MARIOLA ŁAGUNA[a]
WACŁAW BĄK[a]
EWELINA PURC[a]
EMILIA MIELNICZUK[a]
PIOTR K. OLEŚ[b]
[a]The John Paul II Catholic University of Lublin
  Institute of Psychology
[b]University of Social Sciences and Humanities in Warsaw

SHORT MEASURE OF PERSONALITY TIPI-P
IN A POLISH SAMPLE

The paper presents the Polish adaptation of the Ten-Item Personality Inventory (TIPI-P), which is designed to measure five dimensions of personality as described in the Big Five model. The items were translated into Polish from the original English version. The equivalence of the two versions was confirmed in a study conducted on a sample of bilingual students. A series of four independent studies (on a total sample of over 500 students) were conducted to verify the psychometric properties of the Polish version of TIPI. They proved acceptable test-retest reliability of TIPI-P after two weeks. As in the original version of the inventory, the internal consistency coefficients are relatively low. Correlations between TIPI-P and NEO-FFI as well as between self-rating and peer-rating confirm the convergent and discriminant validity of TIPI-P scales. The results of the studies suggest that the Polish adaptation can be used in scientific research conducted on student samples.

Keywords: personality traits, the Big Five, psychological measurement, TIPI.

Personality traits are among the most basic variables in personality description (cf. Goldberg, 1992; McAdams & Pals, 2006; McCrae & Costa, 1999). They are relatively stable dispositions to react in a particular way – this tendency refers to both behaviors and emotional reactions. Traits as such do not lend themselves to observation, but their existence can be inferred on the basis of observable behavior and the emotional reactions displayed. This translates directly into

---

Corresponding author: Mariola Łaguna – Institute of Psychology, The John Paul II Catholic University of Lublin, Al. Racławickie 14, 20-950 Lublin, Poland; e-mail: laguna@kul.pl

the manner of investigating traits. In trait theory, the commonly used measures are self-rating questionnaires (in which the respondent retrospectively describes his or her most typical reactions and behaviors) and peer-rating questionnaires (in which the intensity of a given trait is inferred on the basis of rating provided by an observer). In both cases it is assumed that traits manifest themselves directly in behavior that is overt and accessible to observation.

Although theorists usually share these general assumptions on the nature of personality traits and their methodological implications, they offer different proposals regarding the number and organization of basic traits. The most popular models currently include solutions in which it is considered optimal to distinguish five basic personality traits (Cieciuch & Łaguna, 2014). Various five-factor models have been developed, the ones most frequently used in research being the Big Five model (Goldberg, 1992), in which the basic structure of personality traits comprises Extraversion, Agreeableness, Conscientiousness, Emotional Stability (or its opposite, Neuroticism), and the Intellect or Imagination factor, also referred to (in the psychometric tradition, McCrae & Costa, 1999) as Openness to Experience. Descriptions of this model of personality can be found in the literature (e.g., Oleś, 2000; Siuta, 2009). The aim of this article is to present the adaptation of one of the short instruments for measuring the intensity of the five personality traits.

Measures of Personality Traits
as Identified in the Five-Factor Model

A number of instruments for measuring the intensity of personality traits have been designed. The most popular ones include those rooted in the psychometric approach to personality traits: the NEO-PI-R (NEO Personality Inventory, Revised, Costa, & McCrae, 1992) and the NEO-FFI (NEO Five-Factor Inventory, McCrae & Costa, 2004), both having their Polish adaptations (Siuta, 2006; Zawadzki, Strelau, Szczepaniak, & Śliwińska, 1998). The NEO-PI-R consists of 240 items, and its scores, apart from five general scales, provide information also on six specific traits within each of the five basic dimensions. The NEO-FFI is a shorter instrument – it consists of 60 items and allows to assess the intensity of the basic five personality traits without descending to the level of specific traits. Another frequently used measure is the BFI questionnaire (The Big Five Inventory; John, Donahue, & Kentle, 1991), drawing on the lexical tradition. It consists of 44 sentences reflecting the traits that are the most typical for each dimension. An interesting measure is also the SIFFM inventory (Structured Interview

for the Five-Factor Model of Personality) – a structured interview designed by Trull and Widiger (1997), used mainly in diagnosing personality disorders (Klinkosz & Sękowski, 2009).

All the above measures are relatively long and require a considerable amount of participants' time to complete. This is not a big problem in assessment and counseling, and so these measures are often used in those fields of psychological practice (e.g., Jankowski, Oleś, Bąk, & Oleś, 2009). Still, the time-consuming character of the study does limit the possibilities of applying them in complex research projects that include a set of other methods as well.

In scientific research, there is an increasingly visible tendency to use shorter measures (Credé, Harms, Niehorster, & Gaye-Valentine, 2012; Gosling, Rentfrow, & Swann Jr., 2003). Among other reasons, this stems from the necessity to reduce the time of questionnaire completion, which is important especially in research carried out via the Internet. Short scales are also useful where there is a need for repeated measurement of a particular trait at short time intervals – in experimental studies or in longitudinal studies, such as diary study (Bolger & Laurenceau, 2013). Filling in questionnaires consisting of many items may discourage participants and result in careless responding or resignation from participating in the study. It was for these reasons that a few short scales for measuring personality traits have been developed, some of them even consisting of single items only (for a review of measures, see Credé et al., 2012). Short scales tend to have weaker psychometric properties than longer ones, but in the case of some scales these differences are not very large (Credé et al., 2012) and the possibility of using a shorter measure is often a better solution than abandoning measurement altogether.

Several short measures of personality traits follow the lexical approach, using adjectives. People's important and universal characteristics are encoded in language, often in the form of adjectives as the most natural way to describe oneself or another person (Goldberg, 1992; Gorbaniuk, Budzińska, Owczarek, Bożek, & Juros, 2013). This observation became the starting point for the construction of adjective-based scales: a longer one, consisting of 100 adjectives, Trait Descriptive Adjectives (TDA; Goldberg, 1992), and short ones, consisting of five items – Five-Item Personality Inventory (FIPI; Gosling et al., 2003), and of ten items – Ten-Item Personality Inventory (TIPI; Gosling et al., 2003). The last of these is analyzed in the present article.

Short Measure of Personality TIPI-P

The Ten-Item Personality Inventory (TIPI[1]) is a short measure built of ten items, each of them being a pair of adjectives (Gosling et al., 2003). The adjectives were not selected out of the items of already existing measures, but chosen in such a manner that they: (1) reflect as diverse characteristics constituting a given trait as possible; (2) constitute a description of both the negative and the positive pole of a given trait; (3) are not descriptions of extreme intensity of the trait; (4) do not contain negations; (5) minimize the redundancy of trait descriptors. Since its publication in 2003, the measure has been used in numerous studies and has had many cultural adaptations (e.g., Muck, Hell, & Gosling, 2007; Romero, Villar, Gómez-Fraguela, & López-Romero, 2012).

Participants respond to ten pairs of adjectives on a 7-point scale, from 1 – *strongly disagree* to 7 – *strongly agree*. The score on each scale is the mean of two items, one of which is negatively keyed (see Table 2). The time of responding is very short and does not exceed five minutes.

The article presents the process of developing the Polish adaptation of TIPI, which we refer to using the TIPI-P abbreviation (by analogy to other adaptations, e.g., the German one, known as TIPI-G, Muck et al., 2007), and the results of research on its psychometric properties. Four independent studies concerning its reliability and validity will be described, as well as the analysis on aggregate data from several samples.

The Preparation
of the Polish-Language Version

The pairs of adjectives making up the original scales were translated into Polish by three independent translators. Based on these translations, the Polish version was established and then back-translated. This procedure revealed a few places where slight changes were necessary. The Polish version thus prepared as well as the original English version of the measure were administered to 17 final-year students of English (including 15 women) aged from 22 to 24 years ($M =$ 23.18, $SD = 0.64$). Correlations between the corresponding scales of the two language versions ranged from .50 to .81 (Table 1). Correlations for individual items ranged from .04 to .91. Three items with the lowest correlations (statements 4, 7, and 10) were corrected again in the Polish version.

---

[1] We suggest using the English abbreviation and name of the measure because these are commonly used, also in different cultural adaptations (e.g., Muck et al., 2007; Romero et al., 2012).

Table 1

Correlations Between the Polish and English Versions, Reliability, and Descriptive Statistics for TIPI-P Scales

Scale	Polish-English version correlations N = 53	Reliability				Descriptive statistics			
		Test-retest (after 2 weeks)		Cronbach's α		M		SD	
		Polish version N = 105	Original version[1] N = 180	Polish version N = 399	Original version[1]	Polish version N = 399	Original version[1]	Polish version N = 399	Original version
Extraversion	.82	.66	.77	.54	.68	4.82	4.56	1.39	1.48
Agreeableness	.75	.74	.74	.41	.40	4.99	5.26	1.18	1.12
Conscientiousness	.75	.71	.71	.67	.50	4.88	5.47	1.43	1.13
Emotional Stability	.69	.66	.70	.45	.73	4.28	4.85	1.44	1.45
Openness to Experience	.65	.60	.62	.42	.45	5.25	5.43	1.21	1.06

Note. [1]Values for the original TIPI, based on Gosling et al. (2003).

The translation of test items that have an adjectival form involves a special kind of language difficulties because an adjective without any context usually has more than one equivalent in a different language. When preparing the Polish version of TIPI, we encountered problems similar to those that occurred, for instance, during the translation of the English version into German (Muck et al., 2007). The eventually adopted translations of a few adjectives are somewhat different than the versions that would seem to follow from the direct translation of their contents. For example, one of the two words in Item 1 – *extraverted* – was translated as *towarzyski* (which, literally, means *sociable*). Even though this slightly narrows down the meaning of the original English term, we chose this version because the literal translation, *ekstrawertyczny*, is rarely found in natural Polish and may therefore turn out to be incomprehensible to an average respondent. Similarly, the adjective *anxious* was translated as *niespokojny*, (literal meaning: *uneasy*) because the adjective *lękliwy* showed very low and statistically not significant correlation with the original version and did not correlate with the other item in the same scale. *Wszechstronny* (literally: *versatile*) as the translation of the adjective *complex* (*złożony*) is not a literal rendition, either, but it accurately captures the meaning of the word by referring to the trait of Openness to Experience, which it is supposed to measure. The adjective *sympathetic* was originally translated as *współczujący*, but this translation resulted in negative correlation between the Polish and English versions, so the word was eventually rendered as *sympatyczny*, making up one of the two indicators of Agreeableness together with the adjective *ciepły* (*warm*). This solution may be controversial,

since the adjective *sympatyczny* (literally: *nice*, *amiable*) is not always regarded as a personality descriptor (Gorbaniuk et al., 2013). The adjective *konwencjonalny* as the literal translation of *conventional* resulted in low correlation between the Polish and English versions, and so the translation adopted was *zwyczajny* (literally: *ordinary*). In all, certain difficulties in rendering contents into Polish concerned five out of 20 adjectives making up the 10 items of the measure. The final content of TIPI-P items is presented in Table 2.

Table 2
*Correlations Between TIPI-P Items (*N = 399)

Item content	1	2	3	4	5	6	7	8	9	10
1. Towarzyski, entuzjastyczny (E)	1									
2. Krytyczny, kłótliwy (A-r)	-.05	1								
3. Niezawodny, zdyscyplinowany (C)	.15**	-.05	1							
4. Niespokojny, skłony do zamartwiania się (ES-r)	-.28***	.20***	-.13*	1						
5. Otwarty na nowe doświadczenia, wszechstronny (O)	.47***	-.09	.10	-.27***	1					
6. Powściągliwy, cichy (E-r)	**-.38***	-.11*	-.03	.21***	-.21***	1				
7. Sympatyczny, ciepły (A)	.36***	**-.27***	.09	-.03	.32***	.00	1			
8. Niezorganizowany, niedbały (C-r)	-.06	.13*	**-.51***	.16**	-.03	.04	-.07	1		
9. Spokojny, stabilny emocjonalnie (ES)	.08	-.28***	.26***	**-.29***	.12*	.22***	.19***	-.19***	1	
10. Zwyczajny, mało twórczy (O-r)	-.17**	-.06	-.04	.14**	**-.27***	.11*	-.13*	.10*	.02	1

Note. r – recoded item; E – Extraversion, A – Agreeableness, C – Conscientiousness, ES – Emotional Stability, O – Openness to Experience; bold type – correlations between items of the same scale; *** $p < .001$; ** $p < .01$; * $p < .05$ (two-tailed).

EQUIVALENCE OF THE POLISH AND ENGLISH VERSIONS: STUDY 1

In order to test the level of equivalence between the Polish version of TIPI (TIPI-P), prepared in the way described above, and the original English version, the convergence of responses given by individuals knowing both languages well

was analyzed (Drwal, 1995). Despite certain limitations stemming from different degrees of familiarity of the two cultural contexts, the samples used in this procedure are usually either students of foreign languages or foreign students studying in Poland. In Study 1, the former approach was chosen.

Method

The participants in this study were 53 fourth- and fifth-year students of English (including 38 women, 71.7%) aged from 20 to 25 years ($M = 22.43$, $SD = 1.04$). The study was carried out in groups during academic classes. At the beginning of the class, the participants were supposed to fill in the Polish version of the TIPI-P questionnaire, and after about one and a half hours they were asked to fill in the English-language version. Participation in the study was voluntary and participants received no remuneration.

Results and Discussion

Equivalence between the Polish version and the English one was assessed by analyzing correlations between the results of the two versions. The correlation coefficients (Table 1) between the two versions of all the five scales are fairly high and all of them are statistically significant at $p < .01$, their median being .75. The highest correlation was obtained for Extraversion ($r = .82$) and the lowest for Openness to Experience ($r = .65$). Mean correlation for all the scales, computed after Fisher z-transformation, is .95 (this transformation allows to add the values of coefficients, Fisher, 1921). The analysis of differences between the means obtained for each scale of the Polish and English versions revealed that they did not differ significantly, with the exception of the Emotional Stability scale, for which Polish version scores were higher than English version scores, $t(51) = 2.80$, $p = .007$.

The results of analyses point to a convergence of scores obtained using the Polish version of the measure with those obtained using the English version. The correlations are fairly high, even though correlations between adjectives in two different languages tend to be rather low – when using a different language, a person may, in some sense, activate a different aspect of his or her personality, thereby somewhat changing the characteristics attributed to themselves (Ramírez-Esparza, Gosling, Benet-Martínez, Potter, & Pennebaker, 2006). For this reason, in some adaptations of TIPI, the analysis of comparisons between different language versions filled in by the same people was abandoned (e.g., Muck et al., 2007). Our study nevertheless showed that there is a convergence between

the scores of bilingual individuals, even if the scale consists of adjectives. This argues in favor of the Polish adaptation presented here (TIPI-P).

THE TEMPORAL STABILITY OF THE MEASURE:
STUDY 2

One of the indicators of the reliability of scales is test-retest reliability, namely the repeatability of the scores from the same group obtained using the same measure at a certain time interval (Hornowska, 2005). This method is not free from the influence of memory factors, differences in study conditions, and other situational variables. However, it is particularly recommended in the case of a measure which, due to the small number of items and to their special character (aimed at capturing various dimensions of a given trait) cannot achieve high values of Cronbach's alpha internal consistency (Gosling et al., 2003). The aim of this study was to assess the stability of scores obtained on the Polish adaptation of TIPI.

Method

The temporal stability of the Polish adaptation of the measure was assessed on the basis of a two-time study with a two-week interval between measurements. The participants were 91 individuals (including 61 women, 67%) aged from 18 to 29 years ($M = 21.46$, $SD = 2.43$). In this group, 52.4% were first-year psychology students and the remaining 47.6% were fifth-year students of environmental engineering.

Results and Discussion

The values of correlation coefficients between two measurements ranged from .66 to .74, all correlations being statistically significant at $p < .001$ (Table 1). Due to the high homogeneity of the sample (students of two majors), correlation values may be underestimated.

The correlation coefficients obtained indicate acceptable temporal stability of the scales, given that the scales are very short. Scores were found to be the most stable in the case of the Conscientiousness scale and the least stable on the Extraversion and Emotional Stability scales. The correlation values obtained are similar to those found for the original version of the measure (Table 1), where they ranged from .62 to .77 after an interval of about six weeks (Gosling et al.,

2003). Similar values were also obtained for other TIPI adaptations (e.g., from .52 to .83 for the Spanish version; Romero et al., 2012).

INTERNAL CONSISTENCY
AND ITEM INTERCORRELATIONS

Numerous measures are designed in such a way as to achieve high internal consistency of results. It should be remembered that Cronbach's alpha coefficient depends on the intercorrelation of items as well as on scale length – the longer and the more homogeneous the scale, the higher the internal consistency (Hornowska, 2005). In the case of short scales, high internal consistency cannot therefore be expected, and even the very calculation of Cronbach's alpha is debatable, though its values are traditionally given in study reports. Each of the TIPI scales consists of only two items, designed to represent as diverse indicators of a given trait as possible and to avoid redundancy (Gosling et al., 2003); consequently, correlations between them cannot be high.

Method

The data used in analyses had been collected by means of TIPI-P from the four study samples described in the article: Study 1 sample – scores obtained using the Polish version (TIPI-P); Study 2 sample – the results of the first measurement; Study 3 and Study 4 samples, described further below – the results of self-rating of traits. In total, the entire sample consisted of 399 students (including 257 women, 64.4%), aged from 18 to 34 years ($M = 22.03$, $SD = 2.53$).

Results and Discussion

Cronbach's α values (Table 1) range from .41 to .67. They are the lowest for the Agreeableness and Openness to Experience scales and the highest for the Conscientiousness scale. Basic descriptive statistics were also computed and found to be close to the values obtained using the original measure (Table 1; cf. Gosling et al., 2003).

The values of correlations between the measure's items (Table 2) indicate that the pairs of items included in the same scale are, as expected (one of the items describes the negative pole of a given trait and the other describes the positive one), significantly and negatively correlated, with correlation values ranging from -.27 to -.51. Correlation analysis points to problems with three items, whose

correlations with a different item are higher than correlations with an item included in the same scale, though both are positive.

Correlations between TIPI items confirm a statistically significant and moderately negative relationship between adjectives making up a particular scale. Given that the scales consist of only two items each and that the items are purposefully made diverse (Gosling et al., 2003), the results of correlation analysis are not surprising. They are also similar to those obtained for other adaptations of the measure.

Computing the internal consistency coefficient for scales consisting of only two items is debatable, but since this had been done at the stage of developing the measure as well as its other adaptations, we have presented the results of this analysis too. As in studies using the original version (Table 1; cf. Gosling et al., 2003), in our study relatively low internal consistency coefficients were obtained, too. They are similar in the two studies, only the Emotional Stability scale in the Polish version is markedly less consistent than in the original version. For comparison, in the German adaptation, TIPI-G, the values of Cronbach's α ranged from .42 for the Agreeableness scale to .67 for the Extraversion scale (Muck et al., 2007), and in the Spanish version they ranged from .44 for Agreeableness to .61 for Extraversion (Romero et al., 2012). Likewise, in studies on a sample of American students, the values of Cronbach's alpha were from .45 to .67 (Credé et al., 2012). This means the values obtained for TIPI-P are similar.

CONVERGENT AND DISCRIMINANT VALIDITY:
STUDY 3

One of the methods of estimating the theoretical validity of a measure is the assessment of convergent and discriminant validity. In the study presented here, correlations between TIPI and NOE-FFI scores were analyzed.

Method

The participants were 128 students of various majors (including 66 women, 51.6%), aged from 18 to 34 years ($M = 22.56$, $SD = 2.84$). Apart from the TIPI-P scale, we used the NEO-FFI questionnaire as adapted into Polish by Zawadzki and colleagues (1998). It consists of 60 items, to which participants respond on a 5-point scale from *strongly disagree* to *strongly agree*. Cronbach's α reliability of measurement using the Polish version of NEO-FFI scales ranges from .68 to .82 (Zawadzki et al., 1998).

Results and Discussion

The correlation coefficients are statistically significant, average or high, and point to the convergence of measurements using the two measures (Table 3), the median being .65. Only for the Openness to Experience scale is the correlation relatively low. The mean value of correlations (after Fisher transformation) pointing to the convergent aspect of validity is .71, whereas the mean correlation pointing to the discriminant aspect of validity is much lower and equals .13. Only some of the discriminant correlations are statistically significant – in the case of TIPI-P, significant but low correlations with the remaining scales are found for the Extraversion scale, and they are similar to those between NEO-FFI scales (Zawadzki et al., 1998).

Table 3
Correlations Between TIPI-P and NEO-FFI Scales (N = 128)

Measure	NEO-FFI				
	Extraversion	Agreeableness	Conscien-tiousness	Neuroticism/ Emotional Stability	Openness to Experience
TIPI					
Extraversion	**.70*****	.21*	.04	-.37***	.07
Agreeableness	.33***	**.62*****	.13	-.15	.04
Conscientiousness	.07	.11	**.65*****	-.22*	-.11
Neuroticism/Emotional Stability	.38***	.09	.15	**-.65*****	-.17
Openness to Experience	.37***	-.05	-.05	-.24**	**.38*****

Note. Convergent correlations – on the diagonal (bold type); discriminant correlations – below and above the diagonal; *** $p < .001$; ** $p < .01$; * $p < .05$ (two-tailed).

Analysis of correlations of the Polish version (TIPI-P) with NEO-FFI scores points to the convergence of the measurement of the same traits using the two measures, with the exception of the Openness to Experience scale, and at the same time to the independence of the measurement of different traits. The NEO-FFI questionnaire, applied in this case, is a shortened version of the NEO-PI-R; the results of our study are quite similar to the correlations between TIPI-G and the NEO-PI-R in a German sample, which ranged from .41 for Openness to Experience to -.76 for Neuroticism/Emotional Stability (Muck et al., 2007). The results presented here are also similar to the convergent correlations between the

original measure and the BFI, which ranged from .48 for Openness to Experience to .80 for Extraversion (Gosling et al., 2003). It is worth noting that in all previous studies the Openness to Experience scale showed the lowest measurement convergence. The relatively low correlations we obtained for this scale are not, therefore, a defect of the Polish adaptation.

CONVERGENT VALIDITY ASSESSED THROUGH PEER-RATING: STUDY 4

In research on personality it is assumed that traits are accessible to external observation and find expression in language when we describe another person (cf. Gorbaniuk et al., 2013). Consequently, self-rating should at least to some extent be consistent with peer-rating. The assessment of the convergence of these ratings may therefore be an indicator of the validity of the instrument for measuring personality. In the case of TIPI, this kind of approach was used in the German adaptation of the measure, revealing correlations that ranged from .32 for the Agreeableness scale to .56 for the Extraversion scale (Muck et al., 2007). We applied this approach to assessing validity also in the Polish adaptation presented here.

Method

After being introduced to the procedure and after consenting to participate in the study, each participant received two questionnaires. Each participant filled in one of the questionnaires themselves, rating his or her friend on the TIPI-P scale (peer-rating). The other questionnaire was to be passed on to the person thus described; the participant was to ask that person to fill it in as a self-rating survey. Participants were asked not to reveal their mutual ratings to one another until they have completed the questionnaires. In this manner, a total of 254 people were examined (Sample 4). Half of them rated themselves while the other half rated their friends.

Self-ratings were collected from 127 individuals (including 92 women, 72.4%), aged from 17 to 62 years ($M = 22.22$, $SD = 6.10$). A majority of the participants (88.2%) were students. The group of students who described their friend consisted of 127 individuals (including 99 women, 78%), aged from 18 to 29 years ($M = 21.17$, $SD = 1.93$). Of the 116 people who answered the question of where they knew the rated person from, 63 declared knowing the person on a personal basis, 9 reported they knew the person from the university only, and

44 declared knowing the person both on a personal basis and from the university. The mean length of acquaintance estimated by the peer-raters was 61.45 months (SD = 52.50). The mean quality of the relationship was rated as very good (M = 5.02, SD = 1.01) using a scale from 1 – *neutral* to 6 – *exceptionally good*.

Results and Discussion

All the correlation coefficients between self-rating and peer-rating are statistically significant, though their values are average (Table 4). The convergence of scores is the highest for the Extraversion scale (r = .48), and the lowest for Conscientiousness (r = .31). The mean value of convergent correlations (after Fisher transformation) is .42 and is higher than the mean value of discriminant correlations, which equals .19.

Table 4
Correlations Between Self-Rating and Peer-Rating Scores (N = 127)

		Self-rating					Peer-rating				
		E	A	C	ES	O	E	A	C	ES	O
Self--rating											
	E	1.00									
	A	-.04	1.00								
	C	.10	.20*	1.00							
	ES	.03	.45***	.20*	1.00						
	O	.24***	.28***	18*	24**	1.00					
Peer--rating											
	E	.48***					1.00	-.03	.07	.10	.44***
	A		.41***					1.00	.15	.30***	.27**
	C			.31***					1.00	.11	.16
	ES				.44***					1.00	.26**
	O					.33***					1.00

Note. E – Extraversion; A – Agreeableness, C – Conscientiousness; ES – Emotional Stability; O – Openness to Experience; *** p < .001; ** p < .01; * p < .05 (two-tailed).

The results of the study point to a moderate but statistically significant convergence of self-rating and peer-rating, thus confirming the validity of the Polish

adaptation, TIPI-P. The values of convergent correlations obtained in this study are close to the values obtained – using a similar methodology – in German research (Muck et al., 2007). In both cases, self-rating and peer-rating scores were the most convergent on Extraversion. In the German sample, lower convergence was found for the Agreeableness (r = .32) and Emotional Stability scales (r = .38), whereas in our study it is the Conscientiousness and Openness to Experience scales that show lower convergence between self-rating and peer-rating. The mean values of convergent correlations are very similar in the two samples, however (in the German sample, the mean was .43).

It is also worth noting that individuals rating their friends described their relations with them as good, which may have lowered the variance of ratings (cf. Gorbaniuk, Szczepańska, Ivanova, & Zygnerska, 2014) and may have led to underestimating the value of correlations.

GENERAL DISCUSSION

Recent years have seen a considerable growth of interest in short measures of various psychological constructs (cf. Credé et al., 2012). However, this does not refer to the development of instruments for individual diagnosis but to measures used for research purposes, as research ought to be precise and economical at the same time. The development of multivariate analyses as well as online and intensive longitudinal studies (Bolger & Laurenceau, 2013) makes it necessary to use short measures, but they have to be measures in which the shortening has not been achieved at the cost of considerable deterioration of psychometric properties.

The paper presents the Polish adaptation of TIPI, a short 10-item measure for rating the five personality traits. In the course of work on the adaptation, several independent studies were carried out. Acceptable psychometric properties were found in the study of convergence between the two language versions – the Polish translation and the original English version. The results concerning reliability, assessed as temporal stability, are acceptable as well.

The Polish adaptation, TIPI-P, just like the original version and other adaptations, exhibits relatively low intercorrelations of items making up scales and has relatively low values of Cronbach's α coefficient for each scale. It is, however, worth stressing that this measure was not designed in order to maximize the convergence of responses – quite the opposite: the aim was to ensure a broad representation of traits while avoiding redundancy (Gosling et al., 2003). At the same

time, in the case of short scales, temporal stability rather than internal consisten-cy values are pointed to as worth investigating. The former tend to be higher (cf. Romero et al., 2012), and so they were in our study. However, this does not change the fact that short scales are less reliable than longer measures, and for this reason TIPI-P scales cannot be used in individual assessment. The measure is useful in scientific research in which using a longer instrument would not be possible and would result in the participants' discouragement or refusal to cooperate.

Due to their relatively low reliability, short scales for measuring traits may lead to the underestimation of the role of personality traits (Credé et al., 2012). Still, the TIPI-P measure presented here has better psychometric properties than scales consisting of only one item. Moreover, correlations between TIPI-P and external criteria are similar to the values obtained in studies using longer meas-ures rather than to those obtained using single test items (Credé et al., 2012).

The studies we carried out largely confirm the validity of the measurement of the basic five personality traits using TIPI-P. That validity was estimated both as correlations with the NEO-FFI and as convergence between self-rating and peer-rating. Other studies using TIPI-P show the internal validity of this measure. Analysis of correlations in a sample of 120 adults aged from 18 to 65 years (Bojarska, 2013) revealed positive correlations of TIPI-P scales with self-esteem (measured using Rosenberg's SES, Łaguna, Lachowicz-Tabaczek, & Dzwon-kowska, 2007) and their negative correlations with depression – a dimension of the sense of helplessness (measured using Beck's HS, Oleś & Juros, 1985-86). These results were similar to those obtained using the original version of TIPI (Gosling et al., 2003). In this study, positive correlations were found between TIPI-P scales and satisfaction with life (measured using Diener's SWLS; Juczyński, 2001), showing a pattern similar to that obtained in research using the Spanish version of the measure (Romero et al., 2012).

When analyzing the limitations of our studies, one should note the fact that the participants were mostly students. This limits the scope of generalization of the obtained results to other samples. On the other hand, a substantial part of scientific research in which TIPI-P may prove to be useful is carried out on stu-dents. Therefore, testing the psychometric properties of the measure on this kind of sample seems reasonable. The psychometric properties of the measure in other samples remains an open issue.

The results of our research show that the Polish version of TIPI, as a short measure for investigating five personality traits, gives a fairly good approxima-tion of results obtained using longer scales such as the NEO-FFI questionnaire.

Although TIPI-P scales, consisting of two items each, do not have such good psychometric properties as long tests, for example the NEO-PI-R, the instrument does allow us to measure the general dimensions of personality in an economical and respondent-friendly way.

REFERENCES

Bojarska, M. (2013). *Orientacja pozytywna a orientacja negatywna: Dwa niezależne wymiary czy dwa bieguny tego samego czynnika?* (Unpublished master's thesis). SWPS, Warsaw.

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research.* New York: Guilford.

Cieciuch, J., & Łaguna, M. (2014). The Big Five and beyond: Personality traits and their measurement. *Roczniki Psychologiczne, 17*(2), 249-257.

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.

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.

Drwal, R. (1995). Problemy kulturowej adaptacji kwestionariuszy osobowości. In R. Drwal (Ed.), *Adaptacja kwestionariuszy osobowości* (pp. 12-26). Warsaw: Wydawnictwo Naukowe PWN.

Fisher, R. A. (1921). On the 'probable error' of a coefficient of correlation deduced from a small sample. *Metron, 1*, 3-32.

Goldberg, L. R. (1992). The development of markers of the Big-Five factor structure. *Psychological Assessment, 4*(1), 26-42.

Gorbaniuk, O., Budzińska, A., Owczarek, M., Bożek, E., & Juros, K. (2013). The factor structure of Polish personality-descriptive adjectives: An alternative psycho-lexical study. *European Journal of Personality, 27*(3), 304-318.

Gorbaniuk, O., Szczepańska, N., Suchomska, M., Ivanova, A., & Zygnerska, M. (2014). Adjective markers of Polish indigenous lexical personality factors: A peer-rating study. *Roczniki Psychologiczne, 17*(2), 309-325.

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

Hornowska, E. (2005). *Testy psychologiczne: Teoria i praktyka*. Warsaw: Scholar.

Jankowski, T., Oleś, M., Bąk, W., & Oleś, P. (2009). Przydatność kwestionariusza NEO-PI-R w diagnozie indywidualnej. In J. Siuta (Ed.), *Diagnoza osobowości. Inwentarz NEO-PI-R w teorii i praktyce* (pp. 113-154). Warsaw: Pracowania Testów Psychologicznych PTP.

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

Juczyński, Z. (2001). *Narzędzia pomiaru w promocji i psychologii zdrowia.* Warsaw: Pracownia Testów Psychologicznych PTP.

Klinkosz, W., & Sękowski, A. (2009). Inwentarz NEO4 i ustrukturyzowany wywiad dla pięcio-czynnikowej teorii osobowości w poradnictwie zawodowym i praktyce klinicznej. In J. Siuta

(Ed.), *Diagnoza osobowości. Inwentarz NEO-PI-R w teorii i praktyce* (pp. 155-174). Warsaw: Pracownia Testów Psychologicznych PTP.

Łaguna, M., Lachowicz-Tabaczek, K., & Dzwonkowska, I. (2007). Skala samooceny SES Morrisa Rosenberga – polska adaptacja metody. *Psychologia Społeczna, 2*, 164-176.

McAdams, D. P., & Pals, J. L. (2006). A new Big Five. Fundamental principles for an integrative science of personality. *American Psychologist, 61*, 204-217.

McCrae, R. R., & Costa Jr., P. T. (1999). A five-factor theory of personality. In L. Pervin & O. John (Eds.), *Handbook of personality: Theory and research* (pp. 139-153). New York: Guilford Press.

McCrae, R. R., & Costa Jr., P. T. (2004). A contemplated revision of the NEO Five-Factor Inventory. *Personality and Individual Differences, 36*, 587-596.

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.

Oleś, P. K. (2000). Kontrowersje wokół „Wielkiej Piątki". *Czasopismo Psychologiczne, 6*, 7-18.

Oleś, P., & Juros, A. (1985-86). Symptom poczucia beznadziejności w kognitywno-afektywnej teorii depresji A. T. Becka: Skala "Hopelessness" – polska adaptacja. *Summarium, 34-35*, 289-298.

Ramírez-Esparza, N., Gosling, S. D., Benet-Martínez, V., Potter, J. P., & Pennebaker, J. W. (2006). Do bilinguals have two personalities? A special case of cultural frame-switching. *Journal of Research in Personality, 40*, 99-120.

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

Siuta, J. (2006). *Inwentarz Osobowości NEO-PI-R. Adaptacja polska. Podręcznik.* Warsaw: Pracownia Testów Psychologicznych PTP.

Siuta, J. (Ed.) (2009). *Diagnoza osobowości. Inwentarz NEO-PI-R w teorii i praktyce*. Warsaw: Pracownia Testów Psychologicznych PTP.

Trull, T. J., & Widiger, T. A. (1997). *Structured Interview for the Five-Factor Model of Personality (SIFFM). Professional manual.* Odessa, FL: Psychological Assessment Resources.

Zawadzki, B., Strelau, J., Szczepaniak, P., & Śliwińska, M. (1998). *Inwentarz osobowości NEO-FFI Costy i McCrae. Polska adaptacja.* Warsaw: Pracowania Testów Psychologicznych PTP.