Psychometric properties of the Slovak translation of the NEO-PI-R questionnaire

Elena Lisà* e Milan Kohút°

* Institute of Applied Psychology,
Faculty of Social and Economic Sciences,
Comenius University in Bratislava;
e-mail: elalisa78@gmail.com.
° Milan Kohut s.r.o.
e-mail: milan.kohut.1@gmail.com.

Ricevuto: 07.12.2021 - **Accettato:** 27.07.2022

Pubblicato online: 18.10.2022

Abstract

The study aimed to verify the psychometric properties of the NEO-PI-R Slovak translation. The self-report sample consisted of 1.062, a retested sample of 122, and an observer rater of 371 persons. In self-assessments, Cronbach's alfa of domains ranged from .88 (Openness) to .92 (Neuroticism). In observer ratings, the reliability ranged from .89 (Extraversion) to .95 (Conscientiousness). The stability of the test-retest showed an average correlation of r = .81 after three months. Exploratory factor analysis revealed five factors that explained the variance of 60.45%. The agreement between the observers and the self-assessments was at a mean level of .58. The Slovak translation of NEO-PI-R follows the psychometric standards for reliability and construct validity. The study did not follow the educational level and mental health of the research participants.

Keywords: NEO-PI-R, reliability, validity, exploratory factor analysis, observer ratings

ISSN 0391-6081, ISSNe 1972-5620, Doi:10.3280/rip2022oa14728

Introduction

The history of the NEO questionnaire began back in the mid-1970s. The 1985 NEO personality inventory was the first version, with the abbreviation NEO-PI (Hřebíčková et al., 2002). The Neo Five-Factor Inventory (NEO-FFI), as the shorter version of the NEO-PI was prepared in 1989. Costa and McCrae (1992) revised the NEO-PI as the NEO Personality Inventory-Revised (NEO-PI-R) in 1992. NEO-PI-R contains 240 items, 48 items for each domain, further divided into six facets. Table 1 shows the specification of the domains'facets (McCrae & Costa, 2003).

Table 1 - *The specification of NEO-PI-R domains'facets*

Neuroticism		
N1	Anxiety	nervous, high-strung, tense, prone to worry
N2	Anger	prone to experience anger, irritable and ill-tempered, hard to get along with
N3	Depression	sadness, hopelessness, and loneliness, feeling guilt and of diminished self-worth
N4	Self-Consciousness	prone to the emotion of shame or embarrassment
N5	Impulsiveness	tendency to give in to temptations and to be overwhelmed by desires
N6	Vulnerability	inability to deal adequately with stress
Extraversion		
E1	Warmth	friendly, cordial, intimately in personal interaction
E2	Gregariousness	sociable, liking the crowds
E3	Assertiveness	natural leadership, easily expressing their feelings and desires
E4	Activity	keeping busy, acting vigorously, energetic and forceful
E5	Excitement Seeking	preferring environments that stimulate them
E6	Positive Emotions	experiencing joy, delight
Openness to Experience		
O1	Fantasy	vivid imagination and daydreams
O2	Aesthetics	sensitivity to art and beauty

O3	Feelings	valuing the experience as a source of meaning in life
O4	Actions	the opposite of rigidity, willing to try a new experience
O5	Ideas	curiosity
O6	Values	liberal and tolerant in values
Agreeableness		
A1	Trust	trusting, believing the best of others
A2	Straightforwardness	trustworthiness
A3	Altruism	considerateness and desire to help others
A4	Compliance	deferring to others rather than aggressively pushing
A5	Modesty	humble
A6	Tender-mindedness	sentimentality
Conscientiousness		
C1	Competence	rational, informed
C2	Order	efficient in work
C3	Dutifulness	inhibited, adhering scrupulously to the moral precepts
C4	Achievement Striving	pursuing excellence in everything they do
C5	Self-Discipline	be able to accomplish the goals
C6	Deliberation	making plans in advance and thinking carefully before acting

According to Evers et al. (2012), NEO-PI-R is currently the most widely used personality inventory in Europe. The current study aims to examine its psychometric properties in the Slovak translation.

Reliability

According to Hřebíčková et al. (2002), the internal consistency coefficients of self-report in the Czech NEO-PI-R range from .91 (N, C) to .88 (A). A study from Sardinia presents internal consistency values from .80 (A) to .87 (N) (Costa et al., 2006). These values are acceptable and are similar to the original values from the USA: .92 (N), .89 (E), .87 (O), .86 (A), and .90 (C) (Costa et al., 2006). The high alpha values for five domains are also reported in mainland Italians: .91 (N), .88 (E), .87

(O), .86 (A), and .91(C) (Terracciano, 2003). For the Russian NEO-PI-R (N = 350, 16-63), Martin et al. (2002) report on the mean internal consistency of .88, with the highest C (.91) and the lowest O (.85). Källmen et al. (2010) report internal consistency from .75 (A) to .85 (N), with a mean of .79 for the Swedish standardization study (N = 766). For Romanian NEO-PI-R, Ispas et al. (2014) reported the mean consistency of domains at .87, with the highest N (.91) and the lowest O (.83). A French study by Rolland et al. (1998) produces results from two samples, university students (N = 447) and soldiers (N = 268). In both samples, the consistency is the lowest in O (.83 and .77), the highest in N and C (.90 and .91). The mean consistency is .87 for students and .84 for soldiers.

N and C are the most consistent domains (r ~ .90), regardless of the language and cultural environment. The least internal consistency is in the O domain (r ~ .84). Inter-item variability within domains could explain the differences between Cronbach alphas (Urbina, 2004). The inter-item variability is low for the N and C domains and high for the O domain. The reason is that C and N are more biologically dependent, whereas O is more culturally and environmentally dependent (Strelau, 2001). As McCrae and Sutin (2009) state, O is a very broad construct and difficult to grasp. It is the weakest domain in replication studies. However, the internal consistency of O in the current review is generally high. Table 2 presents an overview of all the internal consistency coefficients for different countries.

At the facet level, the highest internal consistency shows N3-Depression. In the studies mentioned above, the internal consistency for N3 is reported from .66 (RUS) to .81 (USA). The lowest Cronbach alfa values are found in the O6-Values facet, from .48 (FRA students) to .29 (CZE). P6-Tender-Mindedness was the second least consistent, with values from .63 (SVE) to .61 (ITA). In general, the highest consistency within the facets shows N1-Anxiety (r = .82) in Sweden. The Czech normative analysis has the lowest O6 (r = .29).

Table 2 - *Cronbach coefficient of internal consistency – country-by-country overview*

NEO-PI-R domains/facets		Cronbach α							
	Domain/ Facet	USA	CZE	ITA	RUS	SVE	ROM	FRAs	FRAm
N	Neuroticism	.92	.91	.87	.89	.85	.91	.90	.91
E	Extraversion	.89	.90	.82	.90	.76	.85	.86	.82
O	Openness	.87	.89	.84	.85	.78	.83	.83	.77
A	Agreeableness	.86	.88	.80	.86	.75	.85	.87	.81
C	Conscientiousness	.90	.91	.85	.91	.80	.90	.90	.91
The average value of factors		.89	.90	.84	.88	.79	.87	.87	.84
N1	Anxiety	.78	.79	.68	.76	.82	.77	.80	.74
N2	Angry Hostility	.75	.75	.61	.73	.71	.72	.70	.64
N3	Depression	.81	.80	.73	.66	.80	.78	.77	.74
N4	Self-Consciousness	.68	.68	.57	.61	.66	.68	.65	.59
N5	Impulsiveness	.70	.71	.53	.65	.63	.68	.62	.64
N6	Vulnerability	.77	.66	.63	.64	.75	.74	.78	.72
E1	Warmth	.73	.74	.60	.68	.73	.71	.72	.70
E2	Gregariousness	.72	.80	.63	.80	.77	.68	.73	.67
E3	Assertiveness	.77	.80	.52	.77	.81	.74	.75	.71
E4	Activity	.63	.59	.46	.76	.59	.63	.70	.48
E5	Excitement-Seeking	.65	.73	.63	.64	.66	.70	.55	.44
E6	Positive Emotions	.73	.78	.69	.74	.80	.73	.76	.64
O1	Fantasy	.76	.81	.70	.74	.72	.74	.77	.70
O2	Aesthetics	.76	.79	.65	.73	.80	.75	.73	.71
O3	Feelings	.66	.77	.52	.68	.74	.63	.62	.51
O4	Actions	.58	.68	.43	.61	.66	.55	.50	.38

O5	Ideas	.80	.76	.72	.66	.72	.74	.75	.73
O6	Values	.67	.29	.42	.47	.49	.63	.48	.31
A1	Trust	.79	.75	.64	.70	.74	.75	.81	.75
A2	Straightforwardness	.71	.77	.62	.71	.68	.70	.78	.67
A3	Altruism	.75	.72	.54	.65	.64	.71	.66	.62
A4	Compliance	.59	.69	.59	.62	.74	.57	.68	.46
A5	Modesty	.67	.75	.56	.72	.62	.72	.76	.69
A6	Tender-Mindedness	.56	.52	.41	.43	.63	.57	.55	.61
C1	Competence	.67	.67	.46	.55	.58	.66	.61	.59
C2	Order	.66	.70	.52	.68	.65	.64	.77	.72
C3	Dutifulness	.62	.77	.58	.70	.63	.69	.64	.64
C4	Achievement Striving	.67	.72	.53	.73	.79	.65	.75	.65
C5	Self-Discipline	.75	.79	.62	.76	.78	.72	.75	.77
C6	Deliberation	.71	.78	.68	.73	.77	.71	.69	.69
Facet average values		.70	.72	.58	.68	.70	.69	.69	.63

Notes: USA = United States of America (Costa & McCrae, 1992); CZE = Czech Republic, N=1,365 (Hřebíčková et al., 2002); ITA = Italy (Terracciano, 2003); RUS = Russia, N=350 (Martin et al., 2002); SVE = Sweden, N=766 (Källmen et al., 2010); ROM = Romania, N=2,200 (Ispas et al., 2014); FRAs = France – students, N=447 (Rolland et al., 1998); FRAm = France – soldiers, N=268 (Rolland et al., 1998).

Validity

Hřebíčková (2004) reports an agreement between self-assessment and observer's rating as the construct validity indicator. The mean consensus between self-report and spouse rating and an agreement among three observers is .56. E shows the highest and N the lowest agreement between self-report and observer rating (Hřebíčková et al., 2002). Pajtinková (2011) worked with the Slovak research sample. Pearson's correlation coefficient of agreement between self-report and siblings'assessment is r = .605. The agreement between self-assessment and friend's assessment is r = .566, and the parent's r = .336 (Pajtinková, 2011).

Exploratory factor analysis (EFA) is considered the standard tool for analyzing the validity of the NEO-PI-R construct (Caprara et al., 2001; Hřebíčková et al., 2002; Ispas et al., 2014; Martin et al., 2002; Rolland et al., 1998). Factor loadings exceed the threshold of .30 in most cases. The factor structure of the five-factor personality model appears to be stable (Aluja et al., 2005; Hesselmark et al., 2015; Marshall et al., 2005). The construct validity of the Czech NEO-PI-R with Varimax rotation shows five factors. The factor loadings show values of .30 or higher. The factors explain 61.51% of the variance, with N 14.65%; C 12.83%; A 11.89%; E 11.49%; and O 10.65% (Hřebíčková et al., 2002). However, some studies do not confirm the five-factor structure. For example, 28.75% of the NEO-PI-R items have no cultural relevance in South Africa (Branco e Silva & Laher, 2012).

Confirmatory factor analysis (CFA) is the standard tool for verifying a model's structure (Rolland et al., 2010). But CFA often fails to support the five-factor model structure (Aluja et al., 2005). The limitations are well documented in many studies over the last decades (Church & Burke, 1994; McCrae et al., 1996; Furnham et al., 2012; Ispas et al., 2014). CFA can overcome the weaknesses of more complex model structures (Aluja et al., 2005). However, improper use of CFA will result in an incorrect number of personality factors (Furnham et al., 2012). The internal structure of the omnibus personality inventories, such as NEO-PI-R or HEXACO-PI-R, is not verifiable very well by CFA. The reasons are the inherent complexity of personality, issues with the personality measurement, and issues with CFA models'application and interpretation (Hopwood & Donnellan, 2010). Alternative procedures such as exploratory structural equation modelling have several advantages over CFA and can be more appropriate for personality data modelling (Marsh et al., 2010; Furnham et al., 2012).

Goal

The current study aims to examine the psychometric properties of the NEO-PI-R questionnaire in Slovak language. To achieve the goal, we applied the approach of classical item analysis theory (Coaley, 2010). On a convenient sample of 1062 adults, we analyzed reliability (internal consistency – Cronbach alpha, stability in time – correlation analysis) and construct validity (EFA, CFA) of NEO-PI-R in Slovak translation.

Materials and Methods

The research sample was collected between 2011 and 2018. It consists of university students of psychology and their relatives and acquaintances. First, the students completed the questionnaire. Second, they recruited their relatives and friends as research participants. Participation in the research was voluntary and anonymous. Some participants could pass their questionnaires together (self-reports and observer ratings). The researchers instructed the students about the eligibility criteria for participation in the research. The criteria were the following: mental health, a normal state of mind and no altered state of mind (tiredness, intoxication, illness). It should be mentioned that asking participants if they have no mental problems to participate in the study may not be enough, and this could have been a possible limitation in participant recruitment. All completed questionnaires were included in the research study. Students collecting data earned credits from psychology courses (introduction to psychology, personality psychology, and research practice). Participants agreed to process data for research purposes in an anonymized form of group analysis without identifying individuals.

The self-report research sample consisted of 1.062 persons, 237 men (22.5%), and 818 women (77.5%). Seven participants did not state their gender. The average age of this sample was 26.47 years, SD = 10.46, a minimum of 16, and a maximum of 70 years. One hundred and fifty participants did not state their age.

The re-test research sample consisted of 122 university students, 23 men (19%), and 99 women (81%). The average age of this sample was 19.76 years, SD = 1.2 (min 19 years and max 24 years). Fifty-four participants did not indicate their age. The students completed NEO-PI-R at the beginning of the semester and again after 12 weeks.

The observer ratings provided 371 people, 163 men (44%), and 208 women (56%). 98 were parents (26.8%), 130 were friends (35.5%), 66 were siblings (18%), and 72 were spouses/partners of those evaluated (19.7%). The average age of this sample was 35.9 years, SD = 13.98 (min 16, max 70 years).

In classical item theory, the correlation is a basic analysis for reliability and construct validity (Urbina, 2004). Power analysis of sample size estimation for correlation showed N = 46 (input parameters: effect size = .5; Power = .95). The sample size decreases the standard error of the normative mean in psychometrics (Coaley, 2010). It is good between 1000 and 1999 participants, and excellent if bigger than 2000. It should be no less than 600 participants (Coulacoglou & Saklofske,

2017). It should be more than 100 participants for reliability analysis (Kline, 2000). The smallest subject-to-item ratio for EFA purposes is 20: 1 (Costello & Osborne, 2005). NEO-PI-R includes 30 facets/items, which means that the research sample should consist of at least 600 participants.

Two researchers translated the questionnaire. Their independent translations were reconciled into a final version. In disagreement, the researchers reached a consensus to capture the most appropriate meaning. The authors of the current study supervised the theses with NEO-PI-R in Slovak (Benešová, 2012; Lednárová, 2012; Mikulíková, 2012; Pajtinková, 2011; Pažítková, 2011) as pilot studies. The NEO-PI-R items in Slovak are available from the authors upon reasonable request and with the agreement of the license holder (https://hogrefe.cz). Data were collected in a paper pen form. The administrative staff member of the faculty and researchers transcribed the data into MS Excel. The transcribers ran random checks of the transcription quality of the colleagues. Data were processed in IBM SPSS 22, JASP 0.16.1.0, and G*Power 3.1.9.4. Correlation analysis, Cronbach alpha analysis, EFA and CFA (N = 1062), test-retest reliability (N = 122), and comparison of self-assessment versus observer rating (N = 371) were performed.

Results

Reliability

The internal consistency coefficients of the domains ranged from. 92 (N) to .88 (O). The reliability of the facets ranged between .83 (N6-Vulnerability) and .27 (O6-Values). The highest internal consistency of observers'ratings was in C (.95) and the lowest in E (.89). In the retest sample, N showed a Cronbach's alpha of .92, and the lowest value had E (.86). Among the facets in the re-test sample, the highest internal consistency had N6 (.81) and the weakest O6 (.25). Among the facets, the highest Cronbach's α was in the S5-Self-Discipline (.84) and the lowest in O6 (.40). Stability over three months showed an average correlation of r = .81 and N (r = .85) was the most stable domain. The other domains showed approximately the same stability over time, at level r = .80. Table 3 presents all internal consistency values. Facets O6 and A6 did not reach satisfactory internal consistency.

Table 3 - *Coefficients of internal consistency and stability over time of the NEO-PI-R domains and facets*

NEO-PI-R Domains and facets		Reliability		
	Domain/Facet	α Whole sample N=1,062	α Observer ratings N=371	r Test-retest N=122
N	Neuroticism	.922	.913	.849
E	Extraversion	.890	.887	.802
O	Openness	.882	.890	.803
A	Agreeableness	.878	.922	.798
C	Conscientiousness	.907	.945	.801
The domain means		.896	.911	.811
N1	Anxiety	.783	.760	.729
N2	Angry Hostility	.749	.780	.711
N3	Depression	.799	.759	.753
N4	Self-Consciousness	.671	.659	.685
N5	Impulsiveness	.631	.663	.674
N6	Vulnerability	.826	.801	.788
E1	Warmth	.732	.785	.621
E2	Gregariousness	.798	.773	.784
E3	Assertiveness	.800	.730	.798
E4	Activity	.536	.405	.765
E5	Excitement-Seeking	.651	.635	.747
E6	Positive Emotions	.785	.798	.672
O1	Fantasy	.801	.745	.694
O2	Aesthetics	.808	.817	.746
O3	Feelings	.762	.780	.658

O4	Actions	.629	.631	.718
O5	Ideas	.763	.792	.747
O6	Values	.270	.398	.390
A1	Trust	.777	.798	.777
A2	Straightforwardness	.790	.781	.740
A3	Altruism	.717	.820	.615
A4	Compliance	.659	.764	.729
A5	Modesty	.750	.739	.747
A6	Tender-Mindedness	.475	.606	.572
C1	Competence	.612	.757	.540
C2	Order	.712	.766	.719
C3	Dutifulness	.685	.796	.768
C4	Achievement Striving	.720	.771	.767
C5	Self-Discipline	.785	.837	.781
C6	Deliberation	.790	.819	.738
Facet averages		.709	.732	.706

Construct validity

Observer ratings

The mean agreement score was r = .58 when controlled by the rater was r = .553 (Table 4). E showed the highest agreement between the observer's rating and self-assessment. N showed the lowest agreement. The siblings were the best at assessing the E domain. Friends were the worst in assessing the domain of C.

Table 4 - *Self-other agreement for NEO-PI-R domains*

	N	E	O	A	C	Average r
Entire sample (N = 371)	.463	.682	.656	.592	.529	.584
Parents (N = 98)	.437	.577	.494	.553	.526	.517
Friends (N = 130)	.412	.585	.638	.604	.406	.529
Siblings (N = 66)	.510	.739	.625	.538	.614	.605
Partners (N = 72)	.528	.611	.424	.559	.468	.518
Partial correlationa (N=371)	.458	.647	.567	.577	.516	.553

a type of rater controlled

Internal structure

We included 30 facets of the NEO-PI-R in the EFA. Analysis of the main components and Varimax rotation with Kaiser normalization extracted five factors (KMO = .871; Bartlett's sphericity test: Approx. Chi-square = 15860.080; df = 435; $p < .001$). The five factors explained the variance to 60.45% as follows: (1) N 20.28%; (2) C 14.58%; (3) A 11.51%; (4) E 8.51%; (5) O 5.58%. All facets saturated the main factors with factor loading equal to or greater than .40 (Table 5). The N and C facets had the highest loadings. Three facets saturated an extra factor, too: E1-Warmth saturated A, N5-Impulsiveness saturated E, and E3-Assertiveness saturated N. Table 5 shows Skewness and Kurtosis that reached values between -1 and 1.

The factors differed in their scores according to gender. Women scored higher in N, E, and A. EFA in the separated groups of men and women confirmed five factors that explained 59.4% and 60.7% of the total variance. The gender did not influence the results of the internal structure analysis.

Table 5 - *Rotated matrix, factor loadings of NEO-PI-R facets (Rotated Component Matrix)*

Facets		Factor				
		1	2	3	4	5
N1	Anxiety	.866	-.024	-.025	-.078	.065
N2	Angry Hostility	.750	-.003	-.383	.066	.018
N3	Depression	.821	-.190	.021	-.200	.066
N4	Self-Consciousness	.685	-.145	.071	-.396	-.082
N5	Impulsiveness	.408	-.311	-.197	.495	.265
N6	Vulnerability	.809	-.268	.062	-.041	-.031
E1	Warmth	-.112	.089	.459	.702	.120
E2	Gregariousness	-.096	-.071	.046	.792	-.098
E3	Assertiveness	-.430	.251	-.362	.487	.160
E4	Activity	-.302	.346	-.200	.461	.165
E5	Excitement-Seeking	-.095	-.251	-.180	.575	.154
E6	Positive Emotions	-.321	.046	.169	.603	.257
O1	Fantasy	.203	-.279	-.023	.179	.681
O2	Aesthetics	.113	.041	.134	-.020	.759
O3	Feelings	.330	.152	.119	.367	.636
O4	Actions	-.325	-.134	-.052	.232	.470
O5	Ideas	-.107	.120	-.068	-.098	.796
O6	Values	-.239	.025	.161	.165	.446
A1	Trust	-.278	.071	.559	.208	.085
A2	Straightforwardness	.013	.120	.762	-.157	-.049
A3	Altruism	.009	.233	.695	.314	.183
A4	Compliance	-.241	.039	.751	-.208	.013
A5	Modesty	.274	-.101	.537	-.196	-.232

A6	Tender-Mindedness	.120	.004	.578	.133	.263
C1	Competence	-.337	.683	.004	.110	.108
C2	Order	.085	.709	.006	-.089	-.011
C3	Dutifulness	.035	.767	.303	-.096	-.081
C4	Achievement Striving	-.159	.755	-.143	.194	.143
C5	Self-Discipline	-.277	.792	.107	-.020	-.053
C6	Deliberation	-.123	.559	.230	-.368	-.150
Factor's Skewness (SE)		.158 (.075)	-.176 (.075)	-.173 (.075)	-.218 (.075)	-.107 (.075)
Factor's Kurtosis (SE)		.016 (.150)	-.015 (.150)	.026 (.150)	-.006 (.150)	-.430 (.150)
Gender differences						
p		< .001	.144	< .001	< .001	.126
Cohen's d		-.398	-.108	-.398	-.402	-.113

Note: Extraction method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 8 iterations. Men coded as 1. Women coded as 2.

Table 6 shows the intercorrelations of the factors. N correlated with E and C negatively and E with O positively, with all values at the border of medium effect size (r ~.34). Factors correlated with age with a small effect size (Table 6).

Table 6 - *Domains'intercorrelations and their correlations with age*

	Extraversion	Openness	Agreeableness	Conscientiousness	Age
Neuroticism	-.349***	.069**	-.121***	-.359***	-.235***
Extraversion		.326***	-.009	.067**	-.182***
Openness			.079**	-.041	-.229***
Agreeableness				.191***	.100**
Conscientiousness					.215***

Confirmatory factor analysis

The structural parameters of the uncorrelated CFA model were significant at $p < .001$ ($X^2 = 7\ 075.804$; df = 405). The five-factor model provided a poor data fit: CFI = .572; TLI = .541; RMSEA = .125; SRMR = .175. As the next step, we tested individual models for each factor. Table 7 summarizes the fit statistics of all factor models. Individual factor models showed the following average values CFI .90; TLI = .84; RMSEA = .13 and SRMR = .06. Extraversion showed the worst data fit to the model.

Table 7 - *Goodness of fit indices for the five NEO-PI-R factors*

	X^2	p	CFI	TLI	RMSEA	SRMR
Neuroticism	175.357	$p < .001$	.947	.911	.132	.054
Extraversion	309.281	$p < .001$	.818	.697	.177	.079
Openness	122.751	$p < .001$	.914	.857	.109	.053
Agreeableness	187.051	$p < .001$	.878	.797	.136	.064
Conscientiousness	128.121	$p < .001$	.947	.912	.112	.044

Notes: df = 9 for all models

Discussion

NEO-PI-R in Slovak shows the best internal consistency in the N domain. The results are like in other foreign studies (Costa & McCrae, 1992; Hřebíčková et al., 2002; Terracciano, 2003; Källmen et al., 2010). We confirmed a high test-retest stability in all five factors. N showed the highest stability among the domains. In contrast, low internal consistency values were observed in facets O6 and A6. These facets in the Slovak translation are not suitable for individual interpretation. C was the most reliable dimension in the assessment by others. The internal consistency in observers'ratings reached higher values than in self-reports. The results correspond to the research from the USA and Czech Republic (Costa & McCrae, 1992; Hřebíčková et al., 2002). The most consistent facets were N6 and N3 (Costa & McCrae, 1992; Terracciano, 2003; Ispas et al., 2014). The stability testing over time confirmed an average value of .81, with N being the most stable domain.

The results confirmed costruct validity upon the agreement between self-assessment and observer's rating. The agreement was the best in assessing the E domain and the worst in assessing the N domain. This is

consistent with other findings (Hřebíčková & Urbánek, 2002; Funder & Dobroth, 1987; John & Robins, 1993). Family members and friends validly assessed the personality traits of close others through NEO-PI-R in Slovak language (Cheek, 1982; McCrae et al., 2004; Hřebíčková, 2002; Hoffman et al., 2007; Funder et al., 1995).

Factor analysis confirmed the five-factor structure of NEO-PI-R in Slovak translation (Caprara et al., 2001; Hřebíčková et al., 2002; Ispas et al., 2014; Martin et al., 2002; Rolland et al., 1998; Aluja et al., 2005; Hesselmark et al., 2015; Marshall et al., 2005). The CFA did not fit the five-factor model well. The results are typical for NEO-PI-R in other studies (Furnham et al., 2012; Vassend & Scrondall; 2011). An acceptable five-factor NEO-PI-R model appears to have a complicated structure (Aluja et al., 2005; McCrae et al., 1996). Many parameters, including several significant covariance factors, could explain the poor data fit. Including correlated error components can make significant improvements in the model data fit (Vassend et al., 2011; Furnham et al., 2012; Gignac, 2009). This procedure enables researchers to reduce the problematic items. However, the modified version of the NEO-PI-R was not the objective of the current study. Because of that, we did not conduct a further analysis. Models with more than five factors might better fit the data. But they are usually unacceptable for poorly defined factors or factors that cannot be replicated in other samples (Aluja et al., 2005). Leading researchers in the field decided to stick to the five-factor personality model (McCrae et al., 1996). Future research could explore alternatives to one-factor models on which NEO PI-R facets and domains are based (Vassend et al., 2011).

The research sample, consisting of psychology students, their relatives and friends, represents at the same time the limitation of the study. Most of the participants probably had higher education, which could not be confirmed as this variable was not collected. Limitations in results include low internal consistency for two facets with values below .50 and eight facets below .70. These results suggest that not all facets could be interpreted. On the valuable feedback from the reviewers, in future research, we would add to the collected participants'demographics the level of education, socioeconomic status, and the affective state (Terracciano, 2003). More demographics of the research sample would help us better understand how the research sample represents the population. McCrae and Costa (2003) point out the importance of temperamental differences due to problems with physical (brain injuries) or mental (depression) health. It would be useful to focus on personality profiles based on age or activity area in future research and verify the actual NEO-PI-3 questionnaire.

Conclusions

The NEO-PI-R in Slovak meets essential psychometric criteria for a standardized personality questionnaire. The results show that it is possible to standardize this method in the Slovak language.

Acknowledgments

The authors thank all participants involved in the research sample. The authors thank Alena Múčková for help in transcribing the data. The authors thank the reviewers for reviewing the manuscript that contributed to its quality.

Declaration of interest statement

The authors declare that they have no conflict of interest.

References

Aluja, A., García, O., García, L.F., & Seisdedos, N. (2005). Invariance of the "NEO-PI-R" factor structure across exploratory and confirmatory factor analyses. *Personality and Individual Differences, 38*, 1879-1889.

Benešová, A. (2012). Zhoda pri posudzovaní osobnosti partnerských dvojíc. [Consistency in the rating and self-assessment of personality in partners/spouses]. Bachelor thesis, Faculty of Social and Economic Sciences, Comenius University in Bratislava.

Branco e Silva, L., & Laher, S. (2012). Exploring the Utility of the NEO-PI-R in a Sample of South African University Students. *IFE Psychologia*, *20*(1), 19-48.

Caprara, G.V., Barbaranelli, C., Hahn, R., & Comrey, A.L. (2001). Factor analyses of the NEO-PI-R Inventory and the Comrey Personality Scales in Italy and the United States. *Personality and Individual Differences, 30*, 217-228.

Cheek, J.M. (1982). Aggregation, moderator variables, and the validity of personality tests: Apeer-rating study. *Journal of Personality and Social Psychology, 43*(6), 1254-1269.

Church, A. T., & Burke, P. J. (1994). Exploratory and confirmatory tests of the Big Five and Tellegen's three- and four-dimensional models. *Journal of Personality and Social Psychology*, 66, 93-114.

Coaley, K. (2010). *An Introduction to Psychological Assessment and Psychometrics*. London: Sage.

Costa, P.T., & McCrae, R.R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa: Psychological Assessment Resources.

Costa, P.T., Terracciano, A., Uda, M., Vacca, L., Mameli, C., Pilia, G., Zonderman, A.B., Lakatta, E., Schlessinger, D., & McCrae, R.R. (2006). Personality Traits in Sardinia: Testing Founder Population Effects on Trait Means and Variances. *Behavior Genetics, 37*, 376-387. DOI: 10.1007/s10519-006-9103-6.

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation, 10*(7), 1-9. DOI: 10.7275/jyj1-4868.

Coulacoglou, C. & Saklofske, D.H. (2017). Recent Advances in Psychological Assessment and Test Construction. In: *Psychometrics and Psychological Assessment* (3-25). Academic Press. DOI: 10.1016/B978-0-12-802219-1.00001-8.

Evers, A., Muñiz, J., Bartram, D., Boben, D., Egeland, J., Fernández-Hermida, J. R., Frans, Ö., Gintiliené, G., Hagemeister, C., Halama, P., Iliescu, D., Jaworowska, A., Jiménez, P., Manthouli, M., Matesic, K., Schittekatte, M., Sümer, H. C., & Urbánek, T. (2012). Testing practices in the 21st century: Developments and European psychologists'opinions. *European Psychologist, 17*(4), 300-319.

Funder, D.C, Kolar D.C., & Blackman, M.C. (1995). Agreement Among Judges of Personality: Interpersonal Relations, Similarity, and Acquaintanceship. *Journal of Personality and Social Psychology, 69*(4), 656-672.

Funder, D.C., & Dobroth, K.M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52*(2), 409-418.

Furnham, A., Guenole, N., Levine, S.Z., & Chamorro-Premuzic, T. (2012). The NEO Personality Inventory Revised: Factor Structure and Gender Invariance From Exploratory Structural Equation Modeling Analyses in a High-Stakes Setting. *Assessment, 20*, 14-23. DOI: 10.1177/1073191112448213.

Gignac, G.E. (2009). Partial confirmatory factor analysis: Described and illustrated on the NEO-PI-R. *Journal of Personality Assessment, 91*, 40-47. DOI: 10.1080/00223890802484126.

Hesselmark, E., Eriksson, J.M., Westerlund, J., & Bejerot, S. (2015). Autism Spectrum Disorders and Self-reports: Testing Validity and Reliability Using the NEO-PI-R. *Journal of Autism and Developmental Disorders, 45*, 1156-1166. DOI: 10.1007/s10803-014-2275-7.

Hoffman, P.D., Buteau, E., & Fruzzetti, A.E (2007). Borderline Personality Disorder: NeoPersonality Inventory Ratings of Patients and Their Family Members. *International Journal of Social Psychiatry, 53*(3), 204-215.

Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and social psychology review: an official journal of the Society for Personality and Social Psychology, 14*(3), 332-346. DOI: 10.1177/1088868310361240.

Hřebíčková, M. (2002). Vnitřní konzistence čeké verze NEO osobnostního inventáře (NEO-PI-R). [Internal consistency of the Czech version of the NEO personality inventory (NEO-PI-R)]. *Československá psychologie, 46*(6), 521-535.

Hřebíčková, M. (2004). *NEO osobnostní inventář*. [*NEO personality inventory*]. Praha: Testcentrum.

Hřebíčková, M., & Urbánek, T. (2002). Shoda při posuzování osobnosti. Consistency in personality assessment]. In F. Baumgartner, M., Frankovský, M., Kentoš (Eds.): *Sociálne procesy a osobnosť 2002. Zborník príspevkov*. Košice: Ústav experimentálnej psychológie SAV, 129 -133.

Hřebíčková, M., Urbánek, T., Čermák, I. (2002). *Psychometrické charakteristiky NEO osobnostního inventáře (NEO-PI-R) pro sebeposouzení a posouzení druhého*. [*Psychometric characteristics of NEO personality inventory (NEO-PI-R) for self-assessment and assessment of the other*]. Brno: Psychologický ústav Akademie věd ČR.

Ispas, D., Iliescu, D., Ilie, A., & Johnson, R.E. (2014). Exploring the Cross-Cultural Generalizability of the Five-Factor Model of Personality: The Romanian NEO PI-R. *Journal of Cross-Cultural Psychology*, 1-15. DOI: 10.1177/0022022114534769.

John, O.P., & Robins, R.W. (1993). Determinants of Interjudge Agreement on Personality Traits: The Big Five Domains, Observability, Evaluativeness, and Unique Perspective of the Self. *Journal of Personality, 61*(4), 521-551. DOI: 10.1111/j.1467-6494.1993.tb00781.x.

Källmen, H., Wennberg, P., & Bergman. H. (2010). Psychometric properties and norm data of the Swedish version of the NEO-PI-R. *Nordic Journal of Psychiatry,65*(5), 311-314. DOI: 10.3109/08039488.2010.545433.

Kline, P. (2000). *Handbook of psychological testing*. London: Routledge.

Lednárová, M. (2012). Zhoda pri posudzovaní osobnosti súrodeneckých dvojíc. [Consistency in the rating and self-assessment of personality in siblings]. Bachelor thesis, Faculty of Social and Economic Sciences, Comenius University in Bratislava.

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big-five factor structure through exploratory structural equation modeling. *Psychological Assessment*, 22, 471-491.

Marshall, M.B., De Fruyt, F., Rolland, J.P., & Bagby, R.M. (2005). Socially Desirable Responding and the Factorial Stability of the NEO PI-R. *Psychological Assessment, 17*(3), 379-384. DOI: 10.1037/1040-3590.17.3.379.

Martin, T.A., Costa, P.T., Oryol, V.E., Rukavishnikov, A.A., & Senin, I.G. (2002). Application of The Russian NEO-PI-R. In: R.R. McCrae et al. (Eds.), *The Five-Factor Model of Personality Across Cultures*, (261-277). New York: Kluwer Academic / Plenum Publishers.

McCrae, R. R., & Costa, P. T., Jr. (2003). *Personality in adulthood: A five-factor theory perspective* (2nd ed.). Guilford Press. DOI: 10.4324/9780203428412.

McCrae, R. R., & Sutin, A. R. (2009). Openness to experience. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 257-273). The Guilford Press.

McCrae, R. R., Zonderman, A. B., Costa, P. T., Jr., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the revised NEO personality inventory: Confirmatory factor analysis versus procrustes rotation. *Journal of Personality and Social Psychology*, *70*, 552-566.

McCrae, R.R., Costa, P.T., Martin, T.A., Oryol, V.E., Senin, I.G., et al. (2004). Consensual validation of personality traits across cultures. *Journal of Research in Personality, 38*(2), 179-201.

Mikulíková, S. (2012). Zhoda pri posudzovaní osobnosti. [Consistency in the rating and self-assessment of personality]. Bachelor thesis, Faculty of Social and Economic Sciences, Comenius University in Bratislava.

Pajtinková, T. (2011). Zhoda pri posudzovaní a sebaposudzovaní osobnosti u súrodencov. [Consistency in the rating and self-assessment of personality in siblings] Diplomová práca. Bratislava: Faculty of Social and Economic Sciences, Comenius University in Bratislava.

Pažítková, N. (2011). Zhoda medzi sebaposúdením a objektívnym posúdením osobnosti kamarátmi a rodičmi. [Consistency in the rating and self-assessment of personality in friends and parents]. Bachelor thesis, Faculty of Social and Economic Sciences, Comenius University in Bratislava.

Rolland, J.P., Parker, W.D., & Stumpf, H. (1998) A Psychometric Examination of the French Translations of NEO-PI-R and NEO-FFI, *Journal of Personality Assessment, 71*(2), 269-291. DOI: 10.1207/s15327752jpa7102_13.

Strelau, J. (2001). The concept and status of trait in research on temperament. *European Journal of Personality, 15*(4), 311-325. DOI: 10.1002/per.412.

Terracciano, A. (2003). The Italian version of the NEO PI-R: conceptual and empirical support for the use of targeted rotation. *Personality and Individual Differences, 35*(8), 1859-1872. DOI: 10.1016/S0191-8869(03)00035-7.

Urbina, S. (2004). *Essential of Psychological Testing*. Hoboken: John Wiley.

Vassend, O. & Skrondal, A. (2011). The NEO personality inventory revised (NEO-PI-R): Exploring the measurement structure and variants of the five-factor model. *Personality and Individual Differences, 50*, 1300-1304. DOI: 10.1016/j.paid.2011.03.002.