's measure of internal–external locus of control (Garza, 1977), the Minnesota Multiphasic Personality Inventory (Gonzalez Valdes, 1979), and the Eysenck Personality Inventory (García Sevilla, Pérez, & Tobeña, 1979).

The three studies reported here focus on a Spanish version of the BFI. Thus, rather than following an *emic* research strategy that would identify indigenous personality dimensions (e.g., Church & Katigbak, 1989; Yang & Bond, 1990), we used an *imposed-etic* strategy (Berry, 1980; Triandis & Marín, 1983). In Study 1, we compared the Spanish and English BFI in college students from Spain and the United States and assessed factor congruence across languages. In Study 2, we examined the Big Five in a college-educated sample of bilingual Hispanics and tested convergent and discriminant validity across both languages and two instruments. In Study 3, we replicated the BFI findings from Study 2 in a working-class Hispanic bilingual sample.

The interpretation of findings from cross-cultural research is complicated by the fact that differences between different-language versions of an instrument may be due to differences between translations, languages, samples, cultures, or a mixture of all of them. We therefore tried to address these challenges by using culturally sensitive translation procedures and testing psychometric equivalence across samples, languages, and instruments. Furthermore, we studied bilingual samples to help unconfound the effects of language and sample differences (John et al., 1984) and recruited a working-class sample to test the generalizability of the Spanish BFI across socioeconomic groups.

Study 1: College Students in Spain and the United States

This study compared the Spanish and English versions of the BFI using two large samples of college students, one from Spain and one from the United States. One of the limitations of the imposed-etic approach (Church & Katigbak, 1988) used in the development of the Spanish BFI scales is that it might leave out culturally salient aspects of the Spanish Big Five. Thus, we also examined how well the translated Spanish BFI scales converged with a set of Big Five scales defined by indigenous Spanish items.

Method

U.S. and Spanish samples. The U.S. sample consisted of 711 undergraduate students (300 men and 411 women) at the University of California at Berkeley. Their mean age was 21 years (SD = 3.3). A wide range of majors was represented, and the majority of the participants were non-psychology majors. Participants completed the English-language BFI on their own time.

The Spanish sample consisted of 894 native residents of Spain (191 men and 703 women). Participants were undergraduate students attending the Universitat Autónoma de Barcelona, a prestigious public university in northeastern Spain. Their mean age was 21 years (SD = 3.9). As in the U.S. sample, the majority of the Spanish participants represented a wide range of non-psychology majors. In the context of a larger study, these participants completed a series of personality inventories, including a Spanish translation of the BFI and a dictionary-based list of indigenous Spanish personality descriptors. Questionnaires were completed during group testing sessions.

English-language BFI. The BFI (John et al., 1991) uses short phrases to assess the most prototypical traits associated with each of the Big Five dimensions in English (John, 1990). The trait adjectives (e.g., *thorough*) that form the core of each of the 44 BFI items (e.g., "does a thorough job") have been shown in previous studies to be univocal, prototypical markers of the Big Five dimensions (John, 1989, 1990). The English BFI items are reprinted in the Appendix. Participants rate each BFI item on a 5-point scale ranging from 1 (*disagree strongly*) to 5 (*agree strongly*); scale scores are computed as the participant's mean item response (i.e., adding all items scored on a scale and dividing by the number of items on the scale).

Despite its brevity, the BFI does not sacrifice either content coverage or good psychometric properties. For example, the eight-item Extraversion scale includes items from at least four of the six facets postulated by Costa and McCrae (1992)—namely, gregariousness, activity, assertiveness, and positive emotions. In U.S. and Canadian samples, the alpha reliabilities of the BFI scales typically range from .75 to .90 and average above .80; 3-month test–retest reliabilities range from .80 to .90, with a mean of .85. Moreover, the intercorrelations among the five scales tend to be low; most of them are below .20, and it is rare for one or two of them to exceed .30 (John & Donahue, 1998). In terms of convergent validity with other Big Five instruments, the BFI scales correlate more highly with both Costa and McCrae's and Goldberg's (1992) scales (mean rs = .75 and .80, respectively) than these two correlate with each other (mean r = .65). Two peer-rating studies provide further validity evidence: On average, the BFI self-report scales correlated .47 with reports from two peers in a college sample and .61 with reports from five family members and peers in an adult community sample (John & Donahue, 1998).

Although the BFI scales show substantial convergent validity with Costa and McCrae's (1992) factor definitions, there are some subtle but important differences for Extraversion and Openness. Preliminary BFI items intended to represent the Extraversion facets of excitement seeking and warmth did not cohere well enough with the other items to be included in the final BFI Extraversion scale. Similarly, items measuring liberal versus conservative values (for the openness to values facet) and behavioral flexibility (for the openness to actions facet) failed to make it onto the BFI Openness scale. Thus, not surprisingly, the convergent validity correlations between the BFI and Costa and McCrae's measures tend to be somewhat lower for Extraversion and Openness.

Spanish (Castillian) translation of the BFI. The Spanish spoken in Spain (referred to as Castillian) differs slightly from the Spanish used in Latin America and the United States. Therefore, the Spanish participants

Table 1
Study 1: Psychometric Properties of the English and Spanish Big Five Inventory Scales

Scale	*n*	α		*M*		*SD*		Correlations with indigenous Big Five scales[a]	
		United States	Spain	United States	Spain	United States	Spain	*r*	Corrected *r*
Extraversion	8	.88	.85	3.2	3.4	.8	.8	.77	.89
Agreeableness	9	.79	.66	3.8	3.8	.6	.5	.60	.83
Conscientiousness	9	.82	.77	3.6	3.5	.7	.7	.63	.79
Neuroticism	8	.84	.80	3.0	3.2	.8	.8	.68	.83
Openness	10	.81	.79	3.7	3.8	.6	.6	.53	.66
M	9	.83	.78	3.5	3.5	.7	.7	.65	.81

Note. $N = 894$ Spaniards and 711 Americans; n = number of items in the scale.
[a] Correlations in the Spanish sample only; correlations were corrected for attenuation due to unreliability using alpha.

completed a Castillian version of the BFI that was developed using the back-translation method of Brislin (1980). Using standard Spanish–English and English–Spanish dictionaries, Verónica Benet-Martínez (who is bilingual) undertook the translation of the BFI items into Spanish. Using the same dictionaries, a second bilingual individual (with a Ph.D. in Spanish) independently translated the material back into English. We then compared the back-translated version with the initial English version, discussed discrepancies between the translators, and generated further translations until we arrived at a final set of Spanish BFI items that both translators agreed best operationalized the condition of being symmetrically translatable to the English originals.

Indigenous Spanish Big Five markers. Spanish participants also provided self-reports on a list of indigenous Spanish personality descriptors developed by Benet-Martínez and Waller (1997). This list consisted of 299 personality-descriptive adjectives randomly selected from a widely used unabridged Spanish dictionary (Real Academia Española, 1989). Selecting items from this indigenous Spanish item set, Benet-Martínez and John (1998) used rational and factor analytic procedures to develop 12-item markers for each of the Big Five. Using separate derivation and replication samples, the Big Five factors were clearly replicated and the 12-item scales all had substantial alpha reliabilities. The highest-loading items were "comical, funny" for Extraversion, "good natured" for Agreeableness, "thinks before acting" for Conscientiousness, "easily upset" for Neuroticism, and "unconventional" for Openness. We used these 12-item scales to examine how well the imported (etic) Spanish BFI scales converged with Big Five scales defined by indigenous Spanish items.[2]

Results and Discussion

Basic psychometric characteristics and group differences. For each of the BFI scales, Table 1 shows the number of items, internal consistency (alpha) reliability, mean, and standard deviation, separately in the U.S. and Spanish samples. As expected, the internal consistencies for the English-language scales were substantial (mean $\alpha = .83$). The alpha coefficients for the Spanish translations were slightly lower (mean $\alpha = .78$). In both the U.S. and Spanish samples, the Extraversion scale showed the highest alpha reliability and the Agreeableness scale the lowest.

Table 1 also shows that the English and Spanish scales had very similar means and standard deviations in the U.S. and Spanish samples. In both samples, the highest means were found for Agreeableness and Openness, followed by Conscientiousness, then Extraversion, and Neuroticism last. Thus, the rank ordering of the means was the same in the two samples.[3]

To test differences between the Spanish and U.S. participants more formally, we correlated the Big Five scale scores with the cultural background of the participants. The United States was coded as 1 and Spain as 2; thus positive correlations indicate that Spaniards had higher scores than U.S. participants. Note that these analyses are based on the total N of 1,605 and provide a powerful test of group differences. Therefore, even minute differences will attain statistical significance, and interpretation has to focus on effect sizes, rather than significance. Three correlations were significant at $p < .01$: The highest was .12 for Extraversion, followed by −.09 for Conscientiousness, and

[2] Although each of the Spanish personality items is indigenous, the Big Five scales obtained with these items do not provide an indigenous instrument because the items were selected to represent an a priori structure, namely the Big Five. The important notion here is that because only those indigenous Spanish terms that correlated highly with the a priori Big Five scale were selected, the Big Five factor structure identified with these terms cannot be viewed as the naturally emerging indigenous structure of personality description in Spain (see Benet-Martínez & Waller, 1997, and Yang & Bond, 1990, for examples of true emic approaches to the identification of indigenous dimensions).

[3] In most research on the Big Five, gender differences tend to be small and factor structures replicate closely across the sexes (e.g., Borkenau & Ostendorf, 1990). Although gender differences were not the focus of the present research, it is of interest to note that across all three studies, only two of the Big Five dimensions showed gender differences, and these differences were small but consistent across U.S. and Spanish samples, English and Spanish instruments, and the BFI and NEO-FFI (in Study 2). For Neuroticism, all eight correlations with sex (keyed 1 for female and 0 for male) were positive, ranging from .03 to .27, with a mean of .18. For Agreeableness, seven of the eight correlations were positive, ranging from −.11 to +.19, with a mean of .08. These correlations are very similar to these found for the full-length NEO PI-R (Costa & McCrae, 1992, p. 55). In short, women tend to score slightly higher on Neuroticism and Agreeableness regardless of instrument, language, and culture.

Table 2
Study 1: Intercorrelations Between the English and Between the Spanish Big Five Inventory Scales

Scale	E	A	C	N	O
Extraversion (E)	—	.17	.09	−.18	.33
Agreeableness (A)	.14	—	.17	−.23	.16
Conscientiousness (C)	.24	.27	—	−.20	.17
Neuroticism (N)	−.29	−.31	−.18	—	−.14
Openness (O)	.25	.05	.08	−.14	—

Note. Correlations for the Spanish sample (n = 894) are above the diagonal; correlations for the U.S. sample (n = 711) are below the diagonal.

.08 for Neuroticism. For Agreeableness and Openness, the correlations were .03 and .06, respectively. Given the size of these correlations, one might not want to interpret them at all. Alternatively, the slightly higher Extraversion scores for Spaniards might be viewed as consistent with the cultural value of *simpatía* (e.g., expressing positive emotions), and their slightly lower Conscientiousness scores as consistent with the flexible time orientation (vs. delay of gratification) assumed to characterize Latin cultures. However, the Neuroticism effect is difficult to interpret, and the lack of an Agreeableness difference would seem inconsistent with the *simpatía* script of promoting smooth and pleasant relationships.[4] More generally, at the broad level of abstraction represented by the Big Five, we did not find sizable differences between the two cultural groups studied here.

Table 2 shows the intercorrelations among the BFI scales within each language. These findings are important because there has been concern that some of the Big Five dimensions are highly intercorrelated (e.g., Block, 1995). The present results show that in our large samples the BFI scales were fairly independent, and this was true in both languages. For instance, the absolute mean of the intercorrelations was .18 in Spanish and .19 in English. Even the highest intercorrelation was only .33 for the Spanish BFI and −.31 for the English BFI. In summary, the English and Spanish versions of the BFI scales had similar psychometric characteristics. Not surprisingly, the newly developed Spanish scales had slightly lower alpha reliabilities.

Comparison of exploratory English and Spanish factors. The varimax-rotated principal components for the Spanish and English versions are presented in Table 3. The U.S. sample replicated the expected five-factor structure of the English-language BFI. More important, a five-factor solution was also uncovered in the Spanish BFI, as indicated by a clear break after the fifth eigenvalue in the Spanish data. Note that Table 3 shows remarkably clear evidence of simple structure, not only in the English original but also in the Spanish translation. For example, for both Extraversion and Conscientiousness, every targeted item loaded substantially (i.e., over .30) on the intended factor and none of the cross-loadings exceeded .22. As summarized at the bottom of Table 3, the congruence of the factor loadings across languages was excellent, with all cross-language factor congruence coefficients exceeding .90 (mean r = .94).

The factor structures reported in Table 3 show a few item-level departures from perfect congruence that are worth discussing. First, note the zero loading for our Spanish translation of the Openness item "likes to reflect, play with ideas." After reviewing the Spanish translation, it became apparent that the Spanish version differed in a subtle but consequential way. Although both versions involve thought and reflection, our initial Spanish translation ("*es dado a la reflexión, al análisis*") does not denote the playful intellect of the English original but instead refers to careful reflection and analysis more characteristic of Conscientiousness than Openness (indeed, the Spanish item's strongest loading was on the Conscientiousness factor). This factor analytic finding, similar to Peabody and Goldberg's (1989) distinction between expressive and controlled intellect, led us to revise the translation to "*le gusta reflexionar, jugar con las ideas*" for our subsequent studies.

The secondary loading of the Neuroticism item "is depressed, blue" on the low pole of Extraversion replicates earlier U.S. findings and appears in both the U.S. and Spanish samples. This replicated finding is consistent with the view that depressed mood involves both high Neuroticism, or high negative affect, and low Extraversion, or low positive affect (J. A. Johnson & Ostendorf, 1993; Larsen & Diener, 1992; Saucier, 1992a; Tellegen, 1985; Watson & Clark, 1992). The inclusion in the BFI of an item tapping depressed mood was considered sufficiently important to tolerate a factorially complex item on the test (John et al., 1991).

Finally, the Agreeableness items "likes to cooperate with others" and "can be cold and aloof" had secondary loadings on Extraversion in the Spanish structure. Further examination of the Spanish translations for these items suggested no obvious translation problems, and thus cultural or sample differences are probably responsible for these differences (see Marín, Triandis, Betancourt, & Kashima, 1983, for a discussion of how cultural differences in affective meaning may lead to discrepancies between equivalent questionnaire items). One interpretation for the "migration" of these two Agreeableness items toward Extraversion is that Spaniards may attach a stronger communal value to Extraversion than do North Americans. These departures also suggest a more negative view of introversion in the Spanish culture than in the United States, and they are consistent with findings for other personality taxonomies studied in Spain (see Benet & Waller, 1995; Benet-Martínez & Waller, 1997).

Despite these differences, the similarity of factor loadings across the languages (see Table 3) provides evidence for the structural similarity between the English and Spanish BFI. A more formal test of cross-language convergence was conducted using a multisample confirmatory factor analysis (CFA) with the EQS program (Bentler, 1995). CFA can be thought of as a special case of traditional factor analytic models in which the number of factors and the variables that load on particular factors are specified in advance. We conducted a multisample CFA so that the structure in both the Spanish and the U.S. samples could be examined in one joint analysis; this analysis examines

[4] Follow-up analyses at the item level did not provide consistent support for cultural differences, either. For example, the positive emotion item on Extraversion ("generates a lot of enthusiasm") failed to show the expected cultural effect (r = .03, ns). Similarly, some of the Agreeableness items showed effects opposite to those expected from the cultural *simpatía* script; for example, Spaniards described themselves as more cold (r = .12, $p < .01$) and less forgiving (r = −.24, $p < .01$).

Table 3
Study 1: Varimax-Rotated Five-Factor Structure for English Big Five Inventory (BFI) Items (U.S. Samples) and Initial Spanish Translations (Spanish Sample)

BFI items (English text)	Extraversion		Agreeableness		Conscientiousness		Neuroticism		Openness	
	US	SP	US	SP	US	SP	US	SP	US	SP
Extraversion										
43 Is outgoing, sociable	**77**	**73**	11	08	00	−02	−10	−04	00	01
1 Is talkative	**67**	**73**	−08	−10	−02	−02	04	10	00	−03
40 Has an assertive personality	**60**	**44**	−21	00	16	09	−09	−12	09	09
32 Generates a lot of enthusiasm	**59**	**55**	16	05	08	05	−11	−13	−03	16
11 Is full of energy	**53**	**51**	13	05	16	12	−19	−22	−03	15
6 Is reserved	**−69**	**−65**	−04	05	00	07	00	−05	−15	−07
27 Is sometimes shy, inhibited	**−71**	**−59**	07	17	−12	−09	19	13	−16	−14
16 Tends to be quiet	**−78**	**−77**	11	11	−03	03	02	−08	−15	−17
Agreeableness										
37 Is considerate and kind to almost everyone	−01	−01	**67**	**51**	−01	14	−02	−02	−08	00
41 Likes to cooperate with others	14	**32**	**51**	24	01	04	−12	−03	−11	00
7 Is helpful and unselfish with others	00	03	**48**	**46**	08	11	−05	00	−13	−09
28 Has a forgiving nature	−03	02	**46**	**35**	−04	−07	−14	−15	−01	−09
24 Is generally trusting	08	05	**38**	**30**	11	−04	−05	00	−21	−16
2 Tends to find fault with others	01	10	**−47**	**−39**	−08	−01	17	14	−12	−15
13 Starts quarrels with others	09	10	**−49**	**−43**	−04	−01	06	08	−12	−06
33 Can be cold and aloof	−29	**−38**	**−55**	**−37**	−15	01	03	00	−05	−10
22 Is sometimes rude to others	−12	00	**−58**	**−49**	00	−15	01	05	−15	−16
Conscientiousness										
3 Does a thorough job	03	−01	00	06	**66**	**58**	06	04	−06	00
29 Does things efficiently	15	08	−04	02	**59**	**34**	−05	−13	−14	08
34 Makes plans, follows through with them	13	−07	−03	−13	**57**	**50**	−14	−05	−10	−15
14 Is a reliable worker	−06	03	05	18	**52**	**33**	00	−05	−21	−11
21 Perseveres until the task is finished	00	06	07	02	**50**	**53**	07	−02	00	09
42 Is easily distracted	−01	−02	−09	−02	**−53**	**−45**	21	14	−09	−16
8 Can be somewhat careless	−01	00	−15	00	**−58**	**−54**	00	00	−10	03
25 Tends to be lazy	−15	−14	−08	−08	**−60**	**−60**	03	00	−16	−10
18 Tends to be disorganized	−06	01	02	−04	**−62**	**−63**	03	02	00	01
Neuroticism										
26 Worries a lot	−22	−08	−06	11	−01	18	**68**	**41**	−15	−11
15 Can be tense	−14	−10	−23	−21	05	00	**59**	**64**	−12	−11
38 Gets nervous easily	−29	−07	−01	−04	−05	−07	**56**	**69**	−24	−14
4 Is depressed, blue	**−39**	**−41**	−24	−06	−11	−10	**42**	**45**	03	−12
30 Can be moody	−10	−20	**−35**	−22	−07	−20	**42**	**41**	01	−06
35 Remains calm in tense situations	−04	−06	03	07	17	15	**−56**	**−55**	03	−03
19 Is emotionally stable, not easily upset	00	05	14	07	09	14	**−63**	**−59**	−05	−04
9 Is relaxed, handles stress well	04	−06	13	10	−02	06	**−74**	**−66**	00	−08
Openness										
23 Is inventive	11	19	−04	−11	02	00	−07	−12	**58**	**64**
5 Is original, comes up with new ideas	14	17	−09	−12	05	04	−10	−12	**57**	**62**
17 Values artistic, aesthetic experiences	−10	−04	07	17	−10	02	05	04	**55**	**40**
20 Has an active imagination	06	18	−04	−07	−09	−05	00	−12	**53**	**56**
36 Likes to reflect, play with ideas	−07	−21	−11	06	00	24	−12	−09	**51**	05
39 Is sophisticated in art, music, or literature	04	00	04	12	01	00	04	09	**48**	**46**
31 Is ingenious, a deep thinker	03	21	−18	−16	09	04	05	−14	**42**	**49**
10 Is curious about many different things	04	24	03	16	−09	01	−11	−08	**40**	**29**
12 Prefers work that is routine	−04	−10	10	08	03	−10	13	04	**−37**	**−39**
44 Has few artistic interests	−06	00	−03	−08	−04	00	−02	−07	**−58**	**−53**
Cross-language factor congruence coefficient	.95		.92		.94		.96		.91	

Note. N = 711 Americans and 894 Spaniards. US = United States; SP = Spain. All loadings were multiplied by 100; loadings |.30| or larger are set in bold. Item numbers refer to the order of the items in the English BFI and the final Spanish BFI, both given in the Appendix.

the goodness of fit of both the Big Five model and the cultural equivalence of the factor structures. We tested two models. In both models, we specified five latent factors representing the a priori Big Five dimensions: All the primary loadings and the two replicated cross-loadings ("is depressed, blue" and "can be moody") were freely estimated, and the loadings of all non-defining items were fixed to zero. To permit a strong test of the invariance of the Big Five model across the English and Spanish BFI, we constrained all primary loadings and the two cross-loadings to equality across the two samples. Because in previous

CFA research (Borkenau & Ostendorf, 1990) the Big Five were found to be moderately intercorrelated, rather than strictly orthogonal, the covariances among the five factors were set to zero in Model 1, whereas in Model 2 we allowed for correlated factors, with the covariances freely estimated. Evaluation of the fit of the two models was based on multiple criteria (Bentler, 1990).[5]

Results for Model 1 indicated inadequate levels of fit. Model 2 (which allowed for correlated factors) had acceptable fit indices, $\chi^2/df = 2.11$ and CFI = .92, and significantly improved the fit over Model 1, $\Delta\chi^2(21) = 8{,}708$, $p < .001$, suggesting that only the model with correlated factors provided an adequate fit. We also examined the Big Five intercorrelations estimated by the CFA. As one would expect from the observed correlations (Table 2), the estimated latent correlations (which are corrected for unreliability by CFA) were generally small and similar in the two samples; of the 20 correlations, only 4 exceeded .25 and only one exceeded .30 (Extraversion and Openness in the Spanish sample), with absolute means of .17 in the U.S. sample and .19 in the Spanish sample.

In summary, the substantial cross-language congruence of the varimax factors and CFA fit indices support the Big Five structure as an adequate solution for the major sources of variance underlying both the English and the Spanish BFI; in both languages, the Big Five dimensions showed only small intercorrelations, allaying concerns about the overlap among the Big Five dimensions (Block, 1995).

Comparison with indigenous Spanish Big Five markers. How well do the translated BFI scales capture the Big Five dimensions as defined with indigenous items? To find out, we computed the correlations between the Spanish BFI scales and the Big Five marker scales defined by 12 indigenous items. These cross-instrument correlations in the Spanish sample are given in the last column of Table 1 and averaged .65; only the Openness dimension showed a correlation below .60. These substantial convergent validity correlations contrasted with much lower discriminant correlations, which averaged .16. Thus, for at least four of the Spanish BFI scales, our findings suggest considerable convergence and discrimination with a set of Big Five dimensions measured with indigenous items.

To determine the extent to which the validity correlations are limited by the imperfect reliability of the two sets of scales, we also corrected for attenuation using alpha. As shown in Table 1, these correlations averaged .81. The only value lower than .70 was found for Openness.[6] Thus, as in other lexical personality research, the composition of the fifth factor in the Big Five showed the least convergence across languages and cultures (Bond, 1994; Church & Katigbak, 1989; Hofstee, Kiers, de Raad, Goldberg, & Ostendorf, 1997). Apparently, there were some systematic differences in the way this dimension was defined in the English BFI and the Spanish indigenous terms. To further examine the nature of these differences, we compared the BFI Openness items with the indigenous Spanish markers for Openness. As one would expect from the substantial convergent correlation, several basic elements were represented in both instruments—namely, openness to ideas, to fantasy, and to aesthetics. However, in the indigenous Spanish scale, Openness also included interests, preferences, and attitudes that define the open-minded lifestyle: unconventional attitudes and tastes, enjoyment of travel, a bohemian and world-open approach to life, and interest in spiritual issues. This broader definition resembles more closely McCrae and Costa's (1997b) definition of this factor than its definition as Intellect or Imagination found in the English and German lexical studies (Ostendorf, 1990; Saucier, 1992b; Saucier & Goldberg, 1996).

For the other four dimensions, the factor definitions in the BFI and the indigenous Spanish marker items showed only minor differences in emphasis. In Spain, humor seemed a particularly important facet of Extraversion; lack of anger-proneness an important facet of Agreeableness; sound judgment in the daily matters of existence (e.g., sensible, well-balanced, not excessive) an important facet of Conscientiousness; and anxiety, fear, and worry central facets of Neuroticism. Future research should use an emic strategy (e.g., Yang & Bond, 1990) to test how central these facets are in personality description in Spain. In general, then, the Spanish BFI scales showed substantial convergent correlations and considerable overlap in item content with the indigenous markers.

Limitations and Caveats

In addition to the etic-imposed research strategy, two other limitations of this study must be considered. First, the monolingual design is limited by the fact that differences between the U.S. and Spanish findings may be due to translation differences, sample differences, cultural differences, or some combination of the three (John et al., 1984). Second, because the Spanish-speaking sample included participants from Spain only, the generalizability of our findings to Spanish-speaking populations outside Spain remains to be demonstrated. Study 2 was designed to address these limitations by (a) using a bilingual design, (b) testing the psychometric properties of the Spanish BFI in a Hispanic sample from the United States, and (c) comparing the Spanish BFI with another Spanish-language measure of the Big Five.

Study 2: College-Educated Hispanic Bilinguals

The main goals of Study 2 were to examine the cross-language validity of the final version of the Spanish BFI scales in a bilingual Hispanic sample and to assess the convergence of the BFI with a short version of Costa and McCrae's (1992) Spanish NEO PI-R. Both the Spanish and the English versions of the two instruments were administered to the same participants. Such a bilingual design has important advantages over monolingual designs because it can help unconfound the effects of language and sample differences (John et al., 1984). In this

[5] Readers should be aware that in the analysis of personality data these indices often seem to underestimate the fit for the model (especially when samples are large as in this study) and should thus be interpreted as conservative tests of model fit (Borkenau & Ostendorf, 1990; Katigback et al., 1996; McCrae, Zonderman, Costa, Bond, & Paunonen, 1996).

[6] One possible reason for this lower convergence is, as we noted above, that one of the BFI Openness to Experience items was translated incorrectly.

bilingual design, we can directly assess the extent to which the Spanish translations converge with the original English instrument. This use of multiple languages and multiple instruments to measure multiple traits is an extension of the multitrait–multimethod approach (Campbell & Fiske, 1959), a construct validation procedure used to assess the convergence of independent measures of the same trait and the discrimination among measures of different traits. Because the multimethod component of our study involved two sources of method variance (i.e., instruments and languages), we were able to test a series of CFA models specifying different kinds of method effects.

Method

Participants and recruitment procedures to identify bilinguals. The sample included 170 Hispanic individuals (66 men and 104 women) living in the San Francisco Bay Area. Their mean age was 25 years (SD = 10). All participants lived in the United States and were either immigrants from Latin America (n = 80) or U.S.-born descendants of Latin American individuals (90). Immigrant participants were from Mexico (n = 29), El Salvador (13), Argentina (8), Nicaragua (5), Peru (5), Colombia (6), Chile (3), Cuba (3), Guatemala (3), Panama (2), Venezuela (2), and Ecuador (1). U.S.-born participants had backgrounds from Mexico (n = 76), Cuba (3), El Salvador (3), Costa Rica (2), Guatemala (2), Bolivia (2), and Colombia (2). Participants were students (n = 143) who received course credit for their participation and college-educated community residents (27) who volunteered to participate.

Bilingualism is not an either-or category but reflects a set of skills that individuals possess to varying degrees (Reynolds, 1991). Because our study required that participants have good reading comprehension in both English and Spanish, *bilingualism* was defined operationally to the participants as "the ability to read and fully understand novels written in English and Spanish." We used a three-step process in recruitment and screening of bilingual participants. First, students had to report being bilingual in English and Spanish on a prescreening form in introductory psychology courses. Second, they were contacted by phone and asked to confirm their bilingual status. The third step in screening for bilingualism consisted of two translation tests given before administering the personality measures. The goal of these tests was to identify participants who might have misjudged the extent of their bilingualism; in fact, several potential participants found they could not perform the translation tests and elected not to participate further. The translation tests also provided us with an objective measure of our participants' bilingual status. Community residents were contacted by flyers or mail and also completed the translation tests.

Procedures. Participants completed the materials in small group sessions. They completed (a) two translation tests; (b) a demographic background and language-use section; (c) the original English versions of the BFI and the NEO-FFI (Costa & McCrae, 1992); and (d) the Spanish adaptations of these two Big Five instruments, with the items given in a different order than in the English originals. To control for potential order-of-language effects, half of the participants completed the English-language instruments first, whereas the other half completed the Spanish ones first. Because of the possibility of memory effects in this design, we separated the English and Spanish questionnaires with a 5-min unrelated filler task. Moreover, responses in one language were collected before the other language materials were handed out, thus ensuring that participants would not check their earlier responses.

Translation tests and language use. The written translation tests asked participants to translate two short paragraphs—one written in Spanish and the other in English—into the other language. Each paragraph described the personality of a fictitious individual and contained words and expressions of similar difficulty. The translations were scored by a bilingual judge, who deducted points for translation errors (defined as incorrect use of vocabulary and grammar). In order to ensure scoring reliability, a second bilingual individual graded the translations from 10 randomly chosen participants. Because interjudge agreement correlations across the 10 participants were .94 for the English and .97 for the Spanish translations, only the scores of the first judge were used here. On average, participants scored 91% correct (SD = 11%) on the Spanish-to-English test and 83% correct (SD = 9%) on the English-to-Spanish test; these impressive test scores indicate that the participants who passed through our bilingualism screens had a high level of English–Spanish bilingualism, and no further participants had to be excluded.

Participants also reported the percentage of time in their daily lives they spoke Spanish rather than English. On average, participants reported speaking Spanish 32% of the time (SD = 19%), suggesting that the bilingual individuals in the present sample participated actively in the predominantly English-speaking culture but also retained contact with their own language community.

English BFI and final Spanish adaptation. Given the differences between Castillian (the Spanish used in Spain) and the Spanish used in Latin America and the United States, we revised the initial Spanish translations of the BFI items described in Study 1. Four Spanish–English bilingual individuals native from Mexico, Argentina, Venezuela, and Colombia independently revised the original Spanish BFI and modified the wording of those items that included linguistic elements that seemed foreign or unfamiliar in their country's language context. This procedure was repeated by asking each bilingual individual to examine the versions created by the other three bilingual individuals. In each phase, the translators tried to develop a "pan-Spanish" translation of the BFI that could be used by Spanish-speaking individuals of all backgrounds. The translators made an effort to avoid the use of "native" terms and instead offered "generic" Spanish wordings. This process of revision continued until agreement among all four translators was achieved; 23 of the 44 items were modified. The final Spanish version of the BFI is included in the Appendix, along with instructions.

English and Spanish NEO-FFI. The NEO-FFI is a 60-item, abbreviated version of the 240-item English NEO PI-R (Costa & McCrae, 1992) measure of the Big Five dimensions. The NEO PI-R was developed in samples of middle-aged and older adults, using both factor analytic and multimethod validational procedures of test construction. The NEO PI-R scales have shown substantial internal consistency, temporal stability, and notable convergent and discriminant validity against spouse and peer ratings in English-speaking samples (McCrae & Costa, 1990).

As described in a recent manual supplement (Psychological Assessment Resources, 1994), a Spanish version of the 240-item NEO PI-R was developed by a professional translator familiar with Spanish as spoken in the United States. A study of the Spanish NEO PI-R (summarized in the manual supplement) involved 74 bilingual college students who completed the English and Spanish versions in one testing session; convergence between the two versions was substantial, and the internal consistencies for the Spanish scales were adequate.

In this study, rather than using all 240 items from the Spanish NEO PI-R, we used the 60 items that constitute the shorter NEO-FFI. A study of the Spanish version of the NEO-FFI is of interest because for many research applications the full NEO PI-R is too long, and the shorter NEO-FFI has not yet been examined. Moreover, the 60-item NEO-FFI is more appropriate as a comparison for the 44-item BFI. The fact that the NEO-FFI scales are longer than the BFI scales (12 vs. 9 items per scale) and require more than twice as long to complete led us to expect that the NEO-FFI scales would show somewhat higher reliabilities and cross-language convergence than the BFI scales. We administered the NEO-FFI items using the same 1–5 rating scale as the BFI, and scales were scored as mean item responses to make the scale scores directly comparable.

Table 4
Study 2: Psychometric Characteristics of the Spanish and English Big Five Inventory (BFI) and NEO Five Factor Inventory (NEO-FFI) Scales in a College-Educated Hispanic Bilingual Sample

	n		α				*M*				*SD*			
			BFI		NEO-FFI		BFI		NEO-FFI		BFI		NEO-FFI	
Scale	BFI	NEO-FFI	Eng	Spa	Eng	Spa	Eng	Spa	Eng	Spa	Eng	Spa	Eng	Spa
Extraversion	8	12	.87	.84	.81	.72	3.5	3.4	3.6	3.5	.82	.79	.61	.52
Agreeableness	9	12	.80	.65	.78	.69	4.0	3.8	3.8	3.7	.59	.50	.56	.53
Conscientiousness	9	12	.86	.76	.87	.83	3.9	3.7	3.8	3.7	.67	.60	.64	.60
Neuroticism	8	12	.84	.81	.89	.82	2.7	2.7	2.6	2.7	.80	.77	.80	.68
Openness	10	12	.86	.82	.76	.73	3.9	3.9	3.6	3.6	.66	.60	.55	.55
M	9	12	.85	.78	.82	.76	3.6	3.5	3.5	3.5	.71	.66	.63	.58

Note. $N = 170$ English–Spanish bilingual individuals; n = number of items in the scale; Eng = English; Spa = Spanish. Both instruments were administered with a 5-point rating scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Scale scores were computed as the mean rating of the items on each scale (after reversing false-keyed items); thus, scale score metrics are directly comparable across instruments and Big Five scales based on different numbers of items.

Results and Discussion

Basic psychometric characteristics and language differences. For each of the BFI and NEO-FFI scales, Table 4 shows the alpha reliability, mean, and standard deviation for both English and Spanish versions. The alpha reliabilities for the scales in the two instruments and languages ranged from .65 to .84; the alphas for the BFI scales were very similar to those for the monolingual samples in Study 1. As expected, for both instruments, alphas were somewhat higher for the English originals ($M = .84$) than for the Spanish translations (.77). Somewhat surprisingly, the alphas for the NEO-FFI scales ($M = .79$) were not higher than those for the shorter BFI scales (.82). As we had found in Study 1 for the BFI, the Spanish Agreeableness scales on both instruments were the only scales with alphas below .70.

As in Study 1 for the BFI, the means and standard deviations were quite similar for the two languages on both instruments. To test language differences between the Spanish and English versions more formally, we correlated the Big Five scale scores with the language used by the participants; English was coded as 1 and Spanish as 2. Thus, positive correlations indicate that responses given in Spanish were higher than English responses. The correlations for the BFI and NEO-FFI are summarized in Table 5, along with the correlations obtained in Study 1 for the joint culture–language differences between the U.S. and Spanish participants. In Study 2, only 2 of the 10 correlations reached .10, both for the BFI (Agreeableness and Conscientiousness), but they did not replicate for the NEO-FFI. Given the size of these correlations (−.15 and −.24) and their lack of generalizability across instruments, a substantive interpretation should await replication in an independent sample. More generally, then, we did not find strong and consistent differences between the two languages used by the bilingual participants.

Another way to examine the similarity of the response distributions in English and Spanish is to focus on the item level. When we correlated the means obtained for the English items and their Spanish translations across the 44 BFI items, we found high equivalence between the two sets of means ($r = .95$). In samples that include participants of heterogeneous Hispanic national background and generational status, it is important to explore possible subgroup differences. Thus, for each language and instrument, we conducted analyses of variance (ANOVAs) with national background (Mexican vs. non-Mexican) and generational status (U.S.-born vs. immigrant) as between-subjects factors. Generational status had no effect—that is, there were no Big Five differences between U.S.-born and immigrant participants. For national background, the only significant finding involved Openness; across both languages and both instruments, individuals of Mexican background scored lower than non-Mexicans. This effect was apparently due to somewhat lower socioeconomic status (SES; measured by parents' income and education) among the Mexican American participants. When SES was used as a covariate, the differences on Openness disappeared. Overall, then, there was little evidence for Big Five differences among these Hispanic subgroups.

Factor analyses of BFI and NEO-FFI. Using principal-

Table 5
Correlations of Big Five Scale Scores With Culture and Language (Study 1) and With Language (Studies 2 and 3)

	Spanish (vs. English)			
	Study 1	Study 2		Study 3
Scale	BFI	BFI	NEO-FFI	BFI
Extraversion	.12**	−.07	−.05	−.05
Agreeableness	.03	−.24**	−.06	−.06
Conscientiousness	−.09**	−.15**	−.03	−.11
Neuroticism	.08**	.02	.09	.03
Openness	.06	.02	.00	.00

Note. BFI = Big Five Inventory; NEO-FFI = NEO Five Factor Inventory. Positive correlations indicate that responses in Spanish were higher than responses in English.
** $p < .01$.

Table 6
Study 2: Varimax-Rotated Joint Factor Structure for the English and Spanish Big Five Inventory Items in a College-Educated Hispanic Bilingual Sample

Items	E	A	C	N	O
Extraversion					
Is outgoing, sociable	**74**	09	04	−05	30
Es extrovertido, sociable	**80**	10	−07	00	17
Is talkative	**78**	01	−06	00	22
Es bien hablador	**72**	00	−13	05	15
Has an assertive personality	**38**	−11	**35**	−08	23
Es asertivo	**41**	−13	**34**	−20	24
Generates a lot of enthusiasm	**66**	20	11	−10	**31**
Irradia entusiasmo	**58**	21	11	−15	**32**
Is full of energy	**58**	20	08	−02	25
Está lleno de energía	**57**	07	03	−08	13
Is reserved	**−72**	05	07	16	−09
Es reservado	**−66**	07	20	05	−04
Is sometimes shy, inhibited	**−69**	02	−07	24	01
Es a veces tímido, inhibido	**−63**	08	−06	28	−00
Tends to be quiet	**−82**	02	06	12	−12
Tiende a ser callado	**−76**	10	03	11	−11
English loadings absolute mean	**67**	08	10	09	19
Spanish loadings absolute mean	**64**	09	12	11	14
Agreeableness					
Is considerate and kind to almost everyone	01	**64**	21	01	13
Es considerado y amable con casi todos	08	**72**	04	−05	06
Likes to cooperate with others	15	**57**	09	00	22
Le gusta cooperar con los demás	16	**54**	10	02	−06
Is helpful and unselfish with others	07	**71**	17	09	03
Es generoso y ayuda a los demás	04	**53**	20	04	06
Has a forgiving nature	−02	**49**	00	00	20
Es indulgente, no le cuesta perdonar	−08	**27**	−11	02	06
Is generally trusting	16	**38**	10	23	05
Es generalmente confiado	13	**31**	−09	15	15
Tends to find fault with others	00	**−56**	−04	13	12
Tiende a ser criticón	12	**−38**	−04	16	15
Starts quarrels with others	17	**−63**	−19	01	−04
Inicia disputas con los demás	18	**−41**	−19	08	00
Can be cold and aloof	−21	**−48**	−14	17	11
Es a veces frío y distante	**−30**	**−41**	−02	23	05
Is sometimes rude to others	17	**−67**	−22	16	03
Es a veces maleducado con los demás	10	**−52**	−25	12	01
English loadings absolute mean	10	**57**	12	08	10
Spanish loadings absolute mean	13	**45**	11	09	06
Conscientiousness					
Does a thorough job	−03	18	**69**	04	23
Es minucioso en el trabajo	−08	06	**47**	08	21
Does things efficiently	−05	−01	**69**	03	20
Hace las cosas de manera eficiente	07	06	**52**	08	21
Makes plans and follows through with them	06	16	**67**	05	10
Hace planes y los sigue cuidadosamente	05	13	**62**	−04	03
Is a reliable worker	−06	19	**64**	16	09
Es un trabajador cumplidor, de confianza	07	28	**48**	13	−02
Perseveres until the task is finished	−09	15	**74**	−01	16
Persevera hasta terminar el trabajo	−03	17	**61**	00	16
Is easily distracted	16	−05	**−59**	17	00
Se distrae con facilidad	−03	−03	**−54**	25	03
Can be somewhat careless	04	−18	**−62**	22	13
Puede a veces ser algo descuidado	−09	14	**−49**	02	19
Tends to be lazy	00	−13	**−63**	03	−02
Tiende a ser flojo, vago	−03	−20	**−63**	02	02
Tends to be disorganized	00	00	**−64**	06	21
Tiende a ser desorganizado	−04	−06	**−57**	04	**34**
English loadings absolute mean	05	11	**66**	08	12
Spanish loadings absolute mean	05	12	**55**	07	13

(*table continues*)

Table 6 (*continued*)

Items	E	A	C	N	O
Neuroticism					
Can be moody	−10	**−39**	−15	**54**	03
Es temperamental, de humor cambiante	01	**−32**	−16	**60**	02
Is depressed, blue	−29	−11	−10	**50**	00
Es depresivo, melancólico	−18	−20	−12	**56**	−03
Gets nervous easily	−25	11	−01	**75**	−06
Se pone nervioso con facilidad	−17	08	−01	**76**	−04
Can be tense	−20	03	−02	**71**	05
Con frecuencia se pone tenso	−26	05	11	**75**	−03
Worries a lot	−19	10	07	**64**	−14
Se preocupa mucho por las cosas	−15	08	17	**54**	07
Remains calm in tense situations	−07	02	16	**−47**	22
Mantiene la calma en situaciones dificiles	−11	07	05	**−59**	20
Is emotionally stable, not easily upset	09	14	14	**−71**	05
Es emocionalmente estable, dificil de alterar	−02	12	01	**−57**	03
Is relaxed, handles stress well	06	−04	−04	**−72**	10
Es calmado, controla bien el estrés	−11	02	−14	**−65**	07
English loadings absolute mean	15	11	08	**63**	08
Spanish loadings absolute mean	12	11	09	**63**	06
Openness					
Is inventive	18	−03	13	−09	**70**
Es inventivo	00	−05	05	−09	**63**
Is original, comes up with new ideas	25	−01	04	−11	**70**
Es original, se le ocurren ideas nuevas	03	−08	08	−24	**53**
Values artistic, aesthetic experiences	15	20	−05	07	**63**
Valora lo artístico, lo estético	04	28	04	00	**62**
Has an active imagination	17	−09	05	−08	**65**
Tiene una imaginación activa	16	03	10	−10	**62**
Likes to reflect, play with ideas	09	11	07	−01	**64**
Le gusta reflexionar, jugar con las ideas	00	02	−14	−06	**47**
Is sophisticated in art, music, or literature	08	10	−01	10	**65**
Es educado en arte, música, o literatura	07	10	10	10	**56**
Is ingenious, a deep thinker	14	06	−06	04	**62**
Es ingenioso, analítico	06	−11	01	−06	**58**
Is curious about many different things	10	−08	04	−09	**57**
Tiene intereses muy diversos	12	−01	−04	−16	**53**
Prefers work that is routine	**−30**	03	−13	00	**−48**
Prefiere trabajos que son rutinarios	−25	05	−16	−16	**−46**
Has few artistic interests	−04	−16	03	02	**−54**
Tiene pocos intereses artísticos	00	−03	−06	06	**−58**
English loadings absolute mean	07	08	06	06	**62**
Spanish loadings absolute mean	15	07	07	10	**56**

Note. $N = 170$ English–Spanish bilingual individuals. All loadings were multiplied by 100; loadings |.30| or larger are set in bold. Spanish items, their intended factor loadings, and their absolute means are underlined. E = Extraversion; A = Agreeablenss; C = Conscientiousness; N = Neuroticism; O = Openness.

components analysis, we next examined the factorial structure of the final version of the Spanish BFI and compared it with the English version. Because the same participants completed both language versions, it was possible to perform one joint principal-components analysis to test whether language-specific factors might emerge, whether this Hispanic sample would respond to the English BFI items differently than previous Anglo American samples had, and whether any of the Spanish-translated items would load differently than their English originals. The eigenvalues indicated a clear break after the fifth factor. The varimax-rotated loadings of the Spanish and English items is presented in Table 6. Note that the pattern of loadings shows impressive evidence of simple structure. Without exception, every single item (including the revised Openness item from Study 1) had its highest loading on the intended factor, and of the 352 cross-loadings, only 10 (3%) reached .30, and none reached .40. This simple structure is also apparent in the absolute means summarizing the loadings of the items on each factor. The means of the expected loadings all exceeded .45; in contrast, none of the means for the cross-loadings reached .20.[7] These mean values also indicate that, on average, the expected loadings were only slightly higher for the English items than for the Spanish translations.

[7] Although the factor structure presented in Table 6 showed very few item-level departures from simple structure, two of them also appeared in Study 1 with the two monolingual samples. The Spanish translation for the Agreeableness item "Can be cold and aloof" had a negative secondary loading on the Extraversion factor, and the Neuroticism item "Can be moody" had a negative secondary loadings on Agreeableness. The other item-level departures were unique to the present sample and should thus be interpreted cautiously until replicated.

Table 7
Comparing the Big Five Inventory (BFI) and NEO Five Factor Inventory (NEO-FFI) in Study 2: Mean Factor Loadings by Big Five and Language, and Convergent Correlations Across Languages and Instruments

	Mean factor loadings[a]				Convergence correlations			
	English		Spanish		English–Spanish		BFI/ NEO-FFI	
Scale	BFI	NEO-FFI	BFI	NEO-FFI	BFI	NEO-FFI	Eng	Spa
Extraversion	.67	.50	.64	.42	.85	.86	.77	.64
Agreeableness	.57	.44	.45	.40	.79	.86	.75	.68
Conscientiousness	.66	.61	.55	.57	.84	.90	.86	.81
Neuroticism	.63	.62	.63	.52	.84	.88	.80	.75
Openness	.62	.50	.56	.47	.86	.88	.69	.67
M	.63	.53	.57	.48	.84	.88	.77	.71

Note. N = 170 English–Spanish bilingual individuals. Eng = English; Spa = Spanish.
[a] Obtained in two joint principal-components analyses; one included the English and Spanish BFI items, and the other included the English and Spanish NEO-FFI items.

How do these values compare with those for the NEO-FFI? We conducted a joint principal-components analysis of the English and Spanish NEO-FFI items, and the findings are summarized in Table 7. As shown there, the loadings of the targeted items averaged .63 for the BFI and .53 for the NEO-FFI in English, and .57 for the BFI and .48 for the NEO-FFI in Spanish. As in the alpha reliability analyses, the Agreeableness factors were least well-defined, and again this was true for both instruments.

In summary, the pattern of loadings depicted in Table 6 provides strong evidence of the structural similarity between the English and Spanish versions of the BFI in this bilingual Hispanic sample, and Table 7 shows that these values compare favorably with those for the NEO-FFI.

Cross-language convergence for the BFI and the NEO-FFI. The correlations between the English and Spanish scales are shown in Table 7, both for the BFI and for the NEO-FFI. These cross-language convergent validity correlations were all impressively high, with a slightly higher mean for the NEO-FFI scales (.88) than the BFI scales (.84). All of these convergent correlations were significantly larger than the cross-language off-diagonal discriminant correlations, which had a slightly lower mean for the BFI (.14) than for the NEO-FFI (.18).

To further evaluate the convergent correlations in Table 7, we compared them with the alpha reliabilities in Table 4, which reflect the consistency of responses to the same-language items. We found that the convergent correlations were higher than the values expected from the scale reliabilities in the two languages. In other words, if corrected for unreliability, all the cross-language convergence correlations would approach 1.0, indicating that in terms of true scores, the two language versions would lead to the same rank order of individuals.

We also examined the cross-language convergence correlations at the level of the individual items; for the BFI they ranged from .41 to .78 and averaged .60, and for the NEO-FFI they ranged from .19 to .91 and averaged .61.[8] Thus, even at the level of the individual item-translation pairs, average cross-language convergence was substantial and similar for the two instruments.

Finally, we examined cross-language convergence by running separate factor analyses for the English and Spanish versions of the BFI and computing factor congruence coefficients. These coefficients showed the same pattern of results as the other indices of convergence in Table 7. As expected from the smaller sample size in this study of bilinguals, the congruence coefficients were somewhat lower (M = .87) than in Study 1, in which the much more sizable samples yielded more stable and thus congruent factor solutions (M = .94).

Cross-instrument convergence. Table 7 also presents the convergence between the BFI and the NEO-FFI scales, separately for the English and the Spanish language versions. These cross-instrument validity correlations were quite impressive, averaging .71 even for the two translated Spanish versions. As expected, validity correlations were slightly higher for the two English instruments (mean r = .77). These convergence correlations compare favorably to those reported for convergence between adjective scales and the much longer NEO-PI, which averaged .61 in an Anglo American sample (Goldberg, 1992). Moreover, the convergent validity correlations in Table 7 contrast with rather small off-diagonal correlations, which had means of .18 in English and .17 in Spanish, thus providing strong evidence of both convergent and discriminant validity across both languages.[9]

[8] Eight of the NEO-FFI items had cross-language convergence correlations below .40, suggesting that the translations of these items might benefit from revision.

[9] As suggested by a reviewer, we also examined cross-instrument convergence by conducting joint factor analyses of the BFI and NEO-FFI items in English, in Spanish, and in both languages combined. These analyses are not central to the present article but are available from the authors. To summarize briefly, the pattern of item loadings on the factors showed evidence of convergence and discrimination, just as the scale-level findings in Table 7. For example, in the joint English BFI/NEO-FFI factor structure, all 44 BFI items loaded above .30 on the correct factors and 54 of the 60 NEO-FFI items did. In the joint Spanish BFI/NEO-FFI structure, 42 of the 44 BFI items loaded above .30 on the correct factor, as did 50 of the 60 NEO-FFI items. In short, these analyses

Confirmatory factor analysis of the multitrait multimethod matrix. So far, our findings provide strong evidence for the cross-language and cross-instrument validity of the Spanish BFI scales in this Hispanic sample. Together with our findings for the NEO-FFI, our results also suggest that the Big Five serve as the major source of variance in these two instruments in both languages. More formal tests were conducted using CFA. Because this study included two different instruments, each administered in two different languages, we were able to test hypotheses both regarding substantive Big Five effects and regarding two sources of method variance: language-specific effects and instrument-specific effects. Specifically, structural models representing five personality factors plus instrument and language factors were tested and statistically compared using the chi-square difference test. For each of these models, loadings on a particular factor were freely estimated or constrained to 0 on the basis of the particular personality factor, language, or instrument represented. Analyses were performed on scale scores; as shown in Figure 1, each latent Big Five factor was represented by four indicator scales: English BFI, Spanish BFI, English NEO-FFI, and Spanish NEO-FFI. Descriptions of these models and the associated fit indices are given in Table 8.

The first model we tested specified five uncorrelated latent factors representing the Big Five, and the second model five correlated latent factors; neither model included language or instrument method factors. Table 8 shows that both models yielded fit indices that were moderately satisfactory but slightly below what is usually considered acceptable (i.e., comparative fit indices were smaller than .90 and χ^2/df indices were larger than 3). As expected, allowing for correlated Big Five factors (Model 2) resulted in a significant increase in overall fit, $\Delta\chi^2(10, N = 170) = 57.8, p < .05$, thus replicating our CFA findings in Study 1. Model 3, which specified two uncorrelated language factors (English and Spanish) in addition to the five correlated personality factors, resulted in a significant increase in overall fit, $\Delta\chi^2(20, N = 170) = 58.1, p < .05$, but the fit indices were still unacceptable. Interestingly, the standardized solution for this model showed that the two hypothesized language factors were not interpretable as language factors. On both factors, only two loadings were significant, and these loadings showed that there were no language effects shared across either instruments or scales.[10] We next tested the second possible source of method effects, namely, instrument-specific factors. As shown in Table 8, Model 4 represented five correlated Big Five factors plus two uncorrelated instrument factors, one for the BFI scales and one for the NEO-FFI scales. This model had acceptable fit indices and significantly improved overall fit, $\Delta\chi^2(0, N = 170) = 127.8, p < .05$, suggesting that this model provides a better fit than any of the other three models. A path diagram of this model with the complete set of parameters from the standardized solution is depicted in Figure 1.[11]

The five circles in the middle of Figure 1 represent the latent Big Five factors. Each factor influences four measured variables, represented by boxes: the English and Spanish BFI scales on the left, and the English and Spanish NEO-FFI scales on the right. Right next to each box is a lowercase *e*, which represents the influence of error on the measured variable. The BFI method factor on the left-hand side of the figure was found to influence only the BFI Openness scales and, to a lesser extent, the Extraversion scales in both English and Spanish. This method factor thus captures primarily Openness and Extraversion variance that is shared by these BFI scales but not shared with the NEO-FFI and thus cannot be accounted for by the latent Big Five factors that are defined by variance shared across both instruments. Similarly, the NEO-FFI method factor on the right-hand side of the figure was found to influence only the Extraversion and Openness scales and thus represents method variance shared by these two NEO-FFI scales across the two languages. In short, the BFI and NEO-FFI instrument factors are not general factors (i.e., method variance shared by all the traits in the instrument). Rather, these two factors represent instrument-specific trait variance shared by both the English and Spanish versions of the same instrument. This finding is consistent with earlier findings showing that the Extraversion and Openness dimensions are defined somewhat differently on the two instruments.

The parameter estimates for Model 4 in Figure 1 suggest three major conclusions that are consistent with the preceding analyses. First, all 20 scales had substantial loadings on the five latent factors, ranging from a low of .70 to a high of .95. On average, the English-language scales had only slightly higher loadings ($M = .90$) than the Spanish scales ($M = .86$). Second, the substantial size of these loadings did not leave much systematic variance for general language or instrument factors. Instead, the two latent method factors we did uncover related to the two traits (Openness and Extraversion) that showed substantially higher convergence within each instrument across languages than across instruments within each language (see Table 7). In other words, these scales generalized across languages but in ways that differed somewhat across the two instruments. The two method factors capture these instrument-specific sources of variance that are shared across the two languages for the Open-

show much the same picture as do the other correlational and structural-equation analyses at the scale level reported in the text.

[10] The factor specified as an English language factor did not show loadings that could be attributed to a common English-language effect across instruments and scales. The two significant loadings were a positive loading for the English NEO-FFI Extraversion scale and a negative loading for the English BFI Extraversion scale. The Spanish language factor was defined only by BFI scales and included a negative loading for the Spanish BFI Extraversion scale and a positive loading for the Spanish BFI Neuroticism scale.

[11] We also tested a model that specified only trait factors and represented method variance by allowing within-BFI and within-NEO-FFI errors (also labeled unique variances) to intercorrelate. This approach, known as the *correlated uniqueness model* (Bagozzi, 1993; Kenny & Kashy, 1992; Panter, Tanaka, & Hoyle, 1994) yielded a good fit for the data but proved less informative than the models summarized in Table 8. Specifically, because method factors are not specified a priori, evidence of shared method variance is contained only in the 20×20 matrix of intercorrelations among the errors. A close examination of these correlations revealed (a) no evidence for general language effects, (b) no evidence for general instrument effects, and (c) a pattern of significant correlations among the errors for the Extraversion and Openness to Experience BFI scales and among the errors for the Extraversion and Openness to Experience NEO-FFI scales—a pattern consistent with the two method factors specified in Model 4 (Table 8).

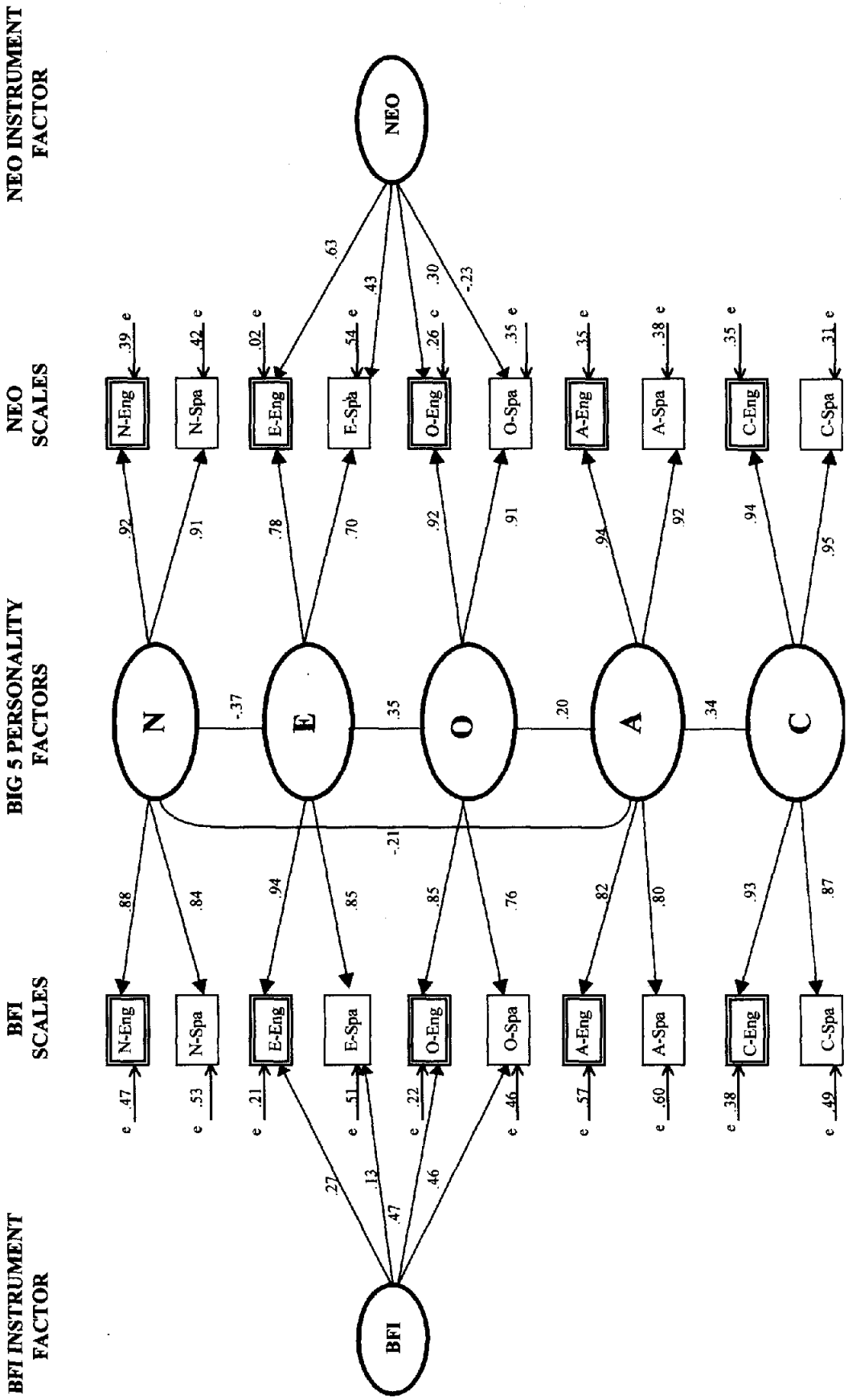
BFI INSTRUMENT FACTOR
BFI SCALES
BIG 5 PERSONALITY FACTORS
NEO SCALES
NEO INSTRUMENT FACTOR
BFI
NEO
N
E
O
A
C
N-Eng
N-Spa
E-Eng
E-Spa
O-Eng
O-Spa
A-Eng
A-Spa
C-Eng
C-Spa
.88
.84
.94
.85
.85
.76
.82
.80
.93
.87
.92
.91
.78
.70
.92
.91
.94
.92
.94
.95
.63
.43
.30
-.23
-.37
.35
.20
.34
-.21
.27
.13
.47
.46
.47
.53
.21
.51
.22
.46
.57
.60
.38
.49
.39
.42
.02
.54
.26
.35
.35
.38
.35
.31

Table 8
Study 2: Summary of Goodness of Fit for the Multitrait Multi-Instrument Multilanguage Data

Model	χ^2	*df*	χ^2/df	CFI	$\Delta\chi^2$
1. Uncorrelated Big Five personality factors, no method factors	581.5	170	3.4	.85	—
2. Correlated Big Five personality factors, no method factors	523.7	160	3.3	.87	57.8*
3. Correlated Big Five personality factors, uncorrelated English and Spanish language factors	465.6	140	3.3	.88	58.1*
4. Correlated Big Five personality factors, uncorrelated BFI and NEO-FFI instrument factors	337.8	140	2.4	.93	127.8*

Note. $N = 170$ English–Spanish bilingual individuals. CFI = comparative fit index (Bentler, 1990). $\Delta\chi^2$ = increase in overall fit. BFI = Big Five Inventory; NEO-FFI = NEO Five Factor Inventory.
$^*p < .05$.

ness and Extraversion dimensions. Nonetheless, in all cases, the loadings on these method factors were considerably smaller than the substantive trait loadings. The third conclusion involves the size of the intercorrelations among the latent Big Five dimensions. As in Study 1, the intercorrelations remained low even when disattenuated for unreliability by CFA; only four correlations exceeded .20, and none reached .40.

Overall, then, the CFA results suggest that five latent, correlated personality factors representing the Big Five structure capture the major sources of variance in our multitrait multilanguage multi-instrument design. In addition to these five substantive factors, two instrument factors (BFI and NEO-FFI) representing trait-specific instrument variance emerged as significant sources of variance. We found no evidence for English and Spanish language factors that would account for variance shared across the Spanish BFI and NEO-FFI scales or across the English BFI and NEO-FFI scales.

General comparison of the BFI and the NEO-FFI. How do these two relatively short Spanish measures of the Big Five compare? The BFI's reliability and factor structure were at least equal to those of the NEO-FFI. In terms of convergence and discrimination across the two languages, the two instruments were again quite similar. However, the BFI takes only about 5 min of administration time (compared with about 15 min for the NEO-FFI) because the BFI scales use fewer items than the NEO-FFI scales (9 vs. 12), and the items are about half as long (*Ms* = 5.1 vs. 9.7 words). Thus, the BFI showed much greater economy (Burisch, 1984). Moreover, the BFI items seemed easier to understand: On each of three indicators (Microsoft, 1993), the BFI required a lower mean reading grade level (M = 5.2) than the NEO-FFI (M = 6.6).[12] We also asked four Hispanic bilingual individuals (who were blind to the purpose of the study) to rate how easy it was to understand each of the items in the two Spanish versions. Items from the two instruments were presented together in a random order, and ratings were made following the procedures devised by Angleitner, John, and Lohr (1986, pp. 85–89). The BFI items were easier to understand, $t(102) = 2.9$, $p < .01$; using Angleitner et al.'s classification rules, 20% of the NEO-FFI items were not immediately understandable, compared to 7% of the BFI items. Despite the BFI/NEO-FFI difference, these findings indicate that both instruments compare favorably with older questionnaires, which on average contained more than 50% items judged not immediately understandable (Angleitner et al., 1986).

External (or predictive) validity was not examined in the present studies, and future research is needed to address this limitation. However, the substantial validity of the English "parent" instruments (e.g., with respect to peer ratings) and the high English–Spanish convergence lead us to be optimistic that validity findings for the BFI and the NEO-FFI will generalize to the Spanish versions of both instruments. At the very least, there is no evidence to suggest otherwise, as the English and Spanish versions did not differ in the bilingual sample of Study 2.

When should researchers use the Spanish BFI and when should they use the NEO PI-R or NEO-FFI? When participant time is not at a premium, participants are well-educated and test-savvy, and the research question calls for the assessment of multiple facets for each of the Big Five dimensions, then the

[12] The NEO PI-R manual (Costa & McCrae, 1992) reports a sixth-grade reading level (p. 4). The three indices used here were the Flesch–Kincaid grade level, Coleman–Liau grade level, and Bormuth grade level, which use word length (number of characters), average number of syllables per word, and sentence length to determine a grade level. Standard writing corresponds to seventh- or eighth-grade level (see readability statistics in Word 6.0).

Figure 1 (opposite). Model 4 for confirmatory factor analysis of multitrait multimethod matrix in Study 2: Five latent personality factors (N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness) plus one Big Five Inventory (BFI) instrument factor and one NEO Five Factor Inventory instrument factor. Eng = English; Spa = Spanish.

full 240-item Spanish NEO PI-R would be most useful. Otherwise, the 44-item BFI (which is given in the Appendix) seems to offer a more efficient and easily understood measure of the Big Five in Spanish-speaking samples than does the 60-item NEO-FFI.

Study 3: Partial Replication in a Hispanic Noncollege Sample

The findings from Study 2 suggest that the revised Spanish BFI scales have respectable psychometric characteristics and substantial cross-language and cross-instrument validity in a bilingual Hispanic sample. One limitation of this study is that it relied on college-educated participants, leaving the generality of our results to less educated samples unexplored. Study 3 was designed to address this limitation; this follow-up study was a partial replication of Study 2 in a working-class sample of Hispanics and focused on the BFI. We examined the reliability and cross-language convergence of the Spanish and English BFI scales; in particular, we asked whether the Spanish BFI scales could be used in adult Hispanic samples with less formal education than typically found in college samples. We thus expected somewhat lower alpha reliabilities and cross-language convergence correlations than in Studies 1 and 2. The main question was whether under these circumstances the psychometric characteristics of the Spanish BFI scales would still be acceptable.

Method

Participants and procedures. This sample included 139 Hispanic adults (54 men and 85 women). Their mean age was 32 years (SD = 13). All participants were either immigrants from Latin American countries or U.S.-born descendants of Latin American individuals. Participants had backgrounds from a wide range of Latin American cultures including Mexico (n = 69), El Salvador (33), Nicaragua (11), Colombia (5), Peru (4), Chile (3), Panama (3), Argentina (3), Cuba (2), Puerto Rico (2), Venezuela (1), Honduras (1), Ecuador (1), and Guatemala (1).

Participants were recruited in a working-class neighborhood in San Francisco, during a 2-day street festival organized for the *Cinco de Mayo* celebration. Participants either approached the experimenter and her assistants (all of whom were Hispanic) to inquire about their poster inviting Spanish–English bilinguals for a study or were approached by the experimenter and her assistants. Those who had not attended college were given the questionnaire materials and asked to complete them in a designated private area next to the information stand for this study.

Instruments. Participants filled out both the English and the final Spanish versions of the BFI; half of them completed the English version first. Participants received a symbolic monetary compensation ($1) for their participation. Given the fears and concerns that recent changes in immigration laws and regulations have triggered among Hispanic immigrants in California, we made an effort to encourage participation by fully assuring participants of their anonymity; thus, the questionnaire included no background questions other than age, sex, and national background.

Results and Discussion

For each of the BFI scales, Table 9 shows the number of items, alpha reliability, mean, and standard deviation for both English and Spanish versions. As can be seen in Table 9, the alpha reliabilities for the scales were all adequate in both languages, ranging from .73 to .80 for the English BFI (m = .78), and from .69 to .77 for the Spanish BFI (m = .74). As expected, these alphas were lower than those in the college samples of Study 1 and 2 (which are also given in Table 9 for comparison purposes), but the size of the difference was not large.

As in Studies 1 and 2, the scale means and standard deviations were again very similar for the two languages, with none of the values being more than one decimal apart. None of the correlations with language used (Spanish vs. English) were significant; they ranged from −.11 to .03 and are given in Table 5. Taken together, the findings from Studies 2 and 3 suggest that language had no consistent effects on bilingual individuals' Big Five scores.

Note also that the means in this study were quite similar to those from the two previous studies. Again, the equivalence of the response distributions in English and Spanish was also apparent at the item level; the correlation between the English and the Spanish item means across the 44 items was .94, similar to the .95 in Study 2. Even across the two studies, the equivalence correlations were all above .91.

As in Study 2, we also explored subsample differences comparing participants of Mexican, Salvadorian, and "other Hispanic" background. For each BFI scale and language, we conducted a one-factorial ANOVA, with the three national backgrounds forming a between-subjects factor. Results indicated no significant differences among the subsamples on any of the BFI scales in either language. Thus, the Openness differences between Mexican and non-Mexican Hispanics from Study 2 failed to replicate.

Table 9 also reports the correlations between the Spanish and English versions of the BFI. These cross-language convergent validity correlations were substantial in size, averaging .65, but lower than the mean of .84 observed in the bilingual college sample of Study 2. There was also evidence for discriminant validity. Only 3 of the 20 off-diagonal correlations exceeded .30, and none of them reached .40 or even approached the magnitude of the convergent validity correlations.

It is interesting to compare the findings for Agreeableness in Study 3 to those in Studies 1 and 2. In the first two studies, the Spanish Big Five scales generally showed levels of alpha reliability and factorial coherence that were only slightly lower than those observed for the original English scales (see Table 9). However, for the Spanish Agreeableness scale, the alphas were generally lower, and this pattern was consistent across the BFI and NEO-FFI and across both the monolingual and the bilingual samples of college students, with alphas ranging from .65 to .69. However, in Study 3, the alpha was .75, a value no different from the other Big Five scales. One possible reason for this discrepancy is that the college samples showed consistently lower standard deviations (mean SD = 0.5) on the Agreeableness scale than did the adults of Study 3 (mean SD = 0.7). Future research should investigate the reasons for this restriction of range in college students' self-reports, such as potential age differences in the social desirability of this construct.

Several limitations of the sample and procedures in this study should be considered in evaluating these findings. Because of the constraints on our participants' time and availability, bilingualism could not be assessed directly as we had in Study 2. As a consequence, some of the participants were probably not

Table 9
Study 3: Psychometric Characteristics and Cross-Language Convergence Correlations for the Spanish and English Big Five Inventory Scales in a Noncollege Hispanic Bilingual Sample

Scale	*n*	α		*M*		*SD*		English–Spanish convergence
		Eng	Spa	Eng	Spa	Eng	Spa	
Extraversion	8	.73	.69	3.6	3.5	0.7	0.7	.69
Agreeableness	9	.78	.75	3.9	3.8	0.7	0.7	.54
Conscientiousness	9	.80	.74	3.9	3.8	0.7	0.6	.61
Neuroticism	8	.80	.75	2.6	2.6	0.8	0.8	.66
Openness	10	.80	.77	3.9	3.9	0.7	0.6	.72
Means of the three studies								
Study 3	9	.78	.74	3.6	3.5	0.7	0.7	.65
Study 2	9	.85	.78	3.6	3.5	0.7	0.7	.84
Study 1	9	.83	.78	3.5	3.5	0.7	0.7	—

Note. N = 136 Spanish–English bilingual individuals; n = number of items in the scale; Eng = English; Spa = Spanish. Means for Studies 2 and 3 were taken from Tables 1 and 4.

truly bilingual (i.e., not fluent in English as well as Spanish). In fact, given that our sample included many recent immigrants with relatively low SES and little formal education in the United States, it is likely that some participants experienced difficulty in responding to the English BFI items, thus lowering the cross-language convergence correlations that could be observed in this study. Another important limitation is the unstructured method of data collection—the clamorous nature of the street festival may have hampered participants' ability to carefully read and respond to the instruments. Given these limitations, the cross-language convergences and alpha reliabilities found here probably represent underestimates of the real effect sizes. Overall, then, the findings of Study 3 show that the Spanish BFI scales can be used successfully in Hispanic minority samples with less formal education than typically found in college samples.

One potential limitation of the bilingual designs used in both Studies 2 and 3 is that results from these samples may not be generalizable to monolingual Hispanic samples (i.e., immigrants who do not speak any English). Bilingual (and bicultural) individuals may differ from monolingual individuals in particular cognitive and interpersonal characteristics (Laframboise, Coleman, & Gerton, 1993; Lambert, 1977; Reynolds, 1991). Thus, the mean levels on the Spanish and English BFI and NEO-FFI scales obtained in Studies 2 and 3 may not apply to monolingual Hispanic samples. Further research is needed to examine this issue.

Finally, although the samples in Studies 2 and 3 were reasonably large, they were nonetheless limited in their representation of the many national backgrounds of Hispanics living in the United States. We conducted subsample analyses comparing Mexicans with non-Mexican Hispanics in Study 2, and between Mexican, Salvadorian, and "other" Hispanics in Study 3, but we did not find any consistent differences. Nonetheless, rather than treating all Hispanic participants as a culturally homogeneous group, future research should further explore within-minority differences (Marín & Marín, 1991). In particular, it would be important to test whether the Spanish versions of the BFI and the NEO PI-R can indeed be used with Hispanics from all national backgrounds. Similarly, although we found no differences between immigrant and U.S.-born Hispanics in the present studies, more careful attention should be paid to acculturation differences (Padilla, 1995; Rogler, Cortes, & Malgady, 1991).

General Discussion

We discuss the implications of our three studies for three issues: (a) measurement of the Big Five personality dimensions in Spanish-speaking individuals, (b) the cross-cultural generality of personality, and (c) research methods in cross-cultural work.

Measuring the Big Five in Spanish

One of our goals was to add to the instruments available in Spanish to encourage more personality research on Hispanic minority populations in the United States. Of particular concern were reliability and structural equivalence of the BFI across languages and cultural groups, which we examined in two kinds of research designs: monolingual cross-cultural samples and bilingual Hispanic samples. The cross-cultural design of Study 1 compared two large monolingual college samples and showed that in Spain, the Spanish BFI scales had adequate psychometric characteristics, with alphas, means, and standard deviations very similar to those of the English-language scales in the U.S. sample. They also showed substantial structural similarity to the English scales in the multisample CFA and convergent and discriminant validity with indigenous Spanish Big Five markers.

The cross-cultural design of Study 1 simultaneously varied both the language and culture group of the respondents. In contrast, the bilingual design of Studies 2 and 3 allowed us to compare the Spanish and English versions within the same U.S. sample of Hispanic individuals. These studies again provided support for cross-language convergent and discriminant validity of the BFI, at both the scale and the item level, and showed that the results from Study 1 generalize to non-European Spanish-speaking samples—both Hispanic college students (Study 2)

and working-class Hispanic adults (Study 3). Study 2 also showed excellent cross-instrument convergence and discrimination with the NEO-FFI, a short form of the 240-item NEO PI-R, in both English and Spanish. This multitrait multilanguage multi-instrument design provided explicit tests of language- and instrument-specific sources of method variance. CFA results showed that (a) both BFI and NEO-FFI scales loaded as expected on five substantive factors, (b) there was no evidence of general language effects, and (c) there was some evidence of instrument-specific factors, involving small but consistent differences in the ways Extraversion and Openness are defined in each of the two instruments. These findings are reassuring from the perspective of instrument equivalence, and they suggest that the Spanish BFI can serve as a useful personality assessment tool for research on various Spanish-speaking populations. Further research is needed to examine the external validity of these new BFI scales, however, before the unrestricted use of the BFI with Spanish-speaking respondents can be advocated.

Cross-Cultural Specificity and Generality of Personality

Our findings also have theoretical implications for research in cross-cultural psychology. The findings from our three studies are surprisingly easy to summarize: Whether we compared samples from Spain and the United States, college students and working-class adults, or Spanish- and English-language versions, there were no consistent differences in factor structures, alphas, and norms. That is, Latin–Anglo differences in such cultural values as collectivism, *simpatía,* and time orientation did not result in systematic differences in the ways the Big Five personality traits covaried in individuals from Latin and Anglo cultural groups.

Consider the culture-level concept of *simpatía* as an example. The college student samples of Studies 1 and 2 showed somewhat lower alpha coefficients for Agreeableness in Spanish and two secondary loadings of Agreeableness items on Extraversion, suggesting that these *simpatía*-related traits might be structured somewhat differently in Spanish. However, Study 3 showed that these differences did not hold in an adult working-class sample, suggesting that age or educational differences may be more important. In terms of mean differences, one might expect that *simpatía* would lead to higher Agreeableness scores in Spanish-speaking samples. However, individuals from Latin cultures did not score higher than Anglo Americans. This is not an isolated finding; Jensen-Campbell, Graziano, and Hair (1996) also failed to obtain mean differences in Agreeableness among 45 Mexican American and 98 Anglo American adolescents.

We also did not find consistent language differences. In Study 2, in which we ensured that participants were truly fluent in both languages, the cross-language validity correlations for the scales (see Table 7) were higher than would be expected from the reliabilities of the scales in the two languages; in fact, the CFA model including separate language factors as method effects failed to fit the data. Moreover, in Studies 2 and 3, in which we used bilingual participants from the same culture, we did not find consistent mean differences between the English and Spanish versions.

How should we interpret these "null" findings, which replicated across our samples and studies? At first glance, these findings would seem to conflict with studies that have shown language effects in bilingual designs (e.g., Ervin, 1964; Marín et al., 1983). However, we think there are three reasons for this apparent difference in findings. First, as Marín et al. observed, "Most of the discrepancies between the answers in English and Spanish were found on those emic items concerned with the meaning of concepts . . . for example, how a Hispanic can show respect to another Hispanic" (p. 181). In contrast, items on most personality trait questionnaires do not focus on culturally specific concepts or meaning systems; instead, they ask about fairly general behavioral and emotional characteristics that may not elicit substantially different interpretations from Anglo Americans, Hispanics, and Spaniards. The imposed-etic research strategy may further limit the extent to which cultural differences are likely to be manifested. Future research should use emic designs (e.g., Church & Katigbak, 1989; Yang & Bond, 1990) to probe Latin–Anglo differences and to compare the BFI scales with indigenous Spanish personality constructs (see Benet-Martínez & Waller, 1997).

A second reason involves differences in the structural properties of both items and responses. For example, Ervin (1964) found that bilinguals differ in the Thematic Apperception Test stories they tell in their two languages. In contrast, the items and response options on the BFI and NEO-FFI are highly structured and do not invite the culturally diverse interpretations possible when telling a story in response to an ambiguous picture. Third, our items are contextually abstract, that is, they are neutral with respect to the context in which relevant behaviors may be manifested; cultural differences that may well exist in the way specific behaviors are manifested in specific contexts are not likely to play an important role here.[13]

In conclusion, the present findings should not be taken to mean that there are no important Latin–Anglo differences in individual personality. Rather, we suggest that whether a personality study shows cultural generality or specificity will depend on the level of abstraction chosen in conceptualizing personality. Even specific traits, such as talkativeness and forgetfulness, are abstractions that summarize general trends in the behavior and experience of the individual over time and situations. The Big Five dimensions represent an even broader level of abstraction, aggregating across numerous more specific trait domains. Thus, it is possible that at the broad trait level, personality structure is quite general, even universal, across cultural groups. At the same time, personality may be much more culturally specific at lower levels of abstraction, such as for middle-level personality constructs (Cantor & Zirkel, 1990) like personal projects and strivings, life goals, and possible selves. This view allows for both cultural specificity and generality in personality, recognizing the crucial role of the level of abstraction at which personality is conceptualized.

Methodology in Cross-Cultural Research

In the studies reported here, we have broadened the multitrait multimethod approach to construct validation by including lan-

[13] For example, the cultural concept *simpatía* emerged from 600 role-differential judgments made by Hispanics and non-Hispanics (Triandis et al., 1984), thus representing a much greater level of specificity and contextual detail.

guage effects as another method factor. This allowed us to apply various kinds of structural equation models to multitrait data from different cultures and languages. In Study 1, we used multisample CFA to test the structural equivalence of English and Spanish Big Five measures simultaneously in a U.S. sample and in a Spanish sample. In Study 2, in which the same bilingual participants completed both English and Spanish instruments, we used a multitrait multilanguage multi-instrument CFA to estimate both substantive personality factors and two kinds of method effects. Another methodological innovation in this research involves the procedures we developed for screening and assessing bilingualism in Study 2, in particular the translation tests. Such tests should prove useful in studies of translation equivalence because true equivalence will be underestimated unless all participants are known to be truly bilingual.

As Panter, Tanaka, & Hoyle (1994) observed, "Much effort is typically expended in collecting high-quality data from multiple observational modes. . . . However, strategies for analyzing the data that emanate from these designs do not always optimize the available information" (pp. 134–135). This observation tends to apply to the analysis of high-quality cross-cultural data as well. We hope that the procedures applied here, especially the multitrait multilanguage matrix approach, will prove useful for personality research in other languages and cultures.

References

Almagor, M., Tellegen, A., & Waller, N. G. (1995). The Big Seven model: A cross-cultural replication and further exploration of the basic dimensions of natural language trait descriptors. *Journal of Personality and Social Psychology, 69,* 300–307.

Angleitner, A., John, O. P., & Lohr, F. (1986). It's What you ask and How you ask it: An itemmetric analysis of personality questionnaires. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaire: Current issues in theory and measurement* (pp. 61–108). Berlin, Germany: Springer-Verlag.

Bagozzi, R. P. (1993). Assessing construct validity in personality research: Applications to measures of self-esteem. *Journal of Research in Personality, 27,* 49–87.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1–26.

Benet, V., & Waller, N. G. (1995). The "Big Seven" model of personality description: Evidence for its cross-cultural generality in a Spanish sample. *Journal of Personality and Social Psychology, 69,* 701–718.

Benet-Martínez, V. (1997). Theoretical perspectives on the five factor model: Agreement about five factors, disagreement about the concept of trait. *Journal of Personality Assessment, 69,* 658–664.

Benet-Martínez, V., & John, O. P. (1998). *Spanish Big Five scales using indigenous Castillian markers.* Unpublished manuscript, University of Michigan, Ann Arbor.

Benet-Martínez, V., & Waller, N. G. (1997). Further evidence for the cross-cultural generality of the "Big Seven" model: Imported and indigenous Spanish personality constructs. *Journal of Personality, 65,* 567–598.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246.

Bentler, P. M. (1995). *EQS: Structural equations program manual.* Los Angeles: BMDP Statistical Software.

Berry, J. W. (1980). Introduction to methodology. In H. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 1–28). Boston: Allyn & Bacon.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117,* 187–215.

Bond, M. H. (1994). Trait theory and cross-cultural studies of person perception. *Psychological Inquiry, 5,* 114–116.

Borkenau, P., & Ostendorf, F. (1990). Comparing exploratory and confirmatory factor analysis: A study of the five-factor model of personality. *Personality and Individual Differences, 11,* 515–524.

Brislin, R. W. (1980). Translation and content analysis of oral and written materials. In H. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology* (Vol. 2, pp. 389–444). Boston: Allyn & Bacon.

Burisch, M. (1984). Approaches to personality inventory construction. *American Psychologist, 39,* 214–227.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Cantor, N., & Zirkel, S. (1990). Personality, cognition, and purposive behavior. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 135–165). New York: Springer-Verlag.

Church, A. T., & Katigbak, M. S. (1988). The emic strategy in the identification and assessment of personality dimensions in a non-Western culture. *Journal of Cross-Cultural Psychology, 19,* 140–163.

Church, A. T., & Katigbak, M. S. (1989). Internal, external, and self-report structure of personality in a non-Western culture: An investigation of cross-language and cross-cultural generalizability. *Journal of Personality and Social Psychology, 57,* 857–872.

Cialdini, R. B., Trost, M. R., & Newson, J. T. (1995). Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology, 69,* 318–328.

Clark, L. A. (1992). *Schedule for Normal and Abnormal Personality (SNAP).* Minneapolis: University of Minnesota Press.

Costa, P. T., & McCrae, R. R. (1980). Influence of Extraversion and Neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology, 38,* 668–678.

Costa, P. T., & McCrae, R. R. (1985). *The NEO Personality Inventory Manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R. The Revised NEO Personality Inventory.* Odessa, FL: Psychological Assessment Resources.

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

Ervin, S. M. (1964). Language and TAT content in bilinguals. *Journal of Abnormal and Social Psychology, 68,* 500–507.

García Sevilla, L., Pérez, J., & Tobeña, A. (1979). Fiabilidad y validez de la versión castellana del EPI [Validity and reliability of the EPI Spanish Version]. *Revista Latinoamericana de Psicología, 11,* 393–402.

Garza, R. T. (1977). Personal control and fatalism in Chicanos and Anglos: Conceptual and methodological issues. In J. L. Martínez (Ed.), *Chicano psychology* (pp. 97–108). New York: Academic Press.

Gergen, K. J., Gulerce, A., Lock, A., & Misra, G. (1996). Psychological science in cultural context. *American Psychologist, 51,* 496–503.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. *Review of Personality and Social Psychology, 2,* 141–165.

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

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48,* 26–34.

Goldberg, L. R., & Kilkowski, J. M. (1984). The prediction of semantic consistency in self-descriptions: Characteristics of persons and of terms that affect the consistency of responses to synonym and

antonym pairs. *Journal of Personality and Social Psychology, 48,* 82–98.

González Valdes, T. L. (1979). Estudio valorativo de algunas características del MMPI en pacientes con trastornos psíquicos [Assessment of the MMPI characteristics in a psychiatric sample]. *Revista del Hospital Psichiatrico de La Habana, 20,* 249–258.

Gross, J. J., & John, O. P. (1995). Facets of emotional expressivity: Three self-report factors and their correlates. *Personality and Individual Differences, 19,* 555–568.

Hofstede, G. (1983). Dimensions of national cultures in fifty countries and three regions. In J. Deregowski, S. Dzuirawiec, & R. Annis (Eds.), *Explications in cross-cultural psychology* (pp. 335–355). Lisse, The Netherlands: Swets & Zeitlinger.

Hofstee, W. K. (1990). The use of everyday personality language for scientific purposes. *European Journal of Personality, 4,* 77–88.

Hofstee, W. K., Kiers, H. A., de Raad, B., Goldberg, L. R., & Ostendorf, F. (1997). A comparison of Big-Five structures of personality traits in Dutch, English, and German. *European Journal of Personality, 11,* 15–31.

Jensen-Campbell, L. A., Graziano, W. G., & Hair, E. C. (1996). Personality and relationships as moderators of interpersonal conflict in adolescence. *Merrill-Palmer Quarterly, 42,* 148–164.

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-Verlag.

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., Caspi, A., Robins, R. W., Moffitt, T. E., & Stouthamer-Loeber, M. (1994). The "Little Five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development, 65,* 160–178.

John, O. P., & Donahue, E. M. (1998). *The Big Five Inventory: Studies of reliability and validity.* Berkeley: University of California, Berkeley, Institute of Personality and Social Research. Manuscript in preparation.

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., Goldberg, L. R., & Angleitner, A. (1984). Better than the alphabet: Taxonomies of personality-descriptive terms in English, Dutch, and German. In H. Bonarius, G. van Heck, & N. Smid (Eds.), *Personality psychology in Europe: Theoretical and empirical developments* (pp. 83–100). Lisse, The Netherlands: Swets & Zeitlinger.

John, O. P., & Robins, R. W. (1993). Gordon Allport: Father and critic of the five-factor model. In K. H. Craik, R. Hogan, & R. N. Wolfe (Eds.), *Fifty years of personality psychology* (pp. 215–236). New York: Plenum.

John, O. P., & Robins, R. W. (1994). Trait and types, dynamics and development: No doors should be closed in the study of personality. *Psychological Inquiry, 5,* 137–142.

Johnson, J. A., & Ostendorf, F. (1993). Clarification of the five-factor model with the abridged Big Five dimensional circumplex. *Journal of Personality and Social Psychology, 65,* 563–576.

Johnson, S., & Wolfe, R. (1995). Personality as a predictor of college performance. *Educational and Psychological Measurement, 55,* 177–185.

Katigbak, M. S., Church, A. T., & Akamine, T. X. (1996). Cross-cultural generalizability of personality dimensions: Relating indigenous and imported constructs in two cultures. *Journal of Personality and Social Psychology, 70,* 99–114.

Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait–multimethod matrix by confirmatory factor analysis. *Psychological Bulletin, 112,* 165–172.

Laframboise, T., Coleman, H. L. K., & Gerton, J. (1993). Psychological impact of biculturalism: Evidence and theory. *Psychological Bulletin, 114,* 395–412.

Lambert, W. E. (1977). The effects of bilingualism on the individual: Cognitive and sociocultural consequences. In P. A. Hornby (Ed.), *Bilingualism: Psychological, social, and educational implications* (pp. 15–28). San Diego, CA: Academic Press.

Larsen, R. J., & Diener, E. (1992). Promises and problems with the circumplex model of emotion. In M. S. Clark (Ed.), *Review of Personality and Social psychology* (Vol. 13, pp. 25–59). Newbury Park, CA: Sage.

Marín, G., & Marín, B. (1991). *Research with Hispanic populations.* Newbury Park, CA: Sage.

Marín, G., Triandis, H. C., Betancourt, H., & Kashima, Y. (1983). Ethnic affirmation versus social desirability: Explaining discrepancies in bilinguals' responses to a questionnaire. *Journal of Cross-Cultural Psychology, 14,* 173–186.

McAdams, D. P. (1992). The five-factor model in personality: A critical appraisal. *Journal of Personality, 60,* 175–215.

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

McCrae, R. R., & Costa, P. T. (1997a). Personality trait structure as a human universal. *American Psychologist, 52,* 509–516.

McCrae, R. R., & Costa, P. T. (1997b). Conceptions and correlates of Openness to Experience. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 826–842). San Diego, CA: Academic Press.

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

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

Microsoft. (1993). Editing and proofing tools. In Microsoft (Ed.), *Microsoft Word user's guide.* Redmond, WA: Author.

Neuberg, S. L., & Newsom, J. T. (1993). Personal need for structure: Individual differences in the desire for simple structure. *Journal of Personality and Social Psychology, 65,* 113–131.

Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit* [Language and personality structure: Toward the validation of the five-factor models of personality]. Regensburg, Germany: S. Roeder Verlag.

Padilla, A. M. (1995). Hispanic psychology: Critical issues in theory and research. Thousand Oaks, CA: Sage.

Panter, A. T., Tanaka, J. S., & Hoyle, R. H. (1994). Structural models for multimode designs in personality and temperament research. In C. F. Halverson, G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 111–138). Hillsdale, NJ: Erlbaum.

Peabody, D., & Goldberg, L. R. (1989). Some determinants of factor structures from personality-trait descriptors. *Journal of Personality and Social Psychology, 57,* 552–567.

Psychological Assessment Resources. (1994). *The Revised NEO Personality Inventory: Manual supplement for the Spanish edition.* Odessa, FL: Author.

Real Academia Española. (1989). *Diccionario manual e ilustrado de la Real Academia de la Lengua Española* [Manual and illustrated dictionary of the Royal Spanish Academy of Language]: Cuarta edición revisada. Madrid, Spain: Espasa-Calpe.

Reynolds, A. G. (1991). *Bilingualism, multiculturalism, and second lan-*

guage learning: The McGill Conference in Honour of Wallace E. Lambert. Hillsdale, NJ: Erlbaum.

Rogler, L. H., Cortes, D. E., & Malgady, R. G. (1991). Acculturation and mental health status among Hispanics. *American Psychologist, 6,* 585–597.

Saucier, G. (1992a). Benchmarks: Integrating affective and interpersonal circles with the Big-Five personality factors. *Journal of Personality and Social Psychology, 62,* 1025–1035.

Saucier, G. (1992b). Openness versus Intellect: Much ado about nothing? *European Journal of Personality, 6,* 381–386.

Saucier, G. (1994). Mini-markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment, 63,* 506–516.

Saucier, G., & Goldberg, L. R. (1996). Evidence for the Big Five in analyses of familiar English personality adjectives. *European Journal of Personality, 10,* 61–77.

Schwartz, S. H. (1994). Beyond individualism/collectivism: New cultural dimensions of values. In U. Kim, H. C. Triandis, Ç. Kâgitçibasi, S. Choi, & G. Yoon (Eds.), *Individualism and collectivism: Theory, methods, and applications* (pp. 85–119). Thousand Oaks, CA: Sage.

Silva, F., Avia, D., Sanz, J., Martínez-Arias, R., Grana, J. L., & Sánchez-Bernardos, L. (1994). The five factor model—I. Contributions to the structure of the NEO-PI. *Personality and Individual Differences, 17,* 741–753.

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

Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the Interpersonal Adjective Scales to include the Big Five dimensions of personality. *Journal of Personality and Social Psychology, 59,* 781–790.

Triandis, H. C. (1990). Cross-cultural studies of individualism and collectivism. In J. J. Berman (Ed.), *Nebraska Symposium on Motivation: Vol. 37. Cross-cultural perspectives* (pp. 41–133). Lincoln: University of Nebraska Press.

Triandis, H. C., Lisansky, J., Marín, G., & Betancourt, H. (1984). Simpatía as a cultural script of Hispanics. *Journal of Personality and Social Psychology, 47,* 1363–1375.

Triandis, H. C., & Marín, G. (1983). Etic plus emic versus pseudoetic: A test of a basic assumption of contemporary cross-cultural psychology. *Journal of Cross-Cultural Psychology, 14,* 489–500.

U.S. Bureau of the Census. (1995). *The Hispanic population in the United States: March 1993* (Current Population Reports, Ser. P20, No. 475). Washington, DC: U.S. Government Printing Office.

Watson, D., & Clark, L. A. (1992). On traits and temperament: General and specific factors of emotional experience and their relation to the five-factor model. *Journal of Personality, 60,* 441–476.

Watson, D., Clark, L. A., & Harkness, A. R. (1994). Structures of personality and their relevance to psychopathology. *Journal of Abnormal Psychology, 103,* 18–31.

Widiger, T. A., & Trull, T. J. (1992). Personality and psychopathology: An application of the five-factor model. *Journal of Personality, 60,* 363–393.

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.

Appendix

English Big Five Inventory

Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who *likes to spend time with others?* Please choose a number for each statement to indicate the extent to which you agree or disagree with that statement.

Disagree strongly	Disagree a little	Neither agree nor disagree	Agree a little	Agree strongly
1	2	3	4	5

I see myself as *someone who* . . .

___ 1. is talkative
___ 2. tends to find fault with others
___ 3. does a thorough job
___ 4. is depressed, blue
___ 5. is original, comes up with new ideas
___ 6. is reserved
___ 7. is helpful and unselfish with others
___ 8. can be somewhat careless
___ 9. is relaxed, handles stress well
___10. is curious about many different things
___11. is full of energy
___12. starts quarrels with others
___13. is a reliable worker
___14. can be tense
___15. is ingenious, a deep thinker
___16. generates a lot of enthusiasm
___17. has a forgiving nature
___18. tends to be disorganized
___19. worries a lot
___20. has an active imagination
___21. tends to be quiet
___22. is generally trusting
___23. tends to be lazy
___24. is emotionally stable, not easily upset
___25. is inventive
___26. has an assertive personality
___27. can be cold and aloof
___28. perseveres until the task is finished
___29. can be moody
___30. values artistic, aesthetic experiences
___31. is sometimes shy, inhibited
___32. is considerate and kind to almost everyone
___33. does things efficiently
___34. remains calm in tense situations
___35. prefers work that is routine
___36. is outgoing, sociable
___37. is sometimes rude to others
___38. makes plans and follows through with them
___39. gets nervous easily
___40. likes to reflect, play with ideas
___41. has few artistic interests
___42. likes to cooperate with others
___43. is easily distracted
___44. is sophisticated in art, music, or literature

Please check: Did you write a number in front of each statement?

Note. Copyright 1991 by Oliver P. John.

(*Appendix continues*)

Spanish Big Five Inventory

Las siguientes expresiones le describen a usted con más o menos precisión. Por ejemplo, ¿está de acuerdo en que usted es alguien "*chistoso, a quien le gusta bromear*"? Por favor escoja un número para cada una de las siguientes expresiones, indicando así hasta que punto está de acuerdo o en desacuerdo en como le describe a usted.

Muy en desacuerdo	Ligeramente en desacuerdo	Ni de acuerdo ni en desacuerdo	Ligeramente de acuerdo	Muy de acuerdo
1	2	3	4	5

Me veo a mi mismo-a como *alguien que . . .*

___ 1. es bien hablador
___ 2. tiende a ser criticón
___ 3. es minucioso en el trabajo
___ 4. es depresivo, melancólico
___ 5. es original, se le ocurren ideas nuevas
___ 6. es reservado
___ 7. es generoso y ayuda a los demás
___ 8. puede a veces ser algo descuidado
___ 9. es calmado, controla bien el estrés
___10. tiene intereses muy diversos
___11. está lleno de energía
___12. prefiere trabajos que son rutinarios
___13. inicia disputas con los demás
___14. es un trabajador cumplidor, digno de confianza
___15. con frecuencia se pone tenso
___16. tiende a ser callado
___17. valora lo artístico, lo estético
___18. tiende a ser desorganizado
___19. es emocionalmente estable, difícil de alterar
___20. tiene una imaginación activa
___21. persevera hasta terminar el trabajo
___22. es a veces maleducado con los demás
___23. es inventivo
___24. es generalmente confiado
___25. tiende a ser flojo, vago
___26. se preocupa mucho por las cosas
___27. es a veces tímido, inhibido
___28. es indulgente, no le cuesta perdonar
___29. hace las cosas de manera eficiente
___30. es temperamental, de humor cambiante
___31. es ingenioso, analítico
___32. irradia entusiasmo
___33. es a veces frío y distante
___34. hace planes y los sigue cuidadosamente
___35. mantiene la calma en situaciones difíciles
___36. le gusta reflexionar, jugar con las ideas
___37. es considerado y amable con casi todo el mundo
___38. se pone nervioso con facilidad
___39. es educado en arte, música, o literatura
___40. es asertivo, no teme expresar lo que quiere
___41. le gusta cooperar con los demás
___42. se distrae con facilidad
___43. es extrovertido, sociable
___44. tiene pocos intereses artísticos

Por favor, compruebe que ha escrito un número delante de cada frase.

Note. Copyright 1996 by Oliver P. John and Verónica Benet-Martínez.

Received October 15, 1997
Revision received March 9, 1998
Accepted March 12, 1998 ■