RUNNING HEAD: Revised Italian TIPI

Psychometric properties of a revised version of the Ten Item Personality Inventory

Carlo Chiorri, Fabrizio Bracco, Tommaso Piccinno, Cinzia Modafferi, Valeria Battini

Department of Educational Sciences, Psychology Unit, University of Genoa, Italy

Correspondence to:

Carlo Chiorri, PhD
Department of Educational Sciences, Psychology Unit
University of Genova
Corso A. Podestà, 2,
16128 Genova (Italy)
e-mail: carlo.chiorri@unige.it
Tel. +3901020953726 / +3901020953726
Tel. direct +3901020953709
Fax +3901020953728

Please cite this paper as:

Chiorri, C., Bracco, F., Piccinno, T., Modafferi C., & Battini, V. (2014). Psychometric Properties of a Revised Version of the Ten Item Personality Inventory. *European Journal of Psychological Assessment*, DOI: 10.1027/1015-5759/a000215.

Summary

Gosling, Rentfrow, and Swann (2003) developed the Ten-Item Personality Inventory (TIPI) to meet the need of very short measures of the Big Five for time-limited contexts or large survey questionnaires. In this paper we show the inadequacy of the Italian version downloadable from Gosling's website and we report the results of four studies in which the psychometric properties of a revised version (I-TIPI-R) were investigated in student and general population samples. This new version showed adequate factor structure, test-retest reliability , self-observer agreement and convergent and discriminant validity with the Big Five Inventory (BFI). Moreover, I-TIPI-R and BFI scores did not differ in their correlations with measures of affect, self-esteem, optimism, emotion regulation and social desirability. Overall, the results suggest that the I-TIPI-R can be considered a valid and reliable alternative to the BFI for the assessment of basic personality traits when very short measures are needed.

Keywords: Five-Factor Model; Ten-Item Personality Inventory; Personality Assessment; Short Measure; Construct Validity.

Introduction

The Five Factor Approach to personality (FFA; as in Block 2010, we use this generic term since it is not related to any specific group of researchers or instruments) has been accepted as a comprehensive conceptual framework for integrating all the research findings and theories in personality psychology, although some contrarian views do exist (Block, 2010). It assumes that individual differences in adult personality characteristics can be organized in terms of five broad trait domains: Extraversion (E), Agreeableness (A), Conscientiousness (C), Neuroticism (N), and Openness to Experience (O). The predictive validity of personality traits on important life outcomes, such as health and longevity, marital success, and educational and occupational attainment has been repeatedly shown (e.g., Ozer & Benet-Martinez 2006), thus suggesting the need to incorporate measures of personality into social surveys (e.g., Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). However, such studies usually require assessments on many constructs and, since most of them have a longitudinal design, participants are asked to complete the measures more than once. As pointed out by Credé, Harms, Niehorster and Gaye-Valentine (2012), respondents can be easily bored and/or fatigued by completing long surveys and this affects their likelihood to attend to item-content with care or to agree to participate in follow-up data collections. From this perspective, the inclusion of a 240-item personality measure such as the NEO Personality Inventory-Revised (NEO-PI-R; Costa & McCrae, 1992), notwithstanding its widely supported validity and reliability, does not seem feasible, and shorter (1 to 8 items per scale) inventories have been developed (for a review, see, e.g., Credé et al., 2012). These scales take little time to complete and are useful not only in addressing the aforementioned issues, but also in providing a measure of personality traits in organizational, clinical or research settings in which assessment time is limited and the choice is between assessing personality with a very short inventory or not assessing personality at all. A number of studies showed that the alleged psychometric shortcomings of short inventories might

actually be limited (for a review, see, e.g., Credé et al., 2012), but in terms of content validity and measurement performance the shortening cannot come at no cost. Short scales are likely to be characterized by substantial content deficiency (e.g., Smith, McCarthy, & Anderson, 2000) and cannot benefit from the minimization of the amount of (random) measurement error through the averaging of responses across multiple items that assess the same construct.

The Ten Item Personality Inventory (TIPI, Gosling, Rentfrow & Swann, 2003) was developed using descriptors from other well-established Big Five instruments. Each item consists of two descriptors, separated by a comma, using the common stem "I see myself as:". Each of the ten items is rated on a 7-point scale ranging from 1 (strongly disagree) to 7 (strongly agree) and the mean administration time is one minute. Gosling et al. (2003) argued that the inventory based on two items per scale could not reach the commonly accepted levels of internal consistency, which is function of mean inter-item correlation and, above all, of the number of items. To tackle these problems they could have used items with a very high correlation (e.g., $r > .70$), which, given their unavoidable redundancy, would have undermined content coverage. Instead, they emphasized content validity considerations, although this resulted in lower inter-item correlations (E=.59, A=.36, C=.42, N=.61, O=.28) and therefore lower Cronbach's αs (E=.68, A=.40, C=.50, N=.73, O=.45) than those typical of more homogeneous scales. They also suggested that researchers wishing to correct TIPI correlations for unreliability should base their corrections on the test-retest reliability coefficient, which has also been recently shown to be more predictive of score validity than estimates of internal consistency (McCrae, Kurtz, Yamagata, & Terracciano, 2011).

In their seminal study, Gosling et al. (2003) also reported promising results with regard to TIPI convergent and discriminant validity in comparison to the Big Five inventory (BFI, John, Donahue, & Kentle, 1991) and the NEO-PI-R. Subsequently, Donnellan, Oswald, Baird, and Lucas (2006) reported "reasonable convergent validity" (p. 196) of the TIPI with their Mini-

IPIP and other measures (50-item IIP-FFM and 120-item IPIP-NEO) from the International Personality Item Pool website (Goldberg, Johnson, Eber, et al., 2006) and Furnham (2008) provided evidence of the convergence of TIPI with the 60-item NEO-Five Factor Inventory (Costa & McCrae, 1992). Erhart, Erhart, Roesch, Chung-Herrera, Nadler and Bradshaw (2009) also found that the correlations of four out of five TIPI scales with measures of locus of control, self-monitoring, life satisfaction and trait anxiety did not significantly differ from those of the IIP-FFM (average difference .052), whereas the results for Agreeableness were somewhat less favorable, as the average difference was larger (.089) and the three pairs of correlations that were significantly different in the validity analyses were all for the Agreeableness measure. TIPI translations in other languages can be freely downloaded from Samuel Gosling's website (http://homepage.psy.utexas.edu/homepage/faculty/gosling/scales_we.htm) and researchers might be tempted to use them for cross-cultural studies. However, such studies can provide valid results only as long as the measures employed to assess personality traits have adequate psychometric properties and measurement invariance across cultures (Benet-Martinez, 2007). It must be noted that not all the TIPI versions available on the website have undergone a formal assessment of their psychometric properties. The few validation studies of the TIPI in other languages (German: Muck, Hell, & Gosling, 2007; Dutch: Hofmans, Kuppens, & Allik, 2008, and Spanish: Romero, Villar, Gómez-Fraguela, & López-Romero, 2012) found adequate psychometric properties in terms of factor structure and content domain representation, but the suitability of many of the other versions as an adequate measuring instrument of the Big Five is still unclear. The website provides an Italian version (I-TIPI) but no information about its psychometric properties is provided, nor are the authors aware of any validation studies in the published literature. In this paper we present a series of studies aimed at investigating the psychometric properties of the Italian TIPI in terms of factor structure, test-retest reliability and construct validity. Specifically, we tested whether the

correlations of TIPI scores with measures of psychological functioning differ from those of the BFI. The lack of differences would imply that, when administration time is an issue, the 44-item BFI could be replaced by the 10-item TIPI with negligible loss in terms of construct and predictive validity.

Study 1

Rationale

In this study we used exploratory factor analysis (EFA) to test the factorial validity of the I-TIPI. Were the expected five-factor structure supported, we could go on and test the other psychometric properties of the inventory. Otherwise, a thorough examination of the possible causes of a lack of conformity to the expected structure would be needed.

Participants

We administered the I-TIPI along with a short socio-demographic schedule, to 189 participants (females = 72%, mean age 28±9 years, range 18-65, median educational level high school diploma) who had responded to advertisements requesting potential volunteers for psychological studies to test its factor structure.

Materials

We noted that I-TIPI instructions were not consistent with the English version, since the opening sentence "Adesso le leggerò alcuni tratti di carattere che possono corrispondere o no a lei" (literally: "Now I will read you some character traits that may or may not apply to you″) is not a correct translation of the original "Here are a number of personality traits that may or may not apply to you". Besides, it implies that the items cannot be read from a hard copy of the questionnaire, but read by an interviewer. Hence, we modified the first words with "Qui di seguito troverà" ("Here you will find"), leaving the remainder unchanged.

Results

Grounding on the theoretical structure of the TIPI, we performed a principal axis (exploratory) factor analysis setting to five the number of factors to extract, Promax-rotated. Eigenvalues began to level off after four, and not five, factors (eigenvalues were 1.86, 1.78, 1.57, 1.18, 0.85, 0.81, 0.66, 0.45, 0.43, 0.41) and the pattern matrix (available from the corresponding author) showed that Extraversion items loaded substantively (i.e., higher than |.30|) on the first factor, but an Agreeableness item loaded substantively only on the second factor together with Neuroticism items. Three items of three different scales loaded substantively on the third factor, while only one item loaded on the fourth factor. The Conscientiousness items loaded on the fifth factor, but item 3 was also loading on the third factor. Although Cronbach's αs and correlations among items of the same scale were consistent with literature (E: α=.65, *r*=.48; A: α =.23, *r*=.14; C: α=.44, *r*=.31; N: α=.39, *r*=.24), the factor solution did not appear to be adequate.

Revision of the Italian TIPI

One reason why we did not obtain the expected factor structure could be the quality of translation. In a previous adaptation study of the TIPI, Muck et al. (2007) pointed out how the shortcomings of a first German version of the TIPI (Herzberg & Brähler, 2006) could be the result of the particular translation of the TIPI rather than of the TIPI itself. We thus introduced the original I-TIPI to a class of undergraduate psychology students (*n*=69, females = 79%, mean age 22±4 years, range 18-41) during a seminar about assessment of personality traits. Students were invited to complete the scale and then asked their opinion about it. In the I-TIPI file on Gosling's website it is reported that "the translations were done by a very large and reputable global market research company and double checked by a second set of native speakers". Though substantially correct, this version appears to have a number of shortcomings that emerged in the discussion carried out with the students (see below) and that suggested the

need of a new translation of the instructions and of the critical items. The first two authors of this article translated independently into Italian the TIPI. Their first language is Italian, but they both have a fluent understanding of the English language. Next, they consensually derived a combined version of the inventory. To address inconsistencies in their translations, they and an Italian-English bilingual speaker considered the meaning of the original English items and the consistency of the new translation with the content domain of each Big Five and a comprehensive list of Italian adjectives used in Big Five studies and reported in DiBlas and Perugini (2002). Subsequently, another Italian-English bilingual speaker blind to the original version back-translated the revised Italian TIPI to English.

Beyond the amendment made before the preliminary administration, students reported that instructions did not actually "sound" Italian, and appeared overly complex. Hence, we reworded them as "Per favore, leggi le seguenti caratteristiche di personalità e indica quanto ti senti descritto da ogni coppia di aggettivi, anche se pensi che una delle due caratteristiche ti descriva meglio dell'altra, utilizzando la seguente scala" ("Please read the following personality traits and rate how well each pair of adjectives describes you, even if you think that one characteristic describes you better than the other, using the following scale"; the seven-point Likert-type scale and its anchors were shown below the instructions).

As for the items (see Table 1), in the web Italian version all the adjectives were correctly translated from a semantic point of view, but some of them raised some issues.

[Insert Table 1 about here]

Enthusiastic (item 1) can be translated "entusiasta", but in Italian being "entusiasta" implies that there must be something or someone toward which one feels this way and it is more a state than a trait. The new translation "esuberante" unequivocally implies enthusiasm, vivacity and expansiveness from a trait point of view. *Critical* (item 2) has been translated "critica", but in everyday language nobody would use this adjective to describe herself, since, again, it means

that there must be something or someone toward which the feeling is addressed. Conversely, the new translation "polemica" can sometimes be used as a synonym for "critical", but implies also being adversarial, not cooperative. *Easily upset* (item 4) can be translated "facile da sconvolgere", but "sconvolgere" implies a much stronger feeling than the English "upset", which can be better translated with "si agita", that implies being upset and worrying. *Complex* can be translated "eclettico", which is correct but the word is not very frequently used - some university students could not even define it correctly. The new translation "Con molti interessi" has the same meaning, but it is far more common. *Careless* (item 8) can be translated "noncurante", but this term is very rarely used for self-description. The new translation "distratta" has the same implications of "careless" and describes someone who pays little attention to detail, forgets assignments and appointments, etc. Lastly, both "comune" and "non particolarmente creativa" appeared inadequate translations of "conventional" and "uncreative" (item 10), respectively, since in current Italian they are commonly used as derogatory terms − not surprisingly, this item loaded on a separate factor. The new translation "tradizionalista" implies following traditions in terms of following cultural norms in, e.g., mating, dressing and eating behavior, whereas "abitudinaria" implies being routine-bound, with little tendency to introduce any sort of change in one's life. The final revised Italian version of the TIPI (I-TIPI-R) is reported in Table 1. The back-translated items were (1) extroverted, exuberant; (2) difficult, adversarial; (3) trustworthy, self-disciplined; (4) worried, anxious; (5) open to new experiences, with many interests; (6) reserved, silent; (7) understanding, affectionate; (8) disorganized, absent-minded; (9) calm, emotionally stable; (10) traditionalist, routine-bound. This result suggests that an adequate trade-off between semantic (i.e., the meaning of the source language item is preserved in the translation) and conceptual (i.e., the degree to which a concept, independent of the words used to operationalize it, exists in the same form in the source and target cultures) equivalence (Behling & Law, 2000) had been achieved. Subsequent

studies aimed at investigating the psychometric properties of this revised version and, specifically, its adequacy as a measure of the Big Five.

Study 2

Rationale

In this study we used exploratory factor analysis (EFA) to test the factorial validity of the I-TIPI-R developed in Study 1 in a group of undergraduate students.

Participants and procedure

A group of 157 undergraduate students of psychology (78% females, mean age 21±5 years, range 18-42) of a North-western Italian university completed the I-TIPI-R, along with a short socio-demographic form, for course credit in a personality assessment class.

Results

As in Study 1, we performed a principal axis (exploratory) factor analysis setting to five the number of factors to extract, Promax-rotated. Results are reported in Table 1. The items were grouped according to the expected five-factor structure. The highest cross-loading was −.29 of item 2 (A) on the Extraversion factor. Eigenvalues began to level off after the fifth factor (2.44, 1.90, 1.37, 1.04, 0.89, 0.57, 0.56, 0.53, 0.38, 0.32). The pattern of factor scores correlations did not appear consistent with the hypothesis of orthogonal factors, since some substantial ($r >$ |.30|) correlations were found (mean inter-factor absolute correlation = .22, after *r*-to-*z* transformation and back-transformation). The average internal consistency index was .57, similar to the .55 average reported by Gosling et al. (2003).

Study 3

Rationale

The I-TIPI-R was developed through discussions with undergraduate students (Study 1) and Study 2 tested its factor structure on a convenience sample drawn from the same population. Since this might have undermined the generalizability of results, in Study 3 we administered

the I-TIPI-R to a sample of volunteers from the general population and tested its factor structure and congruence with Gosling et al. (2003)'s data.

Participants and procedure

A total of 472 participants (52% females, mean age 35±12 years, range 18-77, median educational level high school diploma) took part in the study. They were mostly workers with an open-ended contract (42%) and students (28%), singles (54%) and married/cohabiting (40%). They were recruited and tested with the same procedure of Study 1.

Results

Factor structure

The same statistical analyses of Study 2 were carried out. Results are reported in Table 1. The items were again grouped according to the expected five-factor structure, although eigenvalues began to level off after the sixth, and not the fifth factor (1.99, 1.79, 1.39, 1.11, 0.99, 0.85, 0.62, 0.46, 0.41, 0.39). The higher cross-loading was −.30 of item 9 (N) on the Agreeableness factor. Also in this study the pattern of factor scores correlations did not appear to be consistent with the hypothesis of orthogonal factors (mean inter-factor absolute correlation = .13). The average internal consistency index was .55, similar to Study 2 and previously published studies. We used the pattern matrices in Table 1 to compute congruence coefficients (CCs), which can be interpreted as standardized measures of proportionality of elements in both matrices. Following Lorenzo-Seva and ten Berge (2006), coefficients in the range .85-.94 correspond to a fair similarity, while values higher than .95 implied that the two components compared could be considered equal. Results were E=.99, A=.96, C=.96, N=.99, O=.97, total=.97. Factor intercorrelations were also similar, since their standardized *q* difference did not exceed, in absolute value, .17, which indicates a small effect size (Cohen, 1988). We also computed CCs between our Study 3 and Gosling et al. (2003)'s factor solution[1], derived from

[1] We used the pattern matrices from Promax-rotated Principal Components Analyses, since a Promax-rotated Principal Axis Factoring solution did not converge to an admissible solution with Gosling et al. (2003)'s data.

their correlation matrix. Results were E=.97, A=.82, C=.94, N=.88, O=.90, total=.90. Factor intercorrelations were also similar, since their standardized *q* difference did not exceed, in absolute value, .28, which indicates a small effect size.

Study 4

Rationale

In this study we assessed the construct validity of the I-TIPI-R in terms of its convergence with a well-established but longer measure of the Big Five such as the BFI and of its pattern of associations with social desirability and some important outcomes in psychological functioning such as affect, emotional regulation, self-esteem and optimism (Ozer & Benet-Martinez, 2006). The BFI is a Big Five measure that, although relatively short (44 items), does not sacrifice either content coverage or good psychometric properties, since its items cover most the facets hypothesized by Costa and McCrae (1992) for each Big Five factor. John, Naumann and Soto (2008) reported that BFI scales represent personality at the broadest level of abstraction, each including a large number of distinct and more specific personality and research has shown that the BFI is a measure of the core attributes of the Big Five that is at least as efficient as that of the other longer tools. In their seminal study on the TIPI, Gosling et al. (2003) found that convergent correlations with the BFI were E=.87, A=.70, C=.75, N=.81, O=.65 and absolute discriminant correlations were no higher than .36.

As for association with other constructs, we expected to replicate previous results reported in literature. Social desirability is known to be associated with lower scores in Neuroticism (e.g., Konstabel, Aavik, & Allik, 2006) and with higher scores in Agreeableness (e.g., Graziano & Tobin, 2002). It has been reported that in Italian participants positive and negative affects were positively associated with Extraversion and Neuroticism, respectively (Terracciano, McCrae & Costa, 2003), that reappraisal strategies of emotional regulation were positively associated with all the Big Five except Neuroticism, for which the association was

negative, and that emotion suppression strategies were negatively associated with all the Big Five (Balzarotti, John, & Gross, 2010). Jonason, Teicher and Schmitt (2011) recently showed that TIPI Extraversion and Conscientiousness scores were positively associated with self-esteem (albeit weakly), whereas Neuroticism scores were moderately associated with lower self-esteem. Finally, Sharpe, Martin and Roth (2011) recently reported positive associations (in the .20-.30s) of optimism with all TIPI scales except for Neuroticism, for which the correlation was negative.

The usefulness of a shorter measure can be supported as long as it allows predictions that are not substantially different from those of a longer one. Hence, we also tested whether the associations of TIPI scores with the aforementioned constructs differed from those of the BFI. In Gosling et al. (2003)'s seminal study, the patterns of the TIPI external correlates with variables such as self-esteem, social dominance, math identification, or political values were consistent with those of the BFI, although the size of these correlations was slightly stronger for the BFI.

Participants

A group of 66 undergraduate students of psychology (82% females, mean age 25±8 years, range 19-59) of a North-western Italian university took part in this study for course credit in a research method master course.

Measures

TIPI. As described beforehand.

Big Five Personality Inventory (BFI, John et al., 1991; Italian version in Ubbiali, Chiorri & Donati, in press). The BFI is a 44-item self-report inventory consisting of short phrases that include trait adjectives known to be prototypical markers of the Big Five. A common stem ("I see myself as someone who…") is used for all items, which are made up of one or two prototypical trait adjectives used as the item core, with added elaborative, clarifying, or

contextual information (e.g., "Remains calm in tense situations", "Is sometimes shy, inhibited"). Participants are asked to rate the degree to which each item applies to their personality on a 5-point, Likert-type scale.

Positive and Negative Affect Schedule (PANAS, Watson, Clark & Tellegen, 1988, Italian version in Terracciano, et al., 2003). The PANAS is a 20-item self-report adjective checklist to measure positive (e.g., "Determined", "Inspired") and negative (e.g., "Distressed", "Jittery") affect. Participants were asked to report the frequency (from never to always) of their affect over an extended period of time, i.e., how they felt "in general" on a 5-point, Likert-type scale.

Rosenberg's Self-Esteem Scale (RSES, Rosenberg, 1965; Italian version in Prezza, Trombaccia & Armento, 1997). RSES is a 10-item self-report measure of global self-esteem (e.g., "I feel that I have a number of good qualities"). Participants are asked to indicate the degree to which each item applies to them on a 4-point, Likert-type scale.

Life Orientation Test-Revised (LOT-R, Scheier, Carver & Bridges, 1994; Italian version in Giannini, Schuldberg, Di Fabio & Gargaro, 2008). LOT-R is a 10-item (4 filler items) measure of individual differences in generalized optimism versus pessimism (e.g., "In uncertain times, I usually expect the best"). Participants are asked to report the extent of their agreement with each item on a 5-point, Liker-type scale.

Emotional Regulation Questionnaire (ERQ, Gross & John, 1993; Italian version in Balzarotti et al., 2010). The ERQ is a 10-item measure of two emotion regulation strategies: Reappraisal (6 items, e.g., "I control my emotions by changing the way I think about the situation I’m in") and Suppression (4 items, e.g., "I control my emotions by not expressing them"). Participants are asked to rate each item on a 7-point, Liker-type agreement scale.

Marlowe-Crowne Social Desirability Scale - Short Form (MCSDS-SF, Crowne & Marlowe, 1960, Manganelli-Rattazzi, Canova & Marcorin, 2000). MCSDS-SF is a 9-item version of the original Crowne and Marlowe's self-report scale designed to measure social desirability

independent of psychopathology. Participants are asked to indicate whether statements describing desirable but uncommon behaviors (e.g., admitting mistakes) or undesirable but common behaviors (e.g., lying) are true or false for them.

Procedure

Participants completed the battery in a single session as part of a class activity. The scales included in the battery were administered in counterbalanced fashion to control for order and sequence effects. To ensure anonymity, participants were instructed to generate their unique identification code using their parents' initials and the last two digits of their phone number. This allowed to match the data collected 4 weeks later when the TIPI was administered again for testing test-retest reliability.

Results

Table 2 shows the correlations among the measures employed in this study.

[Insert Table 2 about here]

We first examined convergent and discriminant correlations of TIPI (Time 1 scores) and BFI. $Z_{contrast}$ tests (Westen & Shedler, 2003) showed that convergent correlations were significant and significantly higher than discriminant correlations (E: z=6.48, p<.001; A: z=5.31, p<.001; C: z=7.12, p <.001; N: z=5.26, p <.001; O: z=4.56, p<.001, highest absolute discriminant correlation was .35). TIPI scores showed significant associations also with other constructs. E was associated with emotional regulation measures, A with negative affect, self-esteem and optimism and social desirability, C with positive and negative affect and self-esteem, N with positive and negative affect, self-esteem, optimism and social desirability and O with emotional regulation reappraisal. However, a crucial test for the validity of the TIPI was whether its correlations with other constructs differed from those of the BFI. $Z_{contrast}$ tests showed that this was never the case except for the correlation of openness measures with the Emotional Reappraisal scale of the ERQ (z=2.10, p=.036). However, 35 comparisons were

performed to test the hypothesis, thus yielding a cumulative Type I error rate, initially set at .05, of .83. When correcting for multiple comparisons, no comparison was statistically significant. Average differences were: E: .04, A: .04; C: .02; N: .00; O: .02.

In this study we also tested the temporal stability of TIPI scores at a 4-week interval and results are reported in Table 2. Test-retest correlations ranged from .79 to .90 and mean scores did not significantly differ across administrations.

Study 5

Rationale

In this study we tested self-other agreement of TIPI scores, i.e., the similarity between personality descriptions by the self and by others, which is a further test of construct validity, since it provides evidence of the convergence of scores of the same trait from different sources. As claimed by Funder and West (1993), self-reported personality judgments can be considered accurate to the degree they agree with judgments rendered by others. Other adaptations of the TIPI found adequate convergence between self- and peer-ratings (Muck et al., 2007, Romero et al., 2012).

Participants and procedure

A group of 73 undergraduate psychology students (52% females, mean age 28±9 years, range 20-62) were administered the TIPI and were instructed to ask people who knew them well to rate them on the TIPI scale in order to assess self-other convergence. The rating was done by a friend in 43% of cases, by a partner in 33% and by a close relative in the remainder. The mean length of acquaintance was 11.92±8.29 years, and the mean observers' rating of how well they thought they knew the participants was 8.11±1.30 on a 10-point scale.

Results

Results are reported in Table 2. Self-observer correlations of corresponding scales ranged from .42 to .66, whereas self-observer correlations among different scales ranged from −.32 to .23.

We used the $Z_{contrast}$ test to test whether the correlation of each self-report TIPI scale score with the corresponding observer scale score was higher than the correlation with non-corresponding scales. Results showed that this was always the case (E: z=4.40, p<.001; A: z=3.24, p=.001; C: z=5.22, p<.001; N: z=6.36, p<.001; O: z=3.75, p<.001). Mean self-report scores were not statistically different from mean observer scores of the same scale.

Discussion

The studies reported in this paper aimed at investigating the psychometric properties of the Italian version of the TIPI. In Study 1 we did not find evidence of the adequacy of the Italian version currently downloadable from Gosling's website and we performed a theory-driven revision of the translation that lead to develop the I-TIPI-R, which appeared to be a more adequate adaptation of the original TIPI from both a semantic and conceptual equivalence point of view. Two studies (Study 2 and 3) provided evidence of a clear five-factor structure of the revised I-TIPI-R in both undergraduate and general population samples. Congruence coefficients (CCs) also suggested an excellent generalizability of the factor structure. In Study 3 we also computed CCs to investigate the degree of factor similarity of the I-TIPI-R with the original TIPI and results showed that four out of five I-TIPI-R scales had a fair to excellent congruence, except agreeableness, whose CC was slightly lower than the acceptable value. This result might be due to cultural differences between Italy and the US in the importance of the traits as operationalized by the items (see, e.g., Williams, Satterwhite, & Saiz, 2010), but more research is needed to shed light on this issue.

As in previous studies and as we expected given that the TIPI was not designed to meet high standards of internal consistency, but rather to provide a brief measure of the Big Five without sacrificing validity, we observed low alpha reliability values. However, as suggested by Gosling et al. (2003) and McCrae et al. (2011), test-retest reliability coefficients may provide better estimates of reliability, especially for short scales, and Study 4 provided evidence that

for the I-TIPI-R such values were adequate. The inventory also showed convergent and discriminant validity with a widely used measure of the Big Five such as the BFI. Correlations of TIPI scales with social desirability and some key constructs in psychological functioning such as affect, emotional regulation, self-esteem and optimism were generally consistent with literature, although some associations were not as strong as expected (e.g., Extraversion with positive affect or emotional regulation strategies with Agreeableness, Conscientiousness and Neuroticism). However, we found that the correlations of TIPI scores with such measures did not differ from those of the 4.5 times longer BFI, thus suggesting that, when administration time is limited or very short scales are needed, BFI can be replaced by the TIPI with negligible, if at all, loss in terms of construct validity and a substantial gain in administration time.

Finally, in Study 5 we found an adequate agreement between self- and peer-reported TIPI scores, which can be considered as further evidence of construct validity of the inventory. Coefficients were very similar to previous studies and, interestingly, the most notable difference was in the coefficient of neuroticism, which was high in our and in the Spanish study (.66 and .59, respectively), and lower in the German one (.38). Following, e.g., Vazire (2010), we might have expected that agreement should be lower in traits with low observability (e.g., neuroticism) than in traits with high observability (e.g., openness). This was not the case of this and of the Spanish study, whereas the results of the German study were more consistent with US data. Future research is invited to investigate whether cultural differences may exist in the asymmetry of neuroticism judgments, given also the central role of this trait in personality pathology (Widiger, 2009)

Notwithstanding these promising results, some issues about TIPI content validity should not be overlooked. As pointed out by Credé et al. (2012), the poor content validity of short measures of personality traits may result in a general underestimation of the strength of relationships between traits and criteria. Moreover, given the significantly less variance

accounted for by these measures, the incremental variance explained by a variable above and beyond the variance accounted for by scores on Big Five trait measures is likely to be overestimated. Hence, while Italian users might be prompted to use it with reasonable confidence in longitudinal surveys or in experience-sampling studies where short measures appear to be the only viable option, they must be advised that if the individual contribution of each of the trait facets is needed or tapping the core of all Big Five dimensions is crucial, such as for in-depth diagnostic purposes, a more elaborate measure such as the NEO-PI-R should be used. Actually, we did not investigate the associations of TIPI scores with NEO-PI-R facets scores. However, previous studies that did (e.g., Donnellan et al., 2006; Erhart et al., 2009; Gosling et al., 2003, Hofmans et al., 2008) generally found that TIPI scales were, somewhat inevitably, more correlated with some facets than with others within the same domain. Future studies are therefore invited to investigate whether this is true, and in which terms, also for the I-TIPI-R. A more comprehensive test of the usefulness of the TIPI should also involve longitudinal studies to test its predictive ability for some life outcomes (e.g., health or occupational attainment) at a later time and investigations of the biases that may occur in specific contexts (e.g., selection of job candidates).

It must be acknowledged as a limitation that all but one of the studies presented here have been carried out on homogeneous samples, all being psychology students. Future studies on more diverse and representative samples are thus needed, also to investigate the associations of TIPI scores with socio-demographical variables. Although the sample in Study 3 was relatively large, it still was a convenience sample of volunteers, which could not guarantee the same generalizability of results as a nationally representative sample. Notwithstanding these limitations and the need for further research, we provided evidence that the I-TIPI-R has adequate psychometric properties and, as long as short measures of personality are an

acceptable solution, can be confidently administered in research and assessment contexts in which a short Italian measure of the Big Five is needed.

Acknowledgments

The authors would like to thank Mikaela McKellar for her support in the translation of the scale items and the proofreading of the paper. They are also grateful to the editor and the two anonymous reviewers for their valuable comments and suggestions for improving the quality of the paper.

References

Balzarotti, S., John, O. P., & Gross, J. J. (2010). An Italian adaptation of the Emotion Regulation Questionnaire. *European Journal of Psychological Assessment, 26(1),* 61-67.

Behling, O., & Law, K. S. (2000). *Translating Questionnaires and Other Research Instruments: Problems and Solutions* (Sage University Papers Series on Quantitative Applications in the Social Sciences, series no. 07-131). Thousand Oaks, CA: Sage.

Benet-Martínez, V. (2007). Cross-cultural personality research: Conceptual and methodological issues. In R.W. Robins, R.C. Fraley, & R. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 170-189). New York, NY: Guildford Press.

Block, J. (2010). The Five-Factor Framing of Personality and Beyond: Some Ruminations. *Psychological Inquiry*, *21*(1), 2-25.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Ed., NewYork: Lawrence Erlbaum Associates.

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

Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the big five personality traits. *Journal of Personality and Social Psychology*, *102*(4), 874-888

Crowne, D. P., & Marlowe, D. A. (1960). A new scale for social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349−354.

Di Blas, L., & Perugini, M. (2002). L'approccio psicolessicale nella lingua italiana: due studi tassonomici a confronto [The psycholexical approach in Italian language: comparison of two taxonomic studies]. *Giornale Italiano di Psicologia, XXIX*(1), 67-93.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18*(2), 192-203.

Ehrhart, M. G., Ehrhart, K. H., Roesch, S. C., Chung-Herrera, B. G., Nadler, K., & Bradshaw, K. (2009). Testing the latent factor structure and construct validity of the Ten-Item Personality Inventory. *Personality and Individual Differences, 47*(8), 900-905.

Funder, D. C., & West, S. G. (1993). Consensus, self-other agreement, and accuracy in personality judgment: an introduction. *Journal of Personality, 61*, 457-476.

Furnham, A. (2008). Relationship among four big five measures of different length. *Psychological Reports, 102*(1), 312-316.

Giannini, M., Schuldberg, D., Di Fabio, A., & Gargaro, D. (2008). Misurare l'ottimismo: proprietà psicometriche della versione italiana del Life Orientation Test - Revised (LOT-R) [Measuring optimism: psychometric properties of the Italian version of the Life Orientation Test - Revised (LOT-R)]. *Counseling, 1*(1), 73−83.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., et al. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84-96.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big Five personality domains. *Journal of Research in Personality, 37*, 504-528.

Graziano, W. G., & Tobin, R. M. (2002). Agreeableness: dimension of personality or social desirability artifact? *Journal of Personality, 70*(5), 695-727.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*, 348-362.

Herzberg, P.Y., & Brähler, E. (2006). Assessing the Big-Five personality domains via short forms. A cautionary note and a proposal. *European Journal of Psychological Assessment, 22*, 139-148.

Hofmans, J., Kuppens, P., & Allik, J. (2008). Is short in length short in content? An examination of the domain representation of the Ten Item Personality Inventory scales in Dutch language. *Personality and Individual Differences*, *45*(8), 750-755.

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

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114-158). New York, NY: Guilford.

Jonason, P. K., Teicher, E. A., & Schmitt, D. P. (2011). The TIPI's validity confirmed: Associations with sociosexuality and self-esteem. *Individual Differences Research, 9*(1), 52-60.

Konstabel, K., Aavik, T., & Allik, J. (2006). Social desirability and consensual validity of personality traits. *European Journal of Personality*, *20*(7), 549-566.

Lorenzo-Seva, U., & Ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology*, *2*(2), 57-64.

Manganelli-Rattazzi, A. M., Canova, L., & Marcorin, R. (2000). La desiderabilità sociale. Un'analisi di forme brevi della scala di Marlowe e Crowne [Social desirability: analysis of short forms of Marlowe and Crowne's scale]. *Testing, Psicometria e Metodologia, 7*(1), 5−17.

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review, 15(1),* 28-50.

Muck, P. M., Hell, B., & Gosling, S. D. (2007). Construct validation of a short five-factor model instrument: A self-peer study on the German adaptation of the Ten-Item Personality Inventory (TIPI-G). *European Journal of Psychological Assessment, 23*(3), 166-175.

Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology, 57*(1), 401-421.

Prezza, M., Trombaccia, F. R., & Armento, L. (1997). La Scala dell'Autostima di Rosenberg: Traduzione e validazione Italiana [The Rosenberg Self-Esteem Scale: Italian translation and validation]. *Bollettino di Psicologia Applicata, 223,* 35-44.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: the comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*(4), 313-345.

Romero, E., Villar, P., Gómez-Fraguela, J. A., & López-Romero, L. (2012). Measuring personality traits with ultra-short scales: A study of the Ten Item Personality Inventory (TIPI) in a Spanish sample. *Personality and Individual Differences, 53*(3), 289-293.

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

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A re-evaluation of the Life Orientation Test. *Journal of Personality and Social Psychology, 67,* 1063−1078.

Sharpe, J. P., Martin, N. R., & Roth, K. A. (2011). Optimism and the Big Five factors of personality: Beyond Neuroticism and Extraversion. *Personality and Individual Differences*, *51*(8), 946-951.

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, *12*(1), 102-111.

Terracciano, A., McCrae, R. R., & Costa, P. T. Jr. (2003). Factorial and Construct Validity of the Italian Positive and Negative Affect Schedule (PANAS). *European Journal of Psychological Assessment, 19(2),* 131-141.

Ubbiali, A., Chiorri, C., Hampton, P., & Donati, D. (in press). Psychometric properties of the Italian adaptation of the Big Five Inventory (BFI). *Bollettino di Psicologia Applicata,*

Vazire, S. (2010). Who knows what about a person? The Self–Other Knowledge Asymmetry (SOKA) Model. *Journal of Personality and Social Psychology, 98,* 281-300.

Watson, D., Clark, L. A.,& Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063-1070.

Widiger, T.A. (2009). Neuroticism. In M.R. Leary & R.H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 147-160). New York: Guilford.

Williams, J. E., Satterwhite, R. C., & Saiz, J. L. (2010). *The importance of psychological traits. A cross-cultural study*. Boston: Kluwer Academic Publishers.

Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology, 84*, 608-618.

Table 1 *Factor loadings, factor correlations, scale inter-item correlations and internal consistency of the revised Italian version of the Ten Item Personality Inventory (I-TIPI-R) in student (Study 2, n=157, left values) and general population (Study 3, n=472, right values) samples.*

N.	English	Italian	Scale	*M*	*SD*	E	A	C	N	O
1	Extraverted, enthusiastic.	Estroversa, esuberante	E	4.25 4.44	1.55 1.71	**.98 .95**	.14 .07	-.10 -.02	-.03 .04	-.06 -.03
2	Critical, quarrelsome.	Polemica, litigiosa	A(r)	4.41 4.50	1.59 1.83	-.29 -.17	**.51 .42**	-.13 -.04	-.21 -.23	.03 .01
3	Dependable, self-disciplined.	Affidabile, auto-disciplinata	C	5.71 5.69	1.06 1.16	-.02 -.01	.24 .09	**.75 .89**	.15 .12	.08 .07
4	Anxious, easily upset.	Ansiosa, che si agita facilmente	N	4.57 3.75	1.64 1.96	-.08 -.01	.16 .05	.00 .04	**.75 .86**	-.06 -.10
5	Open to new experiences, complex.	Aperta alle nuove esperienze, con molti interessi	O	5.36 5.63	1.13 1.29	.15 .18	.19 .26	.04 .06	-.05 -.09	**.63 .45**
6	Reserved, quiet.	Riservata, silenziosa	E(r)	3.97 3.90	1.69 1.79	**.54 .54**	-.19 -.26	.07 -.01	-.08 -.10	.13 .05
7	Sympathetic, warm.	Comprensiva, affettuosa	A	5.75 5.52	0.94 1.39	.09 -.01	**.59 .50**	.05 -.07	.14 .04	.01 .01
8	Disorganized, careless.	Disorganizzata, distratta	C(r)	4.64 4.93	1.63 1.73	-.03 -.02	-.16 -.25	**.75 .56**	-.17 -.19	-.09 -.07
9	Calm, emotionally stable.	Tranquilla, emotivamente stabile	N(r)	3.56 3.28	1.53 1.58	.00 -.04	**-.30** -.20	-.12 -.12	**.46 .46**	.05 .18
10	Conventional, uncreative.	Tradizionalista, abitudinaria	O(r)	3.90 3.92	1.60 1.78	-.11 -.08	-.14 -.09	-.06 -.01	.00 .02	**.73 .81**
		correlation with A				-.24 -.12				
		correlation with C				-.04 -.01	.44 .33			
		correlation with N				-.08 -.05	-.30 -.14	-.14 -.21		
		correlation with O				.30 .27	-.16 -.02	-.27 -.17	-.16 -.02	
		α				.72 .70	.38 .44	.67 .57	.50 .56	.58 .50
		r				.56 .54	.27 .29	.55 .43	.33 .39	.43 .35

Note: (r) = reverse item; proportion of variance accounted for by the factor solution = 77.43% and 72.66%; M = mean, SD = standard deviation, E = Extraversion, A = Agreeableness, C = Conscientiousness, N = Neuroticism, O = Openness; α = Cronbach's alpha, *r* = scale inter-item correlation; factor loadings higher than |.30| are bolded for ease of interpretation

Table 2 *Convergent and discriminant correlations between the Italian Ten Item Personality Inventory (Time 1) and the Big Five Inventory in Study 4 (n = 66; italicized values on the diagonal are Cronbach's αs), test-retest correlations (Study 4) and self-observer correlations (Study 5, n = 73).*

	TIPI-E	TIPI-A	TIPI-C	TIPI-N	TIPI-O	BFI-E	BFI-A	BFI-C	BFI-N	BFI-O	PANAS-PA	PANAS-NA	RSES	LOT-R	ERQ-REA	ERQ-SUP	MCSDS-SF
TIPI-E	*.64*																
TIPI-A	-.18	*.31*															
TIPI-C	-.13	.25*	*.60*														
TIPI-N	-.02	-.32**	-.23	*.40*													
TIPI-O	.16	-.07	-.11	-.20	*.35*												
BFI-E	.71***	-.13	-.02	-.03	-.01	*.81*											
BFI-A	-.16	.62***	.17	-.13	.09	-.21	*.76*										
BFI-C	-.16	.14	.79***	-.25*	-.08	-.04	.15	*.86*									
BFI-N	.05	-.35**	-.21	.55***	.02	-.09	-.25*	-.19	*.79*								
BFI-O	.02	.08	.09	-.28*	.58***	-.10	.28*	.01	-.05	*.75*							
PANAS-PA	.14	-.01	.40***	-.30*	.24	.02	.14	.39**	-.20	.23	*.74*						
PANAS-NA	-.02	-.43***	-.26*	.42***	-.16	-.10	-.27*	-.20	.42***	-.27*	-.15	*.84*					
RSES	.11	.25*	.40***	-.42***	.12	.12	.13	.37**	-.39**	.29*	.64***	-.46***	*.87*				
LOT-R	.14	.30*	.21	-.45***	.03	.10	.09	.25*	-.47***	.08	.39**	-.46***	.63***	*.73*			
ERQ-REA	.25*	.05	.03	-.15	.27*	.17	-.14	.04	-.15	.03	.15	-.02	.28*	.33**	*.88*		
ERQ-SUP	-.36**	.03	-.06	-.04	.19	-.38**	.04	.00	-.11	.13	-.04	.06	-.07	-.11	.06	*.78*	
MCSDS-SF	-.10	.36**	.10	-.26*	.00	-.02	.28*	.12	-.33**	.06	.00	-.34**	.15	.16	.14	-.03	*.68*

	Time 1								Observer								
Time 2	TIPI-E	TIPI-A	TIPI-C	TIPI-N	TIPI-O	*M*	*SD*	*α*	Self	TIPI-E	TIPI-A	TIPI-C	TIPI-N	TIPI-O	*M*	*SD*	*α*
TIPI-E	.87***	-.18	-.14	-.07	.26*	4.08	1.50	.66	TIPI-E	.58***	.03	.11	-.11	.22	4.20	1.43	.63
TIPI-A	-.27*	.81***	.30*	-.23	-.08	5.36	1.07	.35	TIPI-A	-.12	.42***	.23	-.18	.07	4.93	1.06	.20
TIPI-C	-.16	.26*	.90***	-.21	-.11	5.50	1.19	.66	TIPI-C	-.01	.07	.61***	.05	-.11	5.15	1.26	.37
TIPI-N	.02	-.32**	-.25*	.79***	-.17	3.73	1.31	.43	TIPI-N	-.12	-.32**	-.02	.66***	.12	3.83	1.46	.57
TIPI-O	.19	-.01	-.10	-.18	.89***	4.69	1.05	.34	TIPI-O	.21	-.02	-.21	-.11	.44***	4.49	1.18	.59
M	3.98	5.26	5.42	3.72	4.61				*M*	4.11	4.97	5.41	3.99	4.24			
SD	1.42	1.04	1.18	1.32	1.00				*SD*	1.28	1.15	1.28	1.61	1.12			
α	.64	.31	.60	.40	.35				α	.56	.20	.57	.68	.47			

Note: * = $p < .05$, ** = $p < .01$; *** = $p < .001$; TIPI = Ten Item Personality Inventory; BFI = Big Five Inventory; PANAS-PA = Positive and Negative Affect Scale - Positive Affect; PANAS-NA = Positive and Negative Affect Scale - Negative Affect; RSES = Rosenberg Self-Esteem Scale; MCSDS-SF: Marlow-Crowne Social Desirability Scale Short Form; ERQ-REA: Emotion Regulation Questionnaire-Reappraisal; ERQ-SUP: Emotion Regulation Questionnaire-Suppression; M = mean; SD = standard deviation.