



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Journal of Research in Personality 37 (2003) 504–528

JOURNAL OF  
RESEARCH IN  
PERSONALITY

[www.elsevier.com/locate/jrp](http://www.elsevier.com/locate/jrp)

## A very brief measure of the Big-Five personality domains<sup>☆</sup>

Samuel D. Gosling,<sup>\*</sup> Peter J. Rentfrow, and William B. Swann Jr.

*Department of Psychology, University of Texas, Seay Psychology Bldg. Rm. 4.212, Austin, TX 78712, USA*

---

### Abstract

When time is limited, researchers may be faced with the choice of using an extremely brief measure of the Big-Five personality dimensions or using no measure at all. To meet the need for a very brief measure, 5 and 10-item inventories were developed and evaluated. Although somewhat inferior to standard multi-item instruments, the instruments reached adequate levels in terms of: (a) convergence with widely used Big-Five measures in self, observer, and peer reports, (b) test–retest reliability, (c) patterns of predicted external correlates, and (d) convergence between self and observer ratings. On the basis of these tests, a 10-item measure of the Big-Five dimensions is offered for situations where very short measures are needed, personality is not the primary topic of interest, or researchers can tolerate the somewhat diminished psychometric properties associated with very brief measures.

© 2003 Elsevier Science (USA). All rights reserved.

---

### 1. Introduction

One obvious way to learn about an individual's standing on a personality trait is simply to enquire directly about that trait. For constructs, such as Extraversion, that are widely understood, it is more straightforward simply to ask a person how extraverted he is than to ask him whether he enjoys the company of others, attends parties frequently, is talkative, outgoing, gregarious, and enthusiastic. That is, why not ask a

---

<sup>☆</sup> Preparation of this article was supported by a research grant from the University of Texas College of Liberal Arts and National Institutes of Mental Health Grant RO3 MH64527-01A1. We thank Veronica Benet-Martinez, Matthias R. Mehl, and Richard W. Robins for their helpful comments on earlier drafts of this paper.

<sup>\*</sup> Corresponding author. Fax: 1-512-471-5935.

E-mail address: [gosling@psy.utexas.edu](mailto:gosling@psy.utexas.edu) (S.D. Gosling).

person one direct question about a trait rather than many questions about the multiple, narrow components that comprise the trait?

The widely accepted answer is that, all things being equal, long instruments tend to have better psychometric properties than short instruments. However, the costs associated with short instruments are not always as great as is feared (Burisch, 1984a, 1984b, 1997). More important, there are some instances when short instruments permit research that would not be possible using long instruments.

### *1.1. Why are short instruments needed?*

In an ideal world, personality researchers would have sufficient time and resources to exploit the superior content validity and reliability of well-established multi-item instruments. Unfortunately, circumstances are often not ideal and researchers may be faced with a stark choice of using an extremely brief instrument or using no instrument at all. For example, one Internet-based study used a single-item measure to obtain ratings of self-esteem from participants who would be unlikely to dwell at the website long enough to complete a multi-item questionnaire (Robins, Trzesniewski, Tracy, Gosling, & Potter, 2002). Studies that require participants to rate themselves and multiple others on several occasions may also profit from the use of short scales. In one longitudinal study of interpersonal perceptions, participants were required to rate several other group members on several traits on several occasions (Paulhus & Bruce, 1992); multi-item scales would have burdened participants excessively so single-item measures were used. Other useful applications for short instruments include large-scale surveys, pre-screening packets, longitudinal studies, and experience-sampling studies (Robins, Hendin, & Trzesniewski, 2001a).

Although single-item scales are usually psychometrically inferior to multiple-item scales, single-item measures do have some advantages. In developing a single-item measure of self-esteem, Robins et al. (2001a) noted that single-item measures "...eliminate item redundancy and therefore reduce the fatigue, frustration, and boredom associated with answering highly similar questions repeatedly" (p. 152; also see Saucier, 1994). Indeed, Burisch (1984b, 1997) showed that short and simple depression scales can be just as valid as long and sophisticated scales. For example, self and peer reports converged just as strongly for a truncated 9-item depression scale ( $r = .54$ ) as for the full 50-item scale ( $r = .51$ ). Burisch's findings suggest that the supposed psychometric superiority of longer scales does not always translate into practice. If the psychometric costs of using short scales are not as steep as might be expected, their relative efficiency make them a very attractive research tool. The widespread use of single-item measures is a testimony to their appeal. Single-item measures have been used to assess such constructs as life-satisfaction (Campbell, Converse, & Rodgers, 1976), subjective well-being (Diener, 1984; Sandvik, Diener, & Seidlitz, 1993), affect (Russell, Weiss, & Mendelsohn, 1989), cultural/ethnic identity (Benet-Martínez, Leu, Lee, & Morris, 2002), relationship intimacy (Aron, Aron, & Danny, 1992), attachment style (Hazan & Shaver, 1987), intelligence (Paulhus, Lysy, & Yik, 1998), and self-esteem (Robins, Tracy, Trzesniewski, Potter, & Gosling, 2001b).

### *1.2. Previous Big-Five instruments*

In this report, we evaluate new 5 and 10-item measures of the Big-Five personality dimensions. The Big-Five framework enjoys considerable support and has become the most widely used and extensively researched model of personality (for reviews, see John & Srivastava, 1999, and McCrae & Costa, 1999), although it has not been accepted universally (Block, 1995).

The Big-Five framework is a hierarchical model of personality traits with five broad factors, which represent personality at the broadest level of abstraction. Each bipolar factor (e.g., Extraversion vs. Introversion) summarizes several more specific facets (e.g., Sociability), which, in turn, subsume a large number of even more specific traits (e.g., talkative, outgoing). The Big-Five framework suggests that most individual differences in human personality can be classified into five broad, empirically derived domains.

Several rating instruments have been developed to measure the Big-Five dimensions. The most comprehensive instrument is Costa and McCrae's (1992) 240-item NEO Personality Inventory, Revised (NEO-PI-R), which permits measurement of the Big-Five domains and six specific facets within each dimension. Taking about 45 min to complete, the NEO-PI-R is too lengthy for many research purposes and so a number of shorter instruments are commonly used. Three well-established and widely used instruments are the 44-item Big-Five Inventory (BFI; see Benet-Martínez & John, 1998; John & Srivastava, 1999), the 60-item NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992), and Goldberg's instrument comprised of 100 trait descriptive adjectives (TDA; Goldberg, 1992). John and Srivastava (1999) have estimated that the BFI, NEO-FFI, and TDA take approximately 5, 15, and 15 min to complete, respectively. Recognizing the need for an even briefer measure of the Big Five, Saucier (1994) developed a 40-item instrument derived from Goldberg's (1992) 100-item set.

### *1.3. Overview of present research*

In two studies, we evaluate new 5 and 10-item measures of the Big Five in terms of convergence with an established Big-Five instrument (the BFI), test–retest reliability, and patterns of predicted external correlates. In Study 1, two samples were assessed using both the new five-item instrument and the BFI. Convergent and discriminant validity was examined in a sample of 1704 undergraduate students who were assessed using both instruments. To compare the pattern of external correlates of the 5-item instrument with the pattern of external correlates of the BFI, we also administered a battery of other instruments. To assess the test–retest reliability of the 5-item instrument and of the BFI, a subset of 118 participants were assessed again two weeks after the initial assessment. To evaluate the performance of the measure when used in observer-report format, a second subset of 60 participants were rated by observers after a brief getting acquainted exercise. To examine the measure when used in peer-report format, we also collected peer reports from a new sample of 83 participants.

In Study 2, one sample was assessed using both the 10-item instrument and the BFI. Convergent and discriminant validity was examined in a sample of 1813 undergraduate students who were assessed using both instruments. To compare the pattern of external correlates of the 10-item instrument with those of the BFI, a battery of other instruments was also administered. To evaluate the foci of the scales from the BFI and the 10-item instrument, we also administered the NEO-PI-R to a subset of 180 participants. To assess the test–retest reliability of the 10-item instrument, the same subset of participants were assessed again, six weeks after the initial assessment.

## 2. Study 1

The aim of Study 1 was to examine a new 5-item instrument designed to assess the Big-Five personality dimensions. We used four tests to evaluate the instrument, each time comparing the 5-item instrument to the BFI. First, to assess convergent and discriminant validity, we obtained self-ratings, observer ratings, and peer ratings using the 5-item instrument and the BFI.

Second, to assess test–retest reliability, a sub-sample of participants took the revised 5-item instrument and the BFI a second time, two weeks after the first test administration. Test–retest correlations are particularly valuable for single-item measures because internal-consistency indices of reliability cannot be computed.

Third, to examine patterns of external correlates, we also obtained self-ratings on several other measures. The construct validity of an instrument can be defined in terms of a nomological network (Cronbach & Meehl, 1955); that is, the degree to which a construct shows theoretically predicted patterns of correlations with other related and unrelated constructs. Our goal here was not to validate the Big-Five constructs but to evaluate the degree to which a very brief measure of the Big-Five constructs assesses the same constructs as those assessed by a longer, established measure. Therefore, the predicted nomological network for the 5-item instrument was provided by the pattern of correlations shown by the standard BFI to a broad range of constructs.

Fourth, to evaluate the convergence between self and observer reports, a sub-sample of participants were rated by observers after a brief getting acquainted exercise. (These data were also used to examine convergent and discriminant correlations in observer reports.)

### 2.1. Method

#### 2.1.1. Instruments

One approach to constructing short tests is to select the best performing items from longer tests on the basis of psychometric criteria, such as item-total correlations. For example, to create an abbreviated set of Big-Five markers from Goldberg's 100-item set, Saucier (1994) relied on psychometric criteria, selecting those items that showed high factor purity and would form reliable scales. The strategy

adopted here was different. Instead of psychometric criteria, we focused on optimizing the content validity of our short measure—we aimed to enhance the bandwidth of the items by including in each item several descriptors selected to capture the breadth of the Big-Five dimensions. Thus, we used a strategy akin to the one used by Hazan and Shaver (1987) who created paragraph-long items that clearly described the heart and breadth of the attachment-style constructs they were assessing. To create items, John and Srivastava (1999) have recommended adding elaborative, clarifying, or contextual information to one or two prototypical adjectives. John and Srivastava (1999) note that augmented items retain the advantages of brevity and simplicity associated with single adjectives, while avoiding some of their pitfalls, such as ambiguous or multiple meanings.

Thus, we consensually selected descriptors to represent each of the domains. Where possible, we called descriptors from existing Big-Five instruments, drawing most heavily on Goldberg's (1992) list of unipolar and bipolar Big-Five markers, adjectives from the BFI, and John and Srivastava's (1999) Adjective Checklist Big-Five markers.<sup>1</sup> Selection was based on the following five guidelines. First, we strove for breadth of coverage, using the facets of the Big Five to guide our selections. Second, we identified items representing both poles of each dimension. Third, where possible we selected items that were not evaluatively extreme. Fourth, for the sake of clarity, we avoided using items that were simply negations. Fifth, we attempted to minimize redundancy among the descriptors. We developed a standard format, in which each item was defined by two central descriptors and clarified by six other descriptors, that together covered the breadth of each domain and included items from the high and low poles. The resulting five items were: Extraverted, enthusiastic (that is, sociable, assertive, talkative, active, *NOT* reserved, or shy); Agreeable, kind (that is, trusting, generous, sympathetic, cooperative, *NOT* aggressive, or cold); Dependable, organized (that is, hard working, responsible, self-disciplined, thorough, *NOT* careless, or impulsive); Emotionally stable, calm (that is, relaxed, self-confident, *NOT* anxious, moody, easily upset, or easily stressed); Open to experience, imaginative (that is, curious, reflective, creative, deep, open-minded, *NOT* conventional). Each of the five items was rated on a 7-point scale ranging from 1 (*disagree strongly*) to 7 (*agree strongly*).

Participants also completed the 44-item BFI (John & Srivastava, 1999). The BFI shows high convergent validity with other self-report scales and with peer ratings of

---

<sup>1</sup> In a pilot study, we used single-item scales based on the labels commonly used to refer to the Big Five dimensions: "Extraverted," "Agreeable, warm," "Conscientious," "Emotionally stable," and "Open to new experiences." The major finding to emerge from this study was that "conscientious" was hard for lay judges to interpret; convergent correlations between ratings on the "conscientious" item and the BFI conscientiousness scale were only .22 for self reports and .36 for peer reports (compared to .81 and .76, .64 and .72, .65 and .70, and .55 and .51, respectively, for Extraversion, Agreeableness, Emotional Stability and Openness). However, when judges who were familiar with the Big Five used the item, the convergent correlations rose dramatically—the convergent correlation between peer ratings by experts on the "conscientious" item and the BFI conscientiousness was .81. The lesson to emerge from these pilot data was that the Big Five definitions are not conveyed to laypersons by the common Big Five labels, and that care should be taken to select items familiar to laypersons.

the Big Five. The BFI items were rated on a 5-point scale ranging from 1 (*disagree strongly*) to 5 (*agree strongly*).

To permit us to examine a broad array of external correlates of the new Five Item Personality Inventory (FIPI) and the BFI, participants also completed a battery of other measures: the Brief Loquaciousness and Interpersonal Responsiveness Test (BLIRT; Swann & Rentfrow, 2001), the Social Dominance Orientation questionnaire (SDO; Pratto, Sidanius, Stallworth, & Malle, 1994), the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965), the Beck Depression Inventory (BDI; Beck, 1972), the Math Identification Questionnaire (MIQ; Brown & Josephs, 1999), the Short Test of Music Preferences (STOMP; Rentfrow & Gosling, 2003), and single-item measures of political values, physical attractiveness, wealth, athletic ability, and intelligence.

### 2.1.2. Recruitment and participants

We examined the FIPI in two samples (A1 and B) of participants, as well as two sub-samples (Samples A2 and A3) drawn from sample A1. Sample A1 was comprised of 1704 University of Texas at Austin undergraduates who volunteered in exchange for partial fulfillment of an introductory psychology course requirement. Participants completed the battery of instruments described above. To reduce the carryover from the FIPI to the BFI, we separated them with several other instruments. Of those who indicated, 1058 (62.6%) were women, 633 (37.4%) were men, 205 (13.5%) were Asian, 205 (13.5%) were Hispanic, 988 (65%) were White, and 122 (8%) were of other ethnicities.

Two weeks later, a sub-sample of 118 of the participants (sample A2) were tested again. In exchange for partial fulfillment of an introductory psychology course requirement, participants completed a second battery of questionnaires, including the FIPI and the BFI. Of those who indicated, 94 (81.7%) were women and 21 (18.3%) were men, 25 (21.9%) were Asian, 11 (9.6%) were Hispanic, 64 (56.2%) were White, and 14 (12.3%) were of other ethnicities.

Between 2 and 9 weeks later, another sub-sample of 60 participants (sample A3) returned to the lab in previously unacquainted same sex pairs.<sup>2</sup> In exchange for partial fulfillment of an introductory psychology course requirement, each pair of participants took part in a zero-acquaintance type exercise in which they had 15 min to introduce themselves and learn a little about one another. After just 15-min, each participant rated the other participant using peer-report versions of the FIPI and the BFI. Of those who indicated, 27 (45.8%) were women and 32 (54.2%) were men, 7 (11.9%) were Asian, 8 (13.6%) were Hispanic, 32 (54.2%) were White, and 12 (20.3%) were of other ethnicities.

Sample B was comprised of 83 University of Texas undergraduates who volunteered in exchange for partial fulfillment of an introductory psychology course requirement. Participants completed a second battery of questionnaires, including peer-report versions of the FIPI and the BFI, which they used to describe “a person

---

<sup>2</sup> Data for one participant were excluded because self-ratings were not provided.

you know well.” Thus, the participants in this study were raters, not targets. Of those who indicated, 26 (49.1%) were women and 27 (50.9%) were men, 9 (17.3%) were Asian, 7 (13.5%) were Hispanic, 34 (65.4%) were White, and 2 (3.8%) were of other ethnicities.

## 2.2. Results and discussion

### 2.2.1. Convergence across measures

In samples A1, A2, A3, and B, we correlated the BFI scale scores with the FIPI scores. The convergent correlations in sample A1 between the BFI and FIPI are shown in the diagonal of Table 1; these convergent correlations (mean  $r = .65$ ) far exceeded the off-diagonal correlations (absolute mean  $r = .18$ ), none of which exceeded .30. More important from a construct validation perspective (Campbell & Fiske, 1959), the pattern of off-diagonal correlations is similar across both measures, correlating .65 across the 10 possible off-diagonal discriminant correlations.

The convergent correlations and off-diagonal discriminant correlations from the self-report data shown individually in Table 1 are summarized in column 1 of Table 2. Convergent and discriminant correlations from samples A2, A3, and B are summarized in data columns 2, 3, and 4 of Table 2 respectively. The convergent correlations were substantial, with the patterns of correlations holding across self (samples A1 and A2), zero-aquaintance observer (sample A3), and peer-ratings (sample B). As in sample A1, the convergent correlations consistently exceeded the off-diagonal discriminant correlations.

As shown in Table 2, the discriminant correlations in both instruments were unusually large for the observer ratings. However, it should be borne in mind that these ratings were made about strangers after only 15 min of getting acquainted, so the strong discriminant correlations probably reflect the limited exposure to relevant behavioral information rather than deficiencies of the instruments themselves.

Table 1  
Convergent correlations between Five-Item Personality Inventory (FIPI) and Big-Five Inventory (BFI): Self-reports

	Big-Five Inventory (BFI)				
	Extrav.	Agree.	Cons.	Em.St.	Open.
Five-Item Personality Inventory (FIPI)					
Extraversion	<b>.80**</b>	.13**	.16**	.27**	.18**
Agreeableness	.17**	<b>.58**</b>	.26**	.26**	.01
Conscientiousness	.04	.30**	<b>.65**</b>	.19**	.02
Emotional Stability	.22**	.22**	.24**	<b>.69**</b>	.16**
Openness to Experience	.25**	.17**	.01	.23**	<b>.48**</b>

Note:  $N = 1704$ . Extrav., Extraversion; Agree., Agreeableness; Cons., Conscientiousness; Em.St., Emotional Stability; Open., Openness to Experience. Convergent correlations are shown in bold typeface on the diagonal. Discriminant correlations are shown below the diagonal for the FIPI and above the diagonal for the BFI.

\*\*  $p < .01$ .

Table 2  
Summary of convergent (diagonal) and discriminant (off-diagonal) correlations across the Big-Five dimensions

	Study 1					Study 2			John and Srivastava (1999)		Mean
	1	2	3	4	5	6	7	8	9		
Data column:	A1	A2	A3	B		C1	TDA <sup>a</sup>	FFI <sup>a</sup>			
Sample:	FIPI	FIPI	FIPI	FIPI		TIPI	Self	Self			
Instrument:	Self	Self	Obsrv.	Peer		Self	Self	Self			
Source:	1704	118	59	83		1813	462	462			
Sample size:											
Convergent (diagonal) correlations with BFI											
Extraversion	.80	.74	.74	.73	.76	.87	.90	.69	.82		
Agreeableness	.58	.62	.69	.70	.59	.70	.78	.76	.77		
Conscientiousness	.65	.56	.70	.65	.63	.75	.81	.79	.80		
Emotional Stability	.69	.64	.58	.70	.63	.81	.76	.76	.76		
Openness to Experience	.48	.53	.69	.60	.57	.65	.75	.64	.70		
Mean	<b>.65</b>	<b>.62</b>	<b>.68</b>	<b>.68</b>	<b>.66</b>	<b>.77</b>	<b>.81</b>	<b>.73</b>	<b>.77</b>		
Discriminant (off-diagonal) correlations (absolute means)											
BFI	.17	.18	.43	.18	.24	.20	—	—	—		
Very brief measure (FIPI or TIPI)	.19	.19	.36	.35	.27	.20	—	—	—		
Mean	<b>.18</b>	<b>.18</b>	<b>.40</b>	<b>.27</b>	<b>.26</b>	<b>.20</b>					

*Note.* Correlations were computed in two samples (A1 and B) in Study 1 and in one sample (C1) in Study 2. Sub-samples of sample A1 were tested a second time to examine test-retest correlations (A2) and observer ratings (A3). Instruments used were the Five-Item Personality Inventory (FIPI) developed in Study 1; the Ten-Item Personality Inventory (TIPI) developed in Study 2; Goldberg's (1992) 100-item measure using trait descriptive adjectives (TDA); and Costa and McCrae's (1992) 60-item NEO Five-Factor Inventory (FFI). BFI, Big-Five Inventory (John & Srivastava, 1999). Ratings were provided by the self (Self), peer acquaintances (Peer), or observers (Obsrv.). Discriminant correlations are the means of the absolute off-diagonal correlations. All means were computed using Fisher r-to-z transformations, and are shown in boldface type; "—" indicates that data were not available.

<sup>a</sup>Convergent validities reported in John and Srivastava (1999, Table 4.3); off-diagonal correlations were not reported.



Overall, the convergent and discriminant correlations appear promising but how do they compare with correlations obtained using well-established multi-item instruments? The fifth data column of Table 2 summarizes the convergent and discriminant correlations obtained using the 5-item instrument tested in Study 1. For purposes of comparison the 7th and 8th data columns in Table 2 show convergent correlations (from John & Srivastava, 1999) between the BFI and two other standard multi-item measures of the Big-Five dimensions<sup>3</sup>—Goldberg's (1992) 100-item adjectival instrument, and Costa and McCrae's (1992) NEO-FFI. Multi-item inventories tend to be more reliable than single-item inventories so one would expect stronger correlations between the BFI and the multi-item instruments than between the BFI and the 5-item instrument; indeed, the convergences among the multi-item inventories were stronger ( $r_s = .81$  and  $.73$ ) than the convergences between the BFI and the 5-item instrument, which averaged in the mid to high .60s.

### 2.2.2. Test–retest reliability

We estimated test–retest reliability for each instrument by correlating scores obtained in the first rating session with the scores obtained in a second rating session, approximately two weeks later. As shown in the first two data columns of Table 3, the test–retest correlations are substantial for both the BFI and the FIPI. However, with the exception of Extraversion, which was equivalent across instruments, the BFI yielded test–retest correlations (mean  $r = .80$ ) that were .10 to .20 points stronger than the FIPI (mean  $r = .68$ ).

### 2.2.3. External correlates

To test whether the patterns of external correlates of the FIPI matched the patterns of external correlates of the BFI, we correlated both instruments with each of the other constructs assessed in the testing battery. These patterns of external correlates are shown separately for each of the Big-Five dimensions in Table 4. As one would expect for such a broad array of constructs the magnitude of correlations varied greatly. To test whether the patterns of correlations were similar across instruments, we computed column-vector correlations for each of the five dimensions. Specifically, we transformed the correlations using Fisher's  $r$ -to- $z$  formula and then computed the correlation between the two columns of transformed correlations. As shown by the column-vector correlations in Table 4, the FIPI displayed patterns of correlations that were virtually identical to the BFI; column-vector correlations ranged from .819 for Openness to .997 for Extraversion.<sup>4</sup> Overall, the FIPI exhibited patterns of correlations that matched very closely those obtained for the BFI. However, because the FIPI was less reliable than the BFI the absolute magnitude of the

<sup>3</sup> John and Srivastava (1999) did not report the discriminant correlations for these instruments.

<sup>4</sup> It should be noted that strong column-vector correlations could be generated merely from the inclusion of a mixture of constructs, some of which correlate strongly and some of which correlate weakly with the Big-Five dimensions.

Table 3  
Test–retest reliability for the BFI, FIPI, and the TIPI

	Test–retest reliability		
	BFI ( <i>N</i> = 114)	FIPI ( <i>N</i> = 114)	TIPI ( <i>N</i> = 180)
Extraversion	.82	.81	.77
Agreeableness	.76	.63	.71
Conscientiousness	.76	.67	.76
Emotional Stability	.83	.65	.70
Openness to Experience	.80	.60	.62
Mean	.80	.68	.72

*Note.* BFI, Big-Five Inventory (John & Srivastava, 1999); FIPI, Five-Item Personality Inventory; TIPI, Ten-Item Personality Inventory. FIPI and BFI retest interval was approximately 2 weeks; TIPI retest interval was approximately 6 weeks.

correlations should be weaker. As shown in the bottom row of Table 4, the BFI consistently yielded stronger correlations than the FIPI, although the magnitude of the difference was small.

#### 2.2.4. Convergence across self- and observer-reports

To examine the convergence between self- and observer-reports, we correlated the self-reports with observer-reports obtained 2–9 weeks later. Recall that the observer-reports were made on the basis of only a 15-min getting acquainted exercise so we did not expect the magnitude of the correlations to be very strong. For the purposes of evaluating the FIPI, we were primarily interested in comparing the self-observer correlations obtained using the FIPI with those obtained using the BFI. As shown in Table 5, the pattern of FIPI correlations generally matched those of the BFI. However, the magnitude of the BFI correlations was stronger than that of the FIPI correlations, especially for Extraversion, Agreeableness, and Openness.

#### 2.2.5. Summary

Although somewhat inferior to the standard Big-Five instrument examined here, the FIPI reached adequate levels in each of the four criteria against which it was evaluated: convergent and discriminant validity, test–retest reliability, patterns of external correlates, and convergence between self- and observer-ratings. The FIPI measure of Extraversion fared the best across the criteria and FIPI measure of Openness fared least well. Overall our findings suggest that the FIPI instrument can stand as a reasonable proxy for a longer Big-Five instrument, especially when research conditions dictate that a very short measure be used.

#### 2.2.6. Limitations

Despite the evidence for the value of the FIPI, single-item measures are subject to some serious limitations. The first limitation is the psychometric cost of using short measures. Compared with standard multi-item measures of the Big Five, the FIPI is

Table 4  
External correlates of the BFI, FIPI, and TIPI

Criterion measures	Extraversion			Agreeableness			Conscientiousness			Emotional Stability			Openness		
	BFI	FIPI <sup>a</sup>	TIPI <sup>b</sup>	BFI	FIPI <sup>a</sup>	TIPI <sup>b</sup>	BFI	FIPI <sup>a</sup>	TIPI <sup>b</sup>	BFI	FIPI <sup>a</sup>	TIPI <sup>b</sup>	BFI	FIPI <sup>a</sup>	TIPI <sup>b</sup>
BLIRT	.58**	.49**	.53**	-.09**	-.07**	-.09**	.16**	.14**	.14**	.31**	.19**	.24**	.22**	.18**	.30**
SDO	-.01	.01	-.05	-.23**	-.14**	-.26**	-.04	-.04	-.06*	.07**	.05	-.02	-.08*	-.10**	-.09*
RSES	.38**	.32**	—	.21**	.16**	—	.31**	.21**	—	.29**	.19**	—	.12**	.15**	—
BDI	-.23**	-.20**	-.16**	-.23**	-.15**	-.15**	-.25**	-.18**	-.20**	-.53**	-.47**	-.52**	-.04**	-.09**	-.06*
MIQ	-.06	-.06*	-.07**	.02	.04	-.04	.12**	.12**	.08*	.04	.05*	.01	-.08**	-.02	-.04
STOMP															
Reflective and Complex	.02	-.02	.03	.02	.02	.07**	-.03	-.09**	-.05	.07*	.02	.08*	.42**	.16**	.28**
Intense and Rebellious	.02	-.01	.06	-.01	.01	.05	-.04	-.06**	-.04	-.03	-.01	-.07*	.17**	.16**	.12**
Upbeat and Conventional	.22**	.21**	.19**	.22**	.19**	.19**	.14**	.17**	.14**	-.06	.04	.01	-.12**	-.07**	-.02
Energetic and Rhythmic	.21**	.21**	.19**	.07**	.06**	.06**	.00	.02	.01	.00	.04	.01	.05	.13**	.14**
Politically liberal	.01	.01	.02	-.07	-.02	-.01	-.10**	-.13**	-.08**	-.02	.07**	-.09**	.21**	.24**	.20**
Politically conservative	.01	.01	-.02	.05	.01	.01	.10**	.11**	.06*	.07*	.09**	.10**	-.15**	-.22**	-.17**
Physically attractive	.31**	.30**	.23**	.08**	.10**	.05	.17**	.07**	.19**	.22**	.18**	.21**	.12**	.13**	.17**
Wealthy	.14**	.13**	.10**	.03	.07**	.01	.06*	.00	.05	.13**	.13**	.09*	-.01	-.02	.00
Athletic	.25**	.21**	.22**	.08**	.06**	.02	.18**	.07**	.19**	.23**	.19**	.19**	.02	.07**	.12**
Intelligent	.19**	.15**	.14**	.03	.07**	.09**	.21**	.09**	.18**	.24**	.18**	.25**	.31**	.15**	.27**
BFI-FIPI column-vector correlations	.993			.977			.890			.968			.830		
BFI-TIPI column-vector correlations	.989			.913			.983			.979			.914		
Mean absolute magnitude of correlations	.18	.16	.14	.10	.08	.08	.13	.09	.11	.15	.13	.14	.14	.13	.14

Note. BFI, Big-Five Inventory (John & Srivastava, 1999); FIPI, Five-Item Personality Inventory; TIPI, Ten-Item Personality Inventory; BLIRT, Brief Loquaciousness Interpersonal Responsiveness Test (Swann & Rentfrow, 2001); SDO, Social Dominance Orientation (Pratto et al., 1994); RSES, Rosenberg Self-Esteem Scale (Rosenberg, 1965); BDI, Beck Depression Inventory (Beck, 1972); MIQ, Math Identification Questionnaire (Brown & Josephs, 1999); STOMP, Short Test Of Music Preferences (Rentfrow & Gosling, 2003). All means column-vector correlations were based on correlations transformed using Fisher's r-to-z formula. "—" indicates that data were not available.

<sup>a</sup>N = 1690.  
<sup>b</sup>N = 1813. The BFI correlations are the means of correlations derived in the two samples (i.e., total N = 3503).  
 \* p < .05.  
 \*\* p < .01.

Table 5  
Correlations between self-reports and subsequent ratings by observers in a zero acquaintance situation

	Self-observer correlations	
	BFI	FIPI
Extraversion	.46	.34
Agreeableness	.24	.02
Conscientiousness	.26	.31
Emotional Stability	.29	.25
Openness to Experience	.49	.38
Mean	.35	.26

*Note.* BFI, Big-Five Inventory (John & Srivastava, 1999); FIPI, Five-Item Personality Inventory. All means were computed using Fisher *r*-to-*z* transformations.

*N* = 59.

less reliable, converges less strongly with other Big-Five measures, and has weaker correlations with other variables.

The second limitation associated with the FIPI concerns its use in analyses involving latent variables, such as structural equation modeling. Such analyses estimate the error terms with which latent variables are measured by using multiple indicators of the same latent variable. Single-item measures provide only a single estimate of the latent variable so error terms must be estimated by other means. For example, researchers may have to estimate the error terms from test–retest correlations, such as the ones reported in Table 3.

In addition, unlike multi-item measures, which can balance positive and negative items, single-item measures cannot assess or control for acquiescence bias. Nor do single items permit researchers to check for errors (e.g., from data entry) using internal consistency estimates or factor analysis.

Although the FIPI fared reasonably well compared to the BFI, there are clearly a number of weakness associated with single-item measures. Therefore, we set out to create a second brief measure of the Big-Five personality dimensions that avoided the pitfalls associated with single-item scales.

### 3. Study 2

The goal of Study 2 was to develop and evaluate a 10-item measure of the Big-Five personality dimensions. It was important that this new measure retain the brevity of the FIPI, while diminishing the limitations associated with it. We used three tests to evaluate the instrument, each time comparing the 10-item instrument with the BFI. First, to assess convergent and discriminant validity, we obtained self-ratings using both the 10-item instrument and the BFI. Second, to assess test–retest reliability, a sub-sample of participants took the revised 10-item instrument a second time, six weeks after the first test administration. Third, to examine patterns of external correlates, we also obtained self-ratings on several other measures.

### 3.1. Method

#### 3.1.1. Instruments

The logic underlying the construction of the new 10-item questionnaire was similar to that adopted for the FIPI. That is, we again strove to retain breadth of coverage, represent both poles of each dimension, and to avoid items that were evaluatively extreme, items that were simply negations, and redundancy among items. With these goals in mind we created a 10-item inventory, with one item representing each pole of the five FFM dimensions. We reduced the length and complexity of the items to make them easier to understand and to allow us to double the number of items on the inventory without increasing the time taken to complete it. As in the construction of the FIPI, we culled descriptors from existing Big-Five instruments, drawing most heavily on Goldberg's (1992) list of unipolar and bipolar Big-Five markers, adjectives from the BFI, and John and Srivastava's (1999) Adjective Checklist Big-Five markers.

The resulting Ten-Item Personality Inventory (TIPI) is reproduced in Appendix A. Each item consists of two descriptors, separated by a comma, using the common stem, "I see myself as:". Each of the five items was rated on a 7-point scale ranging from 1 (*disagree strongly*) to 7 (*agree strongly*). The TIPI takes about a minute to complete.

Many scales are designed with a goal of optimizing internal consistency. The most widely used index of internal consistency, Cronbach's alpha, is a function of the mean inter-item correlation and the number of items comprising the scale. Multi-item scales can afford to bolster internal consistency by using several items with high content overlap. In contrast, with only two items per scale, the TIPI instead emphasized content validity considerations, resulting in lower inter-item correlations than is typical of more homogenous scales. The relatively low inter-item correlations in conjunction with the fact that the TIPI scales have only two items results in some unusually low internal consistency estimates. Specifically, the Cronbach alphas were .68, .40, .50, .73, and .45 for the Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience scales respectively. Thus, these scales provide an example of how validity can exceed reliability (as indexed by alpha). Therefore, researchers wishing to correct TIPI correlations for unreliability should base their corrections on reliability estimates that are less biased by our efforts to retain content validity or the small number of items on each scale; one such estimate would be the test-retest reliability correlations provided in Table 3.

The procedures for Study 2 were similar to those for Study 1. To examine convergence with an established Big-Five instrument, participants completed the 44-item BFI (John & Srivastava, 1999). In addition, to permit us to examine the external correlates of the TIPI and the BFI, participants also completed the same battery of measures included in Study 1: the BLIRT (Swann & Rentfrow, 2001), the SDO (Pratto et al., 1994), the BDI (Beck, 1972), the MIQ (Brown & Josephs, 1999), the STOMP (Rentfrow & Gosling, 2003), and single-item measures of political values, physical attractiveness, wealth, athletic ability, and intelligence.

### 3.1.2. Recruitment and participants

We examined the TIPI in one sample of participants (sample C1), as well as one retest sample (C2). Sample C1 was comprised of 1813 University of Texas at Austin undergraduates who volunteered in exchange for partial fulfillment of an introductory psychology course requirement. Participants completed the battery of instruments described above. To reduce the carryover from the TIPI to the BFI, we separated them with several other instruments. Of those who indicated, 1173 (65%) were women and 633 (35%) were men, 333 (18.5%) were Asian, 229 (12.7%) were Hispanic, 1124 (62.3%) were White, and 117 (6.5%) were of other ethnicities. Norms for the TIPI, derived from this sample, are presented in Appendix B.

Six weeks later, a sub-sample of 180 of the participants (sample C2) were tested again. In exchange for partial fulfillment of an introductory psychology course requirement, participants completed a second battery of questionnaires, including the TIPI and the NEO-PI-R (Costa & McCrae, 1992). Of those who indicated, 121 (69.9%) were women and 52 (30.1%) were men, 30 (17.3%) were Asian, 19 (11%) were Hispanic, 110 (63.6%) were White, and 14 (8.1%) were of other ethnicities.

## 3.2. Results

### 3.2.1. Convergence across measures

In sample C1, we correlated the BFI scale scores with the TIPI scale scores. The convergent validities, shown in the sixth data column of Table 2 (and also along the diagonal of Table 6), are substantial.

The off-diagonal discriminant correlations from the self-report data summarized in column 6 of Table 2 are shown individually in Table 6. The convergent correlations

Table 6  
Convergent correlations between Ten-Item Personality Inventory (TIPI) and Big-Five Inventory (BFI): Self-reports

	Big-Five Inventory (BFI)				
	Extrav.	Agree.	Cons.	Em.St.	Open.
Ten-Item Personality Inventory (TIPI)					
Extraversion	<b>.87**</b>	.13**	.19**	.31**	.25**
Agreeableness	.08*	<b>.70**</b>	.30**	.30**	.07*
Conscientiousness	.10**	.17**	<b>.75**</b>	.25**	.06*
Emotional Stability	.23**	.31**	.21**	<b>.81**</b>	.16**
Openness to Experience	.36**	.19**	.12**	.21**	<b>.65**</b>

Note.  $N = 1813$ . Big-Five Inventory (John & Srivastava, 1999); TIPI, Ten-Item Personality Inventory. Extrav., Extraversion; Agree., Agreeableness; Cons., Conscientiousness; Em.St., Emotional Stability; Open., Openness to Experience. Convergent correlations are shown in bold typeface on the diagonal. Discriminant correlations are shown below the diagonal for the TIPI and above the diagonal for the BFI.

\*  $p < .05$ .

\*\*  $p < .01$ .

(mean  $r = .77$ ) far exceeded the discriminant correlations (absolute mean  $r = .20$ ) and none of the discriminant correlations exceeded .36. As in Study 1, the pattern of off-diagonal correlations is similar across both measures, correlating .55 across the 10 possible off-diagonal discriminant correlations.

The convergent and discriminant correlations appear promising but how do they compare with correlations obtained using well-established multi-item instruments? The convergent and discriminant correlations obtained using the 10-item instrument can be compared to the 7th and 8th data columns, which show convergent correlations between the BFI and two other standard multi-item measures of the Big-Five dimensions. Whereas the convergences between the FIPI and the BFI (mean  $r = .66$ ) were somewhat smaller than the convergences between the BFI and the other multi-item inventories (mean  $r = .77$ ), the TIPI displayed convergences that were comparable to the other multi-item inventories (mean  $r = .77$ ).

We next examined the correlates of the 10 individual TIPI items. Table 7 shows the correlations among the 10 TIPI items, after standard scoring the raw data within participants to control for response biases. Table 8 shows the correlations between the individual TIPI items and the BFI scales, and between the individual TIPI items and composites comprised of just the positive and just the negative BFI items.

To provide a finer grained portrait of the breadth and content domains of the 10-item measure, we examined the convergent and discriminant correlations between the TIPI and the NEO-PI-R scale and facet scores. For purposes of comparison, we also examined the correlations between the BFI and the NEO-PI-R. Note that the NEO-PI-R was administered approximately 6 weeks after the BFI and TIPI. As shown in Table 9, the convergent correlations between the TIPI and the NEO-PI-R dimension scales were strong, ranging from .68 for Conscientiousness to .56 for Openness.

Next, we compared TIPI–NEO-PI-R correlations with the BFI–NEO-PI-R correlations. As should be expected, the correlations were generally stronger between the NEO-PI-R and the 44-item BFI than between the NEO-PI-R and the 10-item TIPI. Nevertheless, the patterns of NEO-PI-R *dimension* correlations were similar across measures. Moreover, the pattern and magnitude of the TIPI–NEO-PI-R and BFI–NEO-PI-R *facet* correlations were very similar, with only a few differences. For instance, the Openness dimension of the BFI appears to place more emphasis on fantasy, aesthetics, and ideas than does the TIPI. Overall, these findings provide good evidence for the construct validity of the 10-item measure.

### 3.2.2. Test–retest reliability

We estimated test–retest reliability for the TIPI by correlating scores obtained in the first rating session with the scores obtained in a second rating session, approximately six weeks later. As shown in the third data column of Table 3, the test–retest correlations for the TIPI are substantial (mean  $r = .72$ ), albeit weaker than the two-week test–retest correlations of the BFI (mean  $r = .80$ ).

Table 7  
Correlations among the 10 TIPI items: Standardized self-reports

Big-Five scale	TIPI item									
	1	2	3	4	5	6	7	8	9	10
Extraversion										
1. Extraverted, enthusiastic	—									
2. Reserved, quiet	<b>-.59**</b>	—								
Agreeableness										
3. Sympathetic, warm	.01	-.02	—							
4. Critical, quarrelsome	-.05*	.18**	<b>-.36**</b>	—						
Conscientiousness										
5. Dependable, self-disciplined	.02	-.04	-.11**	-.07**	—					
6. Disorganized, careless	-.11**	.11**	-.09**	.10**	<b>-.42**</b>	—				
Emotional Stability										
7. Calm, emotionally stable	.14**	-.09**	.01**	-.24**	.05*	-.15**	—			
8. Anxious, easily upset	-.23**	.03**	-.07**	.13**	-.06**	.06*	<b>-.61**</b>	—		
Openness to New Experiences										
9. Open to new experiences, complex	.16**	-.26**	-.04	-.06**	-.09**	.03	.03	-.21**	—	
10. Conventional, uncreative	-.25**	-.08**	-.09**	-.19**	.01	-.18**	-.05*	-.10**	<b>-.28**</b>	—

Note: N = 1,799. Correlations between positively and negatively keyed TIPI items for the same dimension are shown in bold typeface.

\* p < .05.

\*\* p < .01.



Table 8  
Correlations between individual TIPI items and BFI scales and positive and negative BFI-item composites

TIPI item	BFI scales and positive and negative item composites											
	Extraversion		Agreeableness		Conscientiousness		Emotional Stability		Openness			
	Scale	Pos	Neg	Scale	Pos	Neg	Scale	Pos	Neg	Scale	Pos	Neg
<b>Extraversion</b>												
Extraverted, enthusiastic	<b>.75**</b>	.78**	-.54**	.23**	.29**	-.10**	.26**	.26**	.28**	.25**	.24**	.25**
Reserved, quiet	-.77**	-.59**	.83**	-.04	-.03	.04	-.08**	.08**	-.20**	-.08**	-.15**	-.13**
<b>Agreeableness</b>												
Sympathetic, warm	.13**	.21**	.01	<b>.55**</b>	<b>.59**</b>	-.36**	.15**	.23**	.05*	.10**	.17**	.19**
Critical, quarrelsome	.00	.00	.00	-.57**	-.35**	.62**	-.12**	-.06*	-.24**	-.17**	-.01	.01
<b>Conscientiousness</b>												
Dependable,	.15**	.23**	-.01	.23**	.28**	-.13**	<b>.62**</b>	-.36**	.14**	.19**	.09**	.13**
self-disciplined												
Disorganized, careless	-.05*	-.06*	.04	-.16**	-.10**	.18**	-.66**	.70**	-.13**	-.08**	.04	.03
<b>Emotional Stability</b>												
Calm, emotionally stable	.22**	.28**	-.08**	.33**	.32**	-.25**	.25**	.27**	.71**	.72**	.12**	.13**
Anxious, easily upset	-.24**	-.22**	.21**	-.28**	-.19**	.30**	-.18**	-.13**	.19**	-.73**	-.13**	-.10**
<b>Openness to New Experiences</b>												
Open to new experiences,	.34**	.39**	-.20**	.17**	.24**	-.06*	.07**	.19**	.21**	.25**	.43**	.42**
complex												
Conventional, uncreative	-.26**	-.26**	.20**	-.13**	-.11**	.11**	-.12**	.09**	-.14**	-.10**	-.61**	-.60**

Note. N = 1813. Convergent correlations between TIPI items and BFI scales are in bold typeface; Convergent correlations between positively and negatively keyed TIPI items and positive and negative BFI item composites are underlined. Pos, composite of positively keyed BFI items; Neg, composite of negatively keyed BFI items.

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

Table 9  
Correlations of the BFI and TIPI with the NEO-PI-R administered at a different time

NEO-PI-R	Extraversion		Agreeableness		Conscientiousness		Emotional Stab.		Openness	
	BFI	TIPI	BFI	TIPI	BFI	TIPI	BFI	TIPI	BFI	TIPI
Extraversion	<b>.76</b>	<b>.65</b>	.32	.26	.26	.26	.12	.11	.26	.37
Warmth	<b>.66</b>	<b>.59</b>	.49	.47	.19	.15	.09	.08	.23	.32
Gregariousness	<b>.60</b>	<b>.45</b>	.14	.15	.06	.09	.02	-.03	.10	.22
Assertiveness	<b>.64</b>	<b>.61</b>	.10	.01	.33	.38	.25	.19	.18	.25
Activity	<b>.55</b>	<b>.47</b>	.17	.07	.35	.31	-.03	-.04	.15	.23
Excitement Seeking	<b>.35</b>	<b>.26</b>	.10	.13	.02	.02	.06	.06	.16	.26
Positive Emotions	<b>.51</b>	<b>.43</b>	.41	.35	.17	.17	.12	.22	.31	.34
Agreeableness	.09	.07	<b>.66</b>	<b>.59</b>	.16	.10	.14	.14	.00	.07
Trust	.34	.32	<b>.54</b>	<b>.42</b>	.25	.11	.22	.20	-.02	.04
Compliance	.00	-.05	<b>.39</b>	<b>.20</b>	.24	.17	.05	.10	-.12	-.05
Altruism	.21	.15	<b>.60</b>	<b>.59</b>	.12	.16	.08	.11	.12	.16
Straightforwardness	-.14	-.12	<b>.48</b>	<b>.49</b>	.06	.07	.12	.15	-.07	-.01
Modesty	-.27	-.21	<b>.23</b>	<b>.23</b>	-.10	-.13	-.08	-.13	-.01	-.03
Tender-Mindedness	.22	.19	<b>.42</b>	<b>.45</b>	.04	.01	.14	.13	.13	.22
Conscientiousness	.12	.09	.20	.17	<b>.70</b>	<b>.68</b>	.05	.13	.01	-.06
Competence	.22	.18	.20	.20	<b>.54</b>	<b>.53</b>	.20	.32	.17	.12
Order	-.01	-.11	-.05	-.06	<b>.55</b>	<b>.60</b>	-.15	-.12	-.05	-.18
Dutifulness	.00	.03	.23	.18	<b>.51</b>	<b>.52</b>	.05	.17	.00	-.08
Achievement Striving	.22	.17	.08	.10	<b>.55</b>	<b>.47</b>	-.07	-.03	.06	-.02
Self-Discipline	.15	.13	.22	.21	<b>.69</b>	<b>.65</b>	.18	.17	-.07	-.03
Deliberation	-.01	.02	.30	.28	<b>.42</b>	<b>.40</b>	.01	.12	-.05	-.10

Table 9 (continued)

NEO-PI-R	Extraversion		Agreeableness		Conscientiousness		Emotional Stabil.		Openness	
	BFI	TIPI	BFI	TIPI	BFI	TIPI	BFI	TIPI	BFI	TIPI
Neuroticism	-.16	-.13	-.26	-.22	-.31	-.26	-.66	-.66	-.06	-.10
Anxiety	-.09	-.08	-.09	-.06	-.12	-.10	-.61	-.53	-.18	-.16
Angry Hostility	-.02	.01	-.47	-.45	-.20	-.14	-.42	-.46	-.01	-.08
Depression	-.21	-.17	-.17	-.10	-.36	-.34	-.53	-.57	-.05	-.09
Self-consciousness	-.27	-.23	-.03	-.01	-.25	-.23	-.36	-.35	-.02	-.11
Impulsiveness	.00	.00	-.26	-.21	-.19	-.14	-.46	-.42	.10	.06
Vulnerability	-.11	-.11	-.18	-.19	-.27	-.23	-.61	-.66	-.08	-.05
Openness	.21	.23	.00	.07	-.19	-.18	-.12	-.09	.68	.56
Fantasy	-.05	.00	-.09	-.04	-.31	-.27	-.12	-.06	.44	.34
Aesthetics	.24	.24	.11	.12	-.12	-.08	-.20	-.13	.58	.40
Feelings	.33	.32	.09	.13	.07	.10	-.25	-.24	.34	.31
Actions	.25	.23	.02	.06	-.08	-.06	.06	-.01	.44	.51
Ideas	.08	.14	-.08	.02	-.17	-.17	.01	.00	.62	.41
Values	.03	.02	-.04	.01	-.09	-.19	.05	.07	.21	.28

Note. N = 172. BFI, Big-Five Inventory (John & Srivastava, 1999); TIPI, Ten-Item Personality Inventory; NEO-PI-R (Costa & McCrae, 1992). Convergent correlations shown in bold typeface. The BFI and TIPI were administered approximately 6 weeks before the NEO-PI-R. Inter-scale correlations for the NEO-PI-R ranged from .01 (between Neuroticism and Openness) to -.32 (between Conscientiousness and Neuroticism), with an absolute mean of .17. Emotional Stabil., Emotional Stability.

### 3.2.3. *External correlates*

To test whether the patterns of external correlates of the TIPI matched the patterns of external correlates of the BFI, we correlated both instruments with each of the other constructs assessed in the testing battery. These patterns of external correlates are shown separately for each of the Big-Five dimensions in Table 4. As in Study 1, we tested whether the patterns of correlations were similar across instruments, by computing column-vector correlations for each of the five dimensions. As shown by the column-vector correlations in Table 4, the TIPI displayed patterns of correlations that were virtually identical to those of the BFI; with all column-vector correlations exceeding .90. As before, the BFI yielded correlations that were slightly stronger than the TIPI.

## 4. Discussion

### 4.1. *Summary of findings*

Although somewhat inferior to the standard Big-Five instrument examined here, the FIPI and TIPI reached adequate levels in each of the criteria against which it was evaluated: convergent and discriminant validity, test–retest reliability, patterns of external correlates, and (for the FIPI) convergence between self- and observer-ratings. In both measures, Extraversion fared the best across the criteria and Openness and Agreeableness fared least well. Overall our findings suggest that these very brief instruments can stand as reasonable proxies for longer Big-Five instruments, especially when research conditions dictate that a very short measure be used. Of the two instruments, the 10-item instrument is psychometrically superior, it can be used for latent variable modeling, it allows researchers to assess for acquiescence bias and check for errors, and it takes no longer to complete than the 5-item instrument (about 1 min). Should a short instrument be needed, this is the one we recommend.

### 4.2. *Limitations*

Despite the evidence for the value of the TIPI, very short measures are subject to some serious limitations. The first limitation is the psychometric cost of using short measures. Compared with standard multi-item measures of the Big Five, the TIPI is less reliable and correlates less strongly with other variables.

The second limitation of brief measures is their inability to measure individual facets of multi-faceted constructs. The Big-Five dimensions are essentially and explicitly broad constructs encompassing several related but separable facets. Very brief measures, such as the ones examined here, do not provide scores for the narrower facet-level constructs and a much greater investment is needed to obtain facet scores (e.g., by using the 240-item NEO-PI-R). Moreover, by using a short instrument to measure potentially separable facets, the single-item approach

neglects the recent call to turn attention from broad factors to more specific facets, which are often better predictors of specific criteria (Paunonen & Ashton, 2001). However, it is worth noting that the other widely used short measures of the Big Five (e.g., the 44-item BFI and the 60-item NEO-FFI) do not provide facet scores either. Indeed, by using only 10 items (rather than 44 or 60) to measure the Big Five, the TIPI may leave researchers more space and time to focus on other measures of narrow constructs that are more directly related to their research questions.

#### *4.3. Benefits*

The central benefit of the TIPI is that it extends the scope of studies in which the Big Five can be measured. As Saucier (1994) noted in the context of his 40-item instrument, “The availability of this extremely short set of Big-Five markers widens the potential application of the Big Five to assessment situations where brevity is an unusually high priority.” (p. 515). This benefit is particularly appropriate now that the Big-Five framework has been well established; the focus of personality research is now free to shift from the psychometric and structural properties of the Big Five to focus on relations between the Big-Five dimensions and other constructs and outcomes. Whereas the early structural studies needed to devote extensive resources to assessing the Big Five, newer studies can afford to use shorter measures. Indeed, the development of several medium-length instruments such as the BFI, the NEO-FFI, and Saucier’s (1994) marker set reflect this change in research emphasis. Although these instruments require far less time than the full NEO-PI-R, they are still too time-consuming for some research purposes. In cases where researchers are willing to tolerate the somewhat diminished psychometric properties associated with even briefer measures, we offer the TIPI as one potential research tool.

A second potential benefit of the TIPI is that by providing a standard instrument for use by the research community, knowledge about its psychometric properties and its external correlates can accumulate. Without a standard instrument, researchers would be forced to create their own measures, and could not benefit from the work done by others.

A third benefit is that very brief measures eliminate item redundancy, reducing participant boredom and the oft-expressed frustration about “answering the same question again and again.” This benefit can ameliorate the psychometric costs of short measures (Burisch, 1984a, 1984b), and may explain why the TIPI performed so well.

## **5. Conclusion**

Harm-reduction programs such as methadone clinics and condom distribution programs do not condone harmful behaviors but acknowledge that such behav-

iors exist and the programs are designed to alleviate the negative consequences of the behaviors. Analogously, we do not encourage the use of very brief measures, but we acknowledge that when brevity is a high priority, researchers may be driven to create their own very short measures of the Big Five or, even worse, to use no measure at all. For such situations we provide the TIPI, for which the psychometrics are known and are reasonable. Of course, most harm-reduction programs run a risk of backfiring—the possibility that the very existence of the programs will actually increase the frequency of the behaviors they are designed to make safer. Analogously, we hope that this instrument will not be used in place of established multi-item instruments. Instead, we urge that this instrument be used when time and space are in short supply and when only an extremely brief measure of the Big Five will do. Under such circumstances the TIPI, which takes only a minute to complete, represents a sensible option.

#### Appendix A. Ten-Item Personality Inventory-(TIPI)

Here are a number of personality traits that may or may not apply to you. Please write a number next to each statement to indicate the extent to which *you agree or disagree with that statement*. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.

Disagree strongly	Disagree moderately	Disagree a little	Neither agree nor disagree	Agree a little	Agree moderately	Agree strongly
1	2	3	4	5	6	7

*I see myself as:*

1. \_\_\_\_ Extraverted, enthusiastic.
2. \_\_\_\_ Critical, quarrelsome.
3. \_\_\_\_ Dependable, self-disciplined.
4. \_\_\_\_ Anxious, easily upset.
5. \_\_\_\_ Open to new experiences, complex.
6. \_\_\_\_ Reserved, quiet.
7. \_\_\_\_ Sympathetic, warm.
8. \_\_\_\_ Disorganized, careless.
9. \_\_\_\_ Calm, emotionally stable.
10. \_\_\_\_ Conventional, uncreative.

TIPI scale scoring (“R” denotes reverse-scored items): Extraversion: 1, 6R; Agreeableness: 2R, 7; Conscientiousness: 3, 8R; Emotional Stability: 4R, 9; Openness to Experiences: 5, 10R.

## Appendix B

## Normative data for the Ten-Item Personality Inventory (TIPI): Self-reported data

Ethnicity	Whole sample					Female					Male				
	E	A	C	ES	O	E	A	C	ES	O	E	A	C	ES	O
All ethnicities	(N = 1813)														
<i>Mean</i>	4.44	5.23	5.40	4.83	5.38	4.54	5.32	5.51	4.66	5.40	4.25	5.06	5.19	5.13	5.34
<i>SD</i>	1.45	1.11	1.32	1.42	1.07	1.47	1.11	1.11	1.45	1.06	1.41	1.10	1.15	1.31	1.09
White ethnicities	(N = 1173)														
<i>Mean</i>	4.56	5.26	5.47	4.85	5.43	4.68	5.36	5.56	4.65	5.45	4.30	5.05	5.27	5.27	5.39
<i>SD</i>	1.48	1.12	1.13	1.45	1.06	1.47	1.12	1.09	1.46	1.04	1.15	1.11	1.17	1.32	1.09
Hispanic ethnicities	(N = 760)														
<i>Mean</i>	4.43	5.21	5.45	4.90	5.53	4.45	5.32	5.51	4.75	5.58	4.41	4.99	5.34	5.16	5.44
<i>SD</i>	1.41	1.09	1.11	1.42	1.04	1.41	1.10	1.12	1.49	1.03	1.43	1.07	1.09	1.25	1.06
Asian ethnicities	(N = 146)														
<i>Mean</i>	4.12	5.14	5.11	4.64	5.07	4.19	5.23	5.26	4.60	5.07	4.03	5.01	4.90	4.70	5.07
<i>SD</i>	1.31	1.06	1.16	1.32	1.08	1.37	1.03	1.16	1.40	1.08	1.24	1.08	1.13	1.22	1.08
Black ethnicities	(N = 191)														
<i>Mean</i>	4.38	5.37	5.57	5.14	5.53	4.27	5.14	5.67	4.87	5.23	4.57	5.73	5.40	5.60	6.02
<i>SD</i>	1.46	1.17	1.03	1.38	1.05	1.46	1.20	0.92	1.36	1.06	1.47	1.06	1.19	1.32	0.83
Other ethnicities	(N = 35)														
<i>Mean</i>	4.07	5.21	5.34	4.89	5.43	4.21	5.16	5.56	4.83	5.51	3.78	5.33	4.88	5.03	5.28
<i>SD</i>	1.54	1.16	1.06	1.25	1.12	1.64	1.23	1.03	1.29	1.11	1.30	1.03	0.99	1.18	1.15
	(N = 41)														

Note. E, Extraversion; A, Agreeableness; C, Conscientiousness; ES, Emotional Stability; O, Openness.

## References

- Aron, A., Aron, E. N., & Danny, S. (1992). Inclusion of Other in Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*, 596–612.
- Beck, A. T. (1972). *Depression: Causes and treatments*. Philadelphia: University of Pennsylvania Press.
- Benet-Martínez, V., & John, O. P. (1998). 'Los Cinco Grandes' Across cultures and ethnic groups: Multitrait-multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology*, *75*, 729–750.
- Benet-Martínez, V., Leu, J., Lee, F., & Morris, M. (2002). Negotiating biculturalism: Cultural frame-switching in biculturals with 'Oppositional' vs. 'Compatible' cultural identities. *Journal of Cross-Cultural Psychology*, *33*, 492–516.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, *117*, 187–215.
- Brown, R. P., & Josephs, R. A. (1999). The mathematics identification questionnaire. The University of Texas at Austin, Unpublished manuscript.
- Burisch, M. (1984a). Approaches to personality inventory construction. A comparison of merits. *American Psychologist*, *39*, 214–227.
- Burisch, M. (1984b). You don't always get what you pay for: Measuring depression with short and simple versus long and sophisticated scales. *Journal of Research in Personality*, *18*, 81–98.
- Burisch, M. (1997). Test length and validity revisited. *European Journal of Personality*, *11*, 303–315.
- Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). *The quality of American life: Perceptions, evaluations, and satisfactions*. New York: Russell Sage.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, *95*, 542–575.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, *4*, 26–42.
- Hazan, C., & Shaver, P. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology*, *52*, 511–524.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 102–138). New York: Guilford Press.
- Pervin, L. A., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In O. P. John (Ed.), *Handbook of personality theory and research* (pp. 139–153). New York: Guilford Press.
- Paulhus, D. L., & Bruce, M. N. (1992). The effect of acquaintanceship on the validity of personality impressions: A longitudinal study. *Journal of Personality and Social Psychology*, *63*, 816–824.
- Paulhus, D. L., Lysy, D. C., & Yik, M. S. M. (1998). Self-report measures of intelligence: Are they useful as proxy measures of IQ? *Journal of Personality*, *64*, 525–555.
- Paunonen, S. V., & Ashton, M. C. (2001). Big Five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology*, *81*, 524–539.
- Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*, 741–763.
- Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, *84*, 1236–1256.
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001a). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem scale. *Personality and Social Psychology Bulletin*, *27*, 151–161.



- Robins, R. W., Tracy, J. L., Trzesniewski, K. H., Potter, J., & Gosling, S. D. (2001b). Personality correlates of self-esteem. *Journal of Research in Personality*, *35*, 463–482.
- Robins, R. W., Trzesniewski, K. H., Tracy, J. L., Gosling, S. D., & Potter, J. (2002). Self-esteem across the lifespan. *Psychology and Aging*, *17*, 423–434.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, *57*, 493–502.
- Sandvik, E., Diener, E., & Seidlitz, L. (1993). Subjective well-being: The convergence and stability of self-report and non-self-report measures. *Journal of Personality*, *61*, 317–342.
- Saucier, G. (1994). Mini-markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment*, *63*, 506–516.
- Swann, W. B., Jr., & Rentfrow, P. J. (2001). Blirtatiousness: Cognitive, behavioral, and physiological consequences of rapid responding. *Journal of Personality and Social Psychology*, *81*, 1160–1175.