

The assessment of effects of partialling out of socially desirable responding variance on personality traits scores in instructed applicant situations

Goran Opačić¹ & Tatjana Mentus²

¹ *Faculty of Philosophy, University of Belgrade, Belgrade, Serbia*

² *Faculty for Special Education and Rehabilitation,
University of Belgrade, Belgrade, Serbia*

The aim of this study was to examine the extent to which the socially desirable responding (SDR) distorts results of HEDONICA personality inventory (acronym based on eight dimensions of this inventory: Honesty, Disintegration, Impulsiveness, Openness, Extraversion, Neuroticism, Conscientiousness, and Agreeableness). The inventory HEDONICA was merged with components of the Balanced Inventory of Desirable Responding (BIDR) as a control inventory and was administered to a sample of 227 students under two experimental situations/ contexts, operationalized by two instructions: the standard (S) one (such as “be honest”) and the “fake good” (FG) one (such as “portray yourself in a most positive way”). Comparing scores in S and FG situations by using MANOVA, a clear distortion on all personality traits in socially desirable directions were evidenced. When, however, the BIDR subscales in the FG situation were entered into MANOVA as covariates, differences between personality scores in S and FG situations were considerably reduced, and became statistically insignificant on five personality dimensions. When the variance of dimensions of the BIDR inventory was removed from the variance of HEDONICA traits in FG situation, the change between intercorrelations of personality dimensions in S and FG situations did not attain statistical significance. This led to the conclusion that the SDR bias, if even does affect test results (i.e., enhances scores in FG situation), does not affect the scale structure and predictive validity of the examined personality inventory.

Key words: control scale, faking correction, desirable responding, HEDONICA inventory

Corresponding author: mentus@fasper.bg.ac.rs

Acknowledgment. This research was supported by the Ministry of Education, Science, and Technological Development of Republic Serbia (project number 170918).

The authors acknowledge professor Goran Knežević from the University of Belgrade, Faculty of Philosophy who contributed greatly to this study providing computer resources with authorized software, as well as through valuable participation in the discussion of the results.

Highlights:

- The sensitivity of HEDONICA inventory to score manipulation was examined.
- Context was modified by changing test instructions.
- The Balanced Inventory of Desirable Responding was a control inventory.
- Partialling out the control scale variance did not improve test results.

The use of personality tests in applicant selection by various companies, has increased in the 1990s. (Hough & Ones, 2001). Namely, it was found that personality variables may predict job performance and organizational skills, and may serve as a measure of important life outcomes: occupational status, healthy lifestyle behaviours, etc. (Barrick & Mount, 1991; Ones, Viswesvaran, & Schmidt, 1993; Tett et al., 1991). The goal was to select people who are the most promising for various job requirements.

Since personality tests rely primarily on self-report measures, many authors were concerned about their susceptibility to manipulation with scores by test-takers. The reason for this concern was evidenced by experimental testing procedures using personality inventories under the instruction to either “fake good” (FG; for instance: portray yourself as a desirable candidate) or to “fake bad” (FB; for instance: portray yourself as an undesirable candidate), when clear differences between scores on various personality measures were evidenced (Barrick & Mount, 1996; Dunnett, Koun, & Barber, 1981; Furnham & Craig, 1987; Hough et al., 1990; Thornton & Gierasch, 1980). This concern would be meaningless if all examinees had identical personality characteristics. However, the actually existing individual differences may cause persons more prone to faking to deliberately alter their scores in a socially desirable manner and, subsequently, disturb the proper ranking of applicants in a situation of high-stakes testing (Mesmer-Magnus & Viswesvaran, 2006; Viswesvaran & Ones, 1999).

The inclination toward socially desirable responding (SDR) differs for individuals, and may be measured by psychological instruments constructed for that purpose. An older measure of socially desirable responding is the Marlowe-Crowne Social Desirability Scale (M-C SDR; Crowne & Marlowe, 1960). One more recent SDR measure, published by Paulhus (1984, 1994) titled the Balanced Inventory of Desirable Responding (BIDR). It consists of two factors, called self-deceptive enhancement (SDE) and impression management (IM). The SDE factor relates to the tendency of individuals to exaggerate unconsciously (unintentionally) positive attributes of themselves. The second, IM factor relates to the tendency of individuals to deliberately present a socially desirable image on themselves. These factors are measured by properly designed two 20-item subscales. The BIDR is robust and shows favorable internal consistency and test-retest reliability (Paulhus, 1994).

After further study, Paulhus and Reid (1991) revealed that the SDE factor may be divided into two subfactors, one relating to the tendency of individuals to emphasize positive attributes, and the other relating to the tendency of individuals to deny their own unfavorable characteristics, while the IM factor remained unchanged, constituting a three-factor inventory. Recently, Gignac (2013) reported results supporting this three-factor model. Moreover, Paulhus (2002, 2006), provided a more comprehensive four-factor model of socially desirable responding, according to which two main factors, related to moralistic and egoistic bias, are divided into their unconscious and conscious components. This model was operationalized by the Comprehensive Inventory of Desirable Responding (Paulhus, 2006),

Many researchers who use the BIDR nevertheless continue to work with the two originally conceived subscales. Thus, the two-factor BIDR questionnaire (Paulhus, 1984) is currently among the most frequently used questionnaires for SDR measurement, enjoying a widespread use across varied disciplines (Hart, Ritchie, Hepper, & Gebauer, 2015).

Studies aimed at answering the question of whether or not the SDR reduces the validity of personality variables, was a topic of many research papers. There is much of controversy in this area.

Based on their investigations, Butcher and Rouse (1996) recommended that personality questionnaires should include control scales to reveal the inclination of examinees to exaggerate desirable responses. Consequently, some authors used the SDR scales as control variables (Finkel, Burnette, & Scissors, 2007), as scale validity measures (Starzyk, Holden, Fabrigar, & MacDonald, 2006), or as primary variables of interest (Lalwani, Shavitt, & Johnson, 2006).

On the other hand, results of a number of studies did not support the use of SDR scales to improve the validity of personality scales. One of the reasons for this is based on the finding that there is non-zero correlation between social desirability scales and personality trait scales (McCrae & Costa, 1983; Nicholson & Hogan, 1990; Ones, Viswesvaran, & Reiss, 1996). A meta-analysis demonstrated that social desirability correlates with emotional stability and conscientiousness, with estimated population correlations of .18 and .13, respectively. If SDR is merely a source of measurement error, then its relationships with the external criteria should be close to zero. However, relationships between SDR and external criteria, particularly with various personality dimensions, have been found in many studies (Paulhus, 1991; Douglas et al., 1996; Dunnette et al., 1962; Rosse et al., 1998; Schmit & Ryan, 1993; Hough, 1998; Griffith & McDaniel, 2006). Some studies of the effects of social desirability on the criterion related validity of personality measures showed that social desirability moderates the criterion-related validities to some extent, but does not make it negligible (Barrick & Mount, 1996; Hough, 1998; Hough et al., 1990; Ones et al., 1996). Thus, social desirability neither presents any suppressor nor contributes significant variance to the relationship between

personality traits (Hough, 1998; Ones et al., 1996, Ones & Viswesvaran, 1998). A low to negligible increase in validity was found when SDR is controlled, and, in certain cases, the relationships between variables disappeared indicating that SDR behaves as a meaningful personality trait (McCrae & Costa, 1983; Ones et al., 1996).

After Ellingson, Sackett, and Hough (1999), as far as SDR behaves as a personality trait, treating SDR variance as error is questionable. In a study using a within-subjects design, they compared the honest, faked, and corrected personality trait scores, in order to evaluate whether a social desirability correction is really effective, but the correction failed to fulfil this expectation.

Hough, Eaton, Dunnette, Kamp, and McCloy (1990) concluded that job applicant-like individuals do not distort their responses in a considerable manner and personality scales could thus successfully be used in personnel selection. This questioned the construct validity of social desirability and the attempts to define whether social desirability constitutes error variance. The use of personality tests, particularly those based on Big Five models, in applicant selection is encouraged by findings based on meta-analyses, that their validity in prediction of individual traits is not too sensitive to various contexts (Hough & Ones 2001; Murphy & Dzieweczynski, 2005; Ones et al., 1996, Ones, Viswesvaran, & Dilchert, 2005). In a recent paper, Paunonen and LeBell (2012), by means of Monte-Carlo statistical analysis, supported the findings about only a minor decrease in criterion prediction accuracy, even if personality scores were massively infused with desirability bias.

A survey of the available literature shows that there researchers are uncertain about the effects of control scales on the validity of personality tests in applicant situations, and that new contributions would be useful.

The aim of this study was to examine the effects of partialing out of the variance of control scales on the validity of personality tests in applicant instruction conditions. HEDONICA personality inventory (HEDONICA; Knežević, 2008, 2014) was used. Two BIDR subscales – SDE and IM were used as a control inventory.

The strategy of the study was the following: if between HEDONICA dimensions measured in the two experimental situations – the “fake good” and the “standard” one, after partialling out the BIDR variance from the dimensions measured in the FG situation: a) correlations remain unchanged, the effect of response bias would be considered to be negligible; b) correlations increase, it could be concluded that response bias suppresses authentic responding; c) correlations decrease, it could be concluded that a substantial amount of variance is shared between the response bias and personality measures.

To answer the question stated by this strategy, zero-order correlations of HEDONICA scales in fake good and standard conditions, as well as correlation of corresponding HEDONICA scales in fake good and standard condition upon partialling out of BIDR scales variance in FG situation was performed, are compared performing ANCOVA calculations. The obtained results are presented in Results and Discussion sections.

Method

Sample

Participants of the study were first-year students of the Faculty of Special Education and Rehabilitation in Belgrade who were awarded pre-exam points for participation in the study. Since there were considerably more female than male students 227 vs 17, responses of male students were excluded from statistical procedures. Average age of included participants was 19.91 ($SD = 1.62$). The results of this study are, for this reason, generalizable to female population only. However, if a future study may be performed with male participants only, gender influence might be investigated, since the literature data indicate that the BIDR test is sensitive to gender (Bobbio & Manganelli, 2011).

Instruments

Two instruments were used: Paulhus's Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1988), and HEDONICA personality inventory (Knežević, 2008).

The BIDR (Paulhus, 1984) consists of 40 items, that measure two dimensions of socially desirable responding: Self-deception (SDE) and Impression management (IM). An already existing Serbian version of the original BIDR-6, Form 40A inventory was used in this study, adapted according to guidelines of the International Test Commission (2005). The basic concept of these guidelines is incorporated in the Serbian version of test adaptation guidelines (Hedrih, 2018). Psychometric characteristics of the Serbian version of this inventory were presented in Subotić, Dimitrijević, and Lovrić (2016).

As can be seen in table A1, in both conditions, both dimensions of BIDR scale display similar and satisfactory reliabilities.

Available literature shows that impression management, social desirability response and faking are often used as synonyms (see Rust & Golombok, 2009). Although some newer authors mention subtle but important differences are pointed out between these concepts, proposing for instance distinction between concepts of "image enhancement" and "self-deceptive enhancement" (Guion, 2011), we retained older concepts in this study.

HEDONICA personality inventory (HEDONICA; Knežević, 2008, 2014) was designed for the purpose of applicant selection in the public sector in Serbia. HEDONICA is based on the hierarchical model of relations between basic personality dimensions. This implies that the domain dimensions are based on lower range dimensions (modalities). The inventory measures eight personality dimensions. First five dimensions are taken from the NEO-PI-R (Costa & McCrae, 1995). These dimensions are; Neuroticism, Extraversion, Openness to experience, Agreeableness, and Conscientiousness, each consisting of six facets. Sixth dimension, Honesty (called alternatively Amoral¹), is measured through the scale composed of the following six facets: projection and rationalisation of amoral impulses, resentment, machiavelism, sadism, lust for revenge, and passive amorality. Seventh dimension is Disintegration which captures the inclination toward "psychosis" – DELTA-10, consisting of 10 facets: GEI (general executive dysfunctionality), PD (distorsion of perception), P (paranoia), D (depression), FA (level affect), SOD (somatoformic disregulation), EA (enhanced consciousness), MT (magic thinking), M (mania), and SA (social anhedonia). Eighth dimension is measured by the scale Impulsivity which consists of three subscales: low control of impulse, hedonism, and lazyness. For the last three dimensions, the result is displayed as a score on each subscale, as well as the total score. The inventory consists of 257 items, 150 of which originate from NEO-PI-R one and 107 from others. More details on the self-reporttest DELTA-10 may be found in

2 This dimension is separately considered in the reference Mentus, Opačić, & Knežević (2016).

Knežević, Savić, Kutlešić, and Opačić (2017), while more details about the psychometric properties and practical use of the instrument, may be found in a recent monograph by the author of HEDONICA (Knežević, 2014).

The validity of its factor structure was confirmed elsewhere (Knežević, 2008). Eight basic personality dimensions for the sample of this study were calculated as sum scores on items on each dimension of the personality test. Reliability coefficients were found to be satisfactory in both the standard and the fake good condition, see Table A1.

Procedure

The personality inventory and the social desirability inventory were merged and applied as a single questionnaire. Within the repeated measures research design, participants were asked to respond to the complete questionnaire under two instructions, which induced two different experimental situations (contexts).

The first instruction presented to the participants, providing what is called the standard (S) situation, was the following: “Please, read the questionnaire items carefully, and then give an answer that describes you mostly. You don’t have to think a lot about the meaning of the item. You will provide best answers if you give the answer that first came to your mind, after you are sure that you understood what the item means. Do respond to each item. If you made a mistake, just mark again the appropriate answer. Please give answers for all items. Do not miss any of them.”

The second instruction, providing what is called the fake good situation, was the following: “Imagine that you are applying for a job that you consider to be very attractive and you are likely to get it, and that getting the job depends only on the answers in this test. Give your answers in such a way that you maximize your chance for getting the job, presenting yourself in the best possible way”.

Although fake bad and fake good instructions may lead to asymmetric deviations of results relative to the standard situation, in this study we did not consider the instruction fake bad, in order to focus on the situation which is more attractive for practice, and to simplify the whole procedure.

Each subject responded to the battery of two instruments first in the standard, and then in the fake good instruction. It was not possible for participants to return to previous items, once they provided a final answer to each item. The time lag between the test under the first and under the second instruction was two weeks.

Items were presented one at a time, using personal computers equipped with 19-inch monitors. Participants answered test items by selecting a number on a 5-point Likert scale – 1 (*strongly disagree*) and 5 (*strongly agree*). Answers on each scale were collected automatically using the “Psycho” computer program (Knežević, 2014). Answers were stored in scrambled .csv files that were converted to the SPSS database format.

Only the BIDR, SDE, and IM scale results obtained in the S situation were used for further statistical analysis. As the BIDR author (Paulhus, 1991) recommended, the scoring key is balanced. After reversing negatively keyed items, one point is added for each extreme response (4 or 5), and an average score was calculated for the first 20 items (SDE scale) and then for the remaining 20 items (IM scale). Thus, total scores on SDE and IM scales range between zero and 20. Such scoring procedure provided that only participants who exaggerate desirable responses may attain high scores.

The results of statistical analysis are presented in Tables 1 to 5 and Table 1A.

Results and Discussion

At the beginning, the effect of the instructional set was examined, using MANOVA for Repeated measures, in order to examine whether the test scores on eight personality dimensions differ in standard (S) and “fake good” (FG) condition. The within-person factor was the test condition and dependent variables were personality dimensions (H – Amoral tendencies, E – Extraversion, D – Disintegration, O – Openness, N – Neuroticism, I – Impulsivity, C – Conscientiousness, A – Agreeableness). The results of calculations are presented in Tables 1.

Results in table 1 show that mean scores were higher in the FG than in the S situation for Extraversion, Openness, Conscientiousness, and Agreeableness scales. As expected, scores on Amoral tendencies, Disintegration, Neuroticism, and Impulsivity were lower in the fake good situation. The highest discrepancy between test scores was obtained on Extraversion, Impulsivity, Conscientiousness, and Neuroticism scales.

Table 1

Mean scores (M), standard deviations (SD) on HEDONICA self-report scales under S and FG conditions, F statistic of self-report scores in the S and FG situations, and partial eta squared, (η_p^2) as an effect size

	Standard		Fake Good		F	η_p^2
	M	SD	M	SD		
H	2.77	.64	2.31	.70	134***	.35
E	3.71	.59	4.34	.48	278***	.53
D	2.02	.56	1.60	.54	135***	.35
O	3.72	.52	3.94	.49	51***	.17
N	2.80	.66	1.82	.66	489***	.66
I	2.52	.55	1.76	.59	292***	.54
C	3.80	.48	4.60	.49	437***	.64
A	3.65	.49	4.13	.49	172***	.41

Note. H = Amoral tendencies; E = Extraversion, D = Disintegration; O = Openness; N = Neuroticism; I = Impulsivity; C = Conscientiousness; A = Agreeableness.

*** means F was determined for $p < .001$.

Results show that the instructional set had clear and expected effect on results of HEDONICA personality test. Differences between S and FG scores were obtained on all dimensions, in the expected way and size, i.e., the respondents changed their scores on the personality test in a socially desirable manner. These results are consistent with some previous studies that confirmed effects of instructed faking. Viswesvaran and Ones (1999) found that across the Big Five dimensions, effects of mean scale scores are about half of standard deviation for FG instruction. Effect size for mean difference between personality scores in S and FG condition found in this study, although somewhat higher (Table 1), were also similar to those presented in Ellingson, Sackett, and Hough (1999), who obtained the effect size of .51 for Neuroticism, .35 for Agreeableness,

and .53 for Conscientiousness. Along with Topping and O'Goorman (1997) and Holden, Wood, and Tomaszewsky (2001) results obtained in this study indicate that, relative to other personality dimensions, Openness (effect size .17) appears to be less susceptible to the socially desirable bias effect.

We can conclude that the problem of transparency of personality tests (Piedmont, McCrae, Reimann, & Angleitner, 2000), really exists, and individuals can effectively distort their responses in situations where it is advantageous to do so (Griffith & McDaniel, 2006; Viswesvaran & Ones, 1999; Birkeland et al., 2006; Griffith, Chmielowski, & Yoshita, 2007).

Table 2
Correlation coefficients of self-report HEDONICA scales under the S (upper triangle) and FG (lower triangle) conditions

	H	E	D	O	N	I	C	A
H	.56**	.25**	.54**	-.07	.39**	.65**	-.04	-.62**
E	-.19**	.41**	-.04	.44**	-.26**	.16*	.47**	-.05
D	.65**	-.42**	.48**	-.06	.76**	.63**	-.30**	-.35**
O	-.24**	.60**	-.26**	.53**	-.11	-.04	.37**	.19**
N	.58**	-.57**	.84**	-.31**	.44**	.52**	-.48**	-.18**
I	.74**	-.48**	.81**	-.34**	.86**	.26**	-.44**	-.63**
C	-.41**	.75**	-.64**	.51**	-.76**	-.77**	.23**	.15*
A	-.69**	.41**	-.56**	.43**	-.56**	-.73**	.57**	.33**

Note. H = Amoral tendencies; E = Extraversion, D = Disintegration; O = Openness; N = Neuroticism; I = Impulsivity; C = Conscientiousness; A = Agreeableness.

* and **mean statistical significance $p < .01$ and $p < .05$, respectively.

Table 2 illustrates original correlations of self-report HEDONICA scales under S and FG instruction conditions, i.e., before any SDR correction was performed. Correlations under the S condition were presented in top right part of the table (upper triangle), while correlations under the FG condition were presented in the bottom left part of the table (lower triangle). Correlations of self-report scores between corresponding HEDONICA scales are placed on the diagonal.

While in the standard situation HEDONICA dimensions correlate mutually in the expected way and to the expected extent, in the FG situation correlations between all dimensions were higher. Average correlations between scales for S and FG situations were .33 and .56, respectively. Our results confirm earlier findings that faking increases correlations between Big Five dimensions (Pauls & Crost, 2005; Schmit & Ryan, 1993; Ziegler & Buehner, 2009). One should note that in the FG situation some correlations are even higher than autocorrelations. Namely, as a rule, autocorrelations are higher than correlations between scales measuring different constructs. However, since inventory dimensions here are measured in different situations, autocorrelation coefficients are considerably reduced. Consequently, some correlation coefficients between different scales got an opportunity to remain above this limit.

Table 3
Correlations between HEDONICA traits and BIDR (self-deception enhancement (SDE) and impression management (IM)) scales in S and FG situations

	S		FG	
	SDE	IM	SDE	IM
H	-.12*	-.11	-.39**	-.47**
E	.10	.03	.56**	.50**
D	-.16*	-.09	-.51**	-.52**
O	.09	.03	.30**	.28**
N	-.15*	-.05	-.61**	-.59**
I	-.06	-.06	-.55**	-.63**
C	.14*	.10	.57**	.59**
A	.10	.11	.44**	.62**

Note. H = Amoral tendencies; E = Extraversion, D = Disintegration; O = Openness; N = Neuroticism; I = Impulsivity; C = Conscientiousness; A = Agreeableness.

* and **mean statistical significance $p < .05$ and $p < .01$, respectively.

Table 3 shows correlations between HEDONICA personality traits and SDE and IM scores of the BIDR inventory in both S and FG test situations. In the standard situation, SDE scores of the BIDR correlate with H, D, N, and C traits significantly, while IM scores do not correlate significantly with any of the HEDONICA traits. Under the FG condition, correlations between personality traits and SDR scores were all significant and considerably higher from those obtained for the standard condition. As already reported, (Ellingson, Smith, & Sackett 2001; Smith & Ellignson, 2002; Marshall, De Fruyt, Rolland, & Bagby 2005), this fact has implications for the factor structure of the applied instrument.

In the next step of this study, ANCOVA for Repeated Measures was used first in order to obtain differences in personality test scores for eight dimensions (i.e., H, E, D, O, N, I, C, A) in both standard and fake good conditions. Then, personality scores in both conditions were entered into MANCOVA for repeated measures as the within-group factor, while the BIDR scales (i.e., SDE and IM) in the FG condition were entered as covariates. The calculated differences between personaity dimensions, when SDE and IM variances were partialled out from FG scores on personality traits, are presented in Table 4.

Data presented in Table 1, show that before the variance of social desirability was partialled out from personality measures, differences on all personality test scores were statistically significant. However, as now shown in Table 4, after SDE and IM variances were partialled out from FG scores of personality traits, differences on all personality traits were considerably reduced, and became statistically insignificant even in cases of H, D, O, I, and A. Partialling out of the SDE variance caused a more pronounced decrease of statistical indicator sizes than did the partialling out of IM variance. From these results one could conclude that impression management, as a component of socially desirable responding, is a fairly transparent task with a sizeable effect on personality scores, while self-deception enhancement apparently does not have such an obvious effect.

Table 4

Results of MANCOVA indicating differences (F -test, p , η^2) between scores on personality dimensions in S and FG situations afereither SDE or IM variances were partialled out from personality trait scores in the FG situation

	FG score corrected for SDE			FG score corrected for IM		
	F	p	η^2	F	p	η^2
H	.00	.98	.00	26	.00	.10
E	4.6	.03	.02	6.6	.01	.03
D	1.1	.30	.00	20	.00	.07
O	.80	.37	.00	4.6	.03	.02
N	2.3	.13	.01	27	.00	.10
I	4.4	.04	.02	25	.00	.09
C	2.4	.12	.01	14	.00	.05
A	.23	.63	.00	42	.00	.14

Note. H = Amoral tendencies; E = Extraversion, D = Disintegration; O = Openness; N = Neuroticism; I = Impulsivity; C = Conscientiousness; A = Agreeableness.

Results of this analysis are in line with expectations that BIDR scales act as validity scales. However, the crucial test of this expectation should be the answer to the question: does the removal of the variance of BIDR scales under the FG condition increase correlations between the corresponding personality measures in S and FG situations? Partial correlations, expectedly, should be higher than the corresponding zero-order correlations if social desirability, operating through BIDR scores, operates as a suppressor. To test this hypothesis, zero-order and corrected (partial) correlations of corresponding HEDONICA scales under FG and S conditions are compared. Differences between the two correlation coefficients, zero-order and partial ones, are calculated by the algorithm proposed by Raghunathan, Rosenthal, and Rubin (1996) for non-overlapping dependent r . The obtained results in terms of Z statistic and its significance (1-tailed p) are presented in Table 5.

Table 5

Intercorrelations of corresponding HEDONICA personality dimensions in the S and the FG situation and statistical significance of the difference between them in terms of Fisher Z statistics (Z , 1-tailed p)

	r	r -corrected	Z	1-tailed p
H	.56**	.58	2.84	.00
E	.41**	.43	.13	.45
D	.48**	.51	1.36	.09
O	.53**	.53	1.17	.12
N	.44**	.46	.93	.18
I	.26**	.31	1.39	.08
C	.24**	.20	-1.58	.06
A	.33**	.37	.37	.36

Note. H = Amoral tendencies; E = Extraversion, D = Disintegration; O = Openness; N = Neuroticism; I = Impulsivity; C = Conscientiousness; A = Agreeableness. R is taken from the diagonale of the Table 2 and r -corrected is corrected value obtained after partialling out BIDR's scale variance; label ** means statistical significance $p < .05$.

One may see from Table 5 that after SDE and IM variances from the FG condition were partialled out, only the correlation between scores on H in two situations slightly increased, and correlations between scores on C decreased almost significantly. Other changes, in general, did not reach statistical significance. Apparently, the pattern of the observed correlations among personality traits is different for different experimental conditions, but the underlying pattern of correlations obtained by partialling out the response bias variance, remains unchanged. This supports the assumption of multiple “levels” in self-assessment. This result confirms an earlier finding that the effect of using validity scales to “correct” possible biases in personality scores had no effect on validity or even decreased it (Piedmont, McCrae, Reimann, & Angleitner, 2000). Moreover, many studies examining correlations between control scales and external criteria (Paulhus, 1991; Douglas et al., 1996; Dunnette et al., 1962; Rosse et al., 1998; Schmit & Ryan, 1993; Hough, 1998; Griffith & McDaniel, 2006) reported that removing this variance could jeopardize the validity.

Let us recall (the discussion of Table 2) that average absolute correlations between scales for the S and the FG situation were .33, and .56, respectively. After controlling for the BIDR variance, the average correlation between the FG personality scores decreased to .42 but the pattern of correlations remained very similar to that obtained in the FG situation. Similarly, Galić and Jerneić (2013) found that intercorrelations between personality measures obtained after removing SDR variance are more or less similar to those obtained in the FG situation, leading to the conclusion that the correction for SDR response had no effect on personality measures.

The results of the present study are also in line with some other previously published results: (Hough & Ones 2001; Murphy & Dzieweczynski, 2005; Ones et al., 1996, Ones et al., 2005) indicating that validity of personality tests for prediction of personal properties is not too sensitive to various contexts. In other words, a control for socially desirable responding in personality test scores does not harm operational validities, which should mean that social desirability is neither a mediator nor a suppressor variable in personality-performance.

Limitations of the Study

In this section we list some constraints of this study, which may suggest some further investigations on this topic.

First of all, an insufficient number of male participants was the reason why only female participants were included. Thus the conclusions of this study may be generalizable to female population only. This suggests that, in the future, study should be repeated with male participants only, to observe the effects of gender separately, as well as to unite the tests results, i.e., to observe the sample with balanced gender composition. The relatively low number of participants (227) is also a disadvantage.

Furthermore, since the examinations performed in this study were carried out in an experimental (instructed) situation, results obtained in the standard

situation may not be considered as a true representative of the real personnel selection situation. We suggest that results may be further extended and probably improved if results from personnel selection situation are compared to the ones obtained in the FG situation, although no substantially different conclusions may be expected.

Neglecting to consider the fake bad situation, with an intent to focus attention on the practically more important fake good situation, may be added to the limitations of this study.

Finally, in this study an older, two-factor version of Paulhus's model was used, as one of widely used SDR models. Having in mind, as emphasized in the Introduction section, that Paulhus provided a more comprehensive four-factor SDR model operationalized with the Comprehensive Inventory of Desirable Responding (Paulhus, 2006), this study might be repeated on the basis of this new model with more precisely differentiated factor structure.

Conclusion

In this study, effects of partialling out of control scales variance from the personality test variance, on the validity of personality tests were examined, using the HEDONICA inventory to test personality, and the components of the Balanced Inventory of Social Desirable Responding inventory were used as control scales. Different contexts were achieved by use of the two instructions: the standard and the fake good one.

Initial MANOVA confirmed that the instructional set had a clear and expected effect on results of the personality test, i.e., participants changed their scores on the personality test in a socially desirable manner. This confirmed the susceptibility of self-report scores to intentional distortion, and changed the factor structure by increasing correlations between test scores. Results of this analysis are in line with expectations that BIDR scales are functioning as validity scales.

To answer the question: whether the removal of the variance of BIDR scales in fake good condition increases the correlations between the correspondent personality measures in the S and the FG conditions, zero-order correlations of HEDONICA scales in fake good and standard conditions and corrected (partial) correlation of corresponding HEDONICA scales in the fake good and the standard condition were compared after MANCOVA calculations, and the correlations remain unchanged, i.e., the effect of response bias was negligible. Also, this comparison did not confirm the hypothesis that partialling out of the variance of control scales would improve their validity in personnel selection. Thus, results of this study indicate that HEDONICA inventory may be used without any consideration of the socially desirable responding.

Constraints of this study, which might be mitigated by the further studies, are: the single-gender and relatively small sample (227 subjects), the experimental instead of the real application context, the omission from consideration of the multi-factor BIDR versions, and the omission from the consideration of fake bad situation.

References

- Bagby, R. M., & Marshal, M. B. (2003). Positive Impression Management and Its Influence on the Revised NEO Personality Inventory: A Comparison of Analog and Differential Prevalence Group Designs. *Psychological Assessment, 15*(3), 333–339. <http://dx.doi.org/10.1037/1040-3590.15.3.333>
- Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81*, 261–272.
- Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment, 14*(4), 317–333. <https://doi.org/10.1111/j.1468-2389.2006.00354.x>
- Bobbio, A., Manganelli, A. M. (2011) Measuring social desirability responding. A short version of Paulhus' BIDR 6. *Testing, Psychometrics Methodology in Applied Psychology, 18*(2), 117–135. (<http://goo.gl/0xSsEo>)
- Borkenau, P., & Ostendorf, F. (1992). Social desirability scales as moderator and suppressor variables. *European Journal of Personality, 6*, 199–214. <http://dx.doi.org/10.1002/per.2410060303>
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI-2 manual for administration and scoring*. Minneapolis: University of Minnesota Press.
- Butcher, J. N., & Rouse, S. V. (1996). Personality: Individual differences and clinical assessment. *Annual Review of Psychology, 47*, 87–111. <http://dx.doi.org/10.1146/annurev.psych.47.1.87>
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349–354. <http://dx.doi.org/10.1037/h0047358>
- Douglas, E. F., McDaniel, M. A. & Snell, A. F. (1996, August). *The validity of non-cognitive measures decays when applicants fake*. Paper presented at the annual meeting of the Academy of Management, Cincinnati, OH.
- Dunnnett, S., Koun, S., & Barber, P. J. (1981). Social desirability in the Eysenck Personality Inventory. *British Journal of Psychology, 72*, 19–26. <http://dx.doi.org/10.1111/j.2044-8295.1981.tb02158.x>
- Dunnette, M. D., McCartney, J., Carlson, H. C., & Kirchner, W. K. (1962). A study of faking behavior on forced-choice self-description checklist. *Personnel Psychology, 15*, 13–24.
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden Press.
- Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology, 84*, 155–166. [10.1037/0021-9010.84.2.155](http://dx.doi.org/10.1037/0021-9010.84.2.155)
- Ellingson, J. E., Smith, D. B., & Sackett, P. R. (2001). Investigating the influence of social desirability on personality factor structure. *Journal of Applied Psychology, 86*, 122–133. <http://dx.doi.org/10.1037/0021-9010.86.1.122>
- Furnham, A., & Craig, S. (1987). Fakeability and correlates of the Perception and Preference Inventory. *Personality and Individual Differences, 8*, 459–470. [http://dx.doi.org/10.1016/0191-8869\(87\)90207-8](http://dx.doi.org/10.1016/0191-8869(87)90207-8)
- Finkel, E. J., Burnette, J. L., & Scissors, L. E. (2007). Vengefully ever after: Destiny beliefs, state attachment anxiety, and forgiveness. *Journal of Personality and Social Psychology, 92*, 871–886. <http://dx.doi.org/10.1037/0022-3514.92.5.871>

- Galić Z., & Jerneić Ž., (2013). Measuring faking on Five-Factor personality questionnaires The usefulness of Communal and Agentic Management Scales, *Journal of Personnel Psychology* 12, 115–123. <http://dx.doi.org/10.1027/1866-5888/a000087>
- Gignac G. E. (2013). Modeling the Balanced Inventory of Desirable Responding: Evidence in Favor of a Revised Model of Socially Desirable Responding, *Journal of Personality Assessment*, 95, 645–656. <http://dx.doi.org/10.1080/00223891.2013.816717>
- Goffin, R. D., & Christiansen, N. D. (2003). Correcting Personality Test for Faking: A Review of Popular Personality Test and Initial Survey of Researchers. *International Journal of Selection and Assessment*, 11(1/2), 340–344. doi: 10.1111/j.0965-075X.2003.00256.x
- Griffith, R. L., Chmielowski, T., & Yoshita, Y. (2007). Do applicants fake? An examination of the frequency of applicant faking behaviour. *Personnel Review*, 36, 341–355. <http://dx.doi.org/10.1108/00483480710731310>
- Griffith, R. L., & McDaniel, M. A. (2006). The nature of deception and applicant faking behavior. In R.L. Griffith (Ed.). *A closer examination of applicant faking behavior*. Greenwich, CT: Information Age Publishing.
- Guion, R. M. (2011). Assessment, measurement and prediction for personnel decisions (2nd ed.). New York, NY, US: Routledge/Taylor&Francis Group
- Hart, C. M., Ritchie, T. D., Hepper, E. G., & Gebauer, J. E. (2015). The Balanced Inventory of Desirable Responding Short Form (BIDR-16). *SAGE Open*, 1-9. <https://doi.org/10.1177/2158244015621113>
- Hedrih, V. (2018). *Kroskulturalna adaptacija psiholoških mernih instrumenata [Cross-cultural adaptation of psychological measurement instruments]*, Niš: Filozofski fakultet.
- Holden, R. R. (2007). Socially desirable responding does moderate personality scale validity both in experimental and in non-experimental contexts. *Canadian Journal of Behavioural Science*, 39(3), 184–201. <http://dx.doi.org/10.1037/cjbs2007015>
- Hough, L. M., & Ones, D. S. (2001). The structure, measurement, validity, and use of personality variables in industrial, work and organizational psychology. In: D.S. Anderson, H. K.Ones, & C. V., Sinangil (Eds.), *Handbook of industrial work and organizational psychology* (Vol. 1, pp. 233–377). London: Sage.
- Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology*, 75, 581–595. <http://dx.doi.org/10.1037/0021-9010.75.5.581>
- International Test Commission. (2017). *The ITC Guidelines for Translating and Adapting Tests (Second edition)*. [www.InTestCom.org]
- Hart, C. M., Ritchie, T. D., Hepper, E.G., & Gebauer, J. E. (2015) The Balanced Inventory of Desirable Responding Short Form (BIDR-16), *SAGE Open*, 1–9, doi: 10.1177/2158244015621113
- Knežević G. (2008). HEDONICA – Battery of Personality Tests. Belgrade: Ministry of Interior of Serbia: Technical report
- Knežević, G., Savić, D., Kutlešić, V., & Opačić, G. (2017) Disintegration: A reconceptualization of psychosis proneness as a personality trait separate from the Big Five, *Journal of Research in Personality* 70, 187–201. <https://doi.org/10.1016/j.jrp.2017.06.001>
- Knežević, G. (2014). *Profesionalna selekcija, elementi, izazovi i jedno praktično rešenje. [Professional selection, concepts, challenges and one of practical solutions]* Beograd: Centar za primenjenu psihologiju.
- Komar, S., Brown, D. J., Komar, J. A., & Robie, C. (2008). Faking and the validity of conscientiousness: a Monte Carlo investigation. *Journal of Applied Psychology*, 93(1), 140–154. doi: 10.1037/0021-9010.93.1.140
- Lalwani, A. K., Shavitt, S., & Johnson, T. (2006). What is the relation between cultural orientation and socially desirable responding? *Journal of Personality and Social Psychology*, 90, 165–178. <http://dx.doi.org/10.1037/0022-3514.90.1.165>

- 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), 379–84. doi:10.1037/1040-3590.17.3.379
- McCrae, R. R., & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology, 51*, 882–888. doi: 10.1037//0022-006X.51.6.882
- Costa, P. T., McCrae R.R., & Robert R. (1985). *The NEO personality inventory manual*. Odessa, FL: Psychological Assessment Resources.
- McCrae R. R., Costa P. T., & Martin T. A. (2005). “The NEO PI-3: A more readable revised NEO personality inventory”. *Journal of Personality Assessment, 84*(3), 261–270. doi:10.1207/s15327752jpa8403_05.
- McCrae, R. R., & Costa, P. T., Jr. (2010). *NEO Inventories: Professional manual*. Lutz, FL: Psychological Assessment Resources, Inc.
- Mentus T., Opačić G., & Knežević G. (2016). The investigation of the manipulation with scores on Amoral dimension of the HEDONICA inventory. *Psihologija, 49*(2), 129–148. doi: 10.2298/PSI1602129M
- Mesmer-Magnus, J., Viswesvaran, C., Deshpande, S., & Joseph, J. (2006). Social desirability: the role of over-claiming, self-esteem, and emotional intelligence. *Psychology Science, 49*, 336–356.
- Murphy, K. R., & Dzieweczynski, J. L. (2005). Why don't measures of broad dimensions of personality perform better as predictors of job performance? *Human Performance, 18*(4), 343–357. http://dx.doi.org/10.1207/s15327043hup1804_2
- Nicholson, R. A., & Hogan, R. (1990). The construct validity of social desirability. *American Psychologist, 45*, 290–292. <http://dx.doi.org/10.1037/0003-066X.45.2.290>
- Ones, D. S., Viswesvaran, C., & Dilchert, S. (2005). Personality at work: Raising awareness and correcting misconceptions. *Human Performance, 18*(4), 389–404. doi: 10.1207/s15327043hup1804_5
- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance [Monograph]. *Journal of Applied Psychology, 78*, 679–703. doi:10.1037/0021-9010.78.4.679
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology, 81*, 660–679. <http://dx.doi.org/10.1037/0021-9010.81.6.660>
- Ones, D. S., & Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance, 11*, 245–269. <https://doi.org/10.1080/08959285.1998.9668033>
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*, 598–609. doi:10.1037/0022-3514.46.3.598
- Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology, 60*, 307–317, <http://dx.doi.org/10.1037/0022-3514.60.2.307>
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of Personality and Social Psychological Attitudes* (pp. 17–59). San Diego, CA: Academic Press.
- Paulhus, D. L. (1994). *Balanced inventory of desirable responding: Reference manual for BIDR version 6*, Department of Psychology, University of British Columbia, Vancouver, Canada: Unpublished manuscript.
- Paulhus, D. L. (2002). Social desirable responding: The evolution of a construct. In: H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69), Mahwah NJ: Erlbaum.

- Paulhus, D. L. (2006). A Comprehensive Inventory of Desirable Responding (CIDR). Poster presented at the meeting of the Association for Research in Personality, New Orleans.
- Pauls, C. A., & Crost, N. W. (2005). Cognitive Ability and Self-Reported Efficacy of Self-Presentation Predict Faking on Personality Measures. *Journal of Individual Differences*, 26(4), 194–206 <http://dx.doi.org/10.1027/1614-0001.26.4.194>
- Paunonen S. V., & LeBel E. P. (2012) Socially desirable responding and its elusive effects on the validity of personality assessments. *Journal of Personality and Social Psychology*. 103, 158–75. doi: 10.1037/a0028165.
- Piedmont, R. L., McCrae, R. R., Riemann, R., & Angleitner, A. (2000). On the invalidity of validity scales: Evidence from self-reports and observer ratings in volunteer samples. *Journal of Personality and Social Psychology*, 78, 582–593. <http://dx.doi.org/10.1037/0022-3514.78.3.582>
- Raghunathan, T. E., Rosenthal, R., & Rubin, D. B., (1996). Comparing correlated but nonoverlapping correlations. *Psychological Methods*, 1, 178–183. <http://dx.doi.org/10.1037/1082-989X.1.2.178>
- Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of preemployment personality testing and hiring decisions. *Journal of Applied Psychology*, 83, 634–644. <http://dx.doi.org/10.1037/0021-9010.83.4.634>
- Rust, J., & Golombok, S. (2009). *Modern psychometrics: The science of psychological assessment*. (3rd ed.). New York, NY, US: Routledge/Taylor & Francis Group
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262–274.
- Schmit, M. J., & Ryan, A. M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology*, 78, 966–974. <http://dx.doi.org/10.1037/0021-9010.78.6.966>
- Smith, D. B., & Ellingson, J. E. (2002). Substance vs. Style: A new look at social desirability in motivating contexts. *Journal of Applied Psychology*, 87, 211–220.
- Starzyk, K. B., Holden, R. R., & Fabrigar, L. R. (2006). The personal acquaintance measure: A tool for appraising one's acquaintance with any person. *Journal of Personality and Social Psychology*, 90, 833–847. doi: 10.1037/0022-3514.90.5.833
- Subotić, S., Dimitrijević, S., & Radetić Lovrić S., (2016). Psychometric evaluation and short form development of the Balanced Inventory of Desirable Responding (BIDR–6), *Psihologija* 49(3), 263–276. doi: 10.2298/PSI1603263S
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703–742. <https://doi.org/10.1111/j.1744-6570.1991.tb00696.x>
- Thornton, G. C., III., & Gierasch, P. F., III. (1980). Fakability of an empirically derived selection instrument. *Journal of Personality Assessment*, 44, 48–51. http://dx.doi.org/10.1207/s15327752jpa4401_7
- Topping, G. D., & O’Gorman, J. G. (1997). Effects of faking set on validity of the NEO-FFI. *Personality and Individual Differences*, 23(1), 117–124. [http://dx.doi.org/10.1016/S0191-8869\(97\)00006-8](http://dx.doi.org/10.1016/S0191-8869(97)00006-8)
- Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational & Psychological Measurement*, 59, 197–210. <http://dx.doi.org/10.1177/00131649921969802>
- Zickar, M. J., & Drasgow F. (1996). Detecting faking on a personality instrument using appropriateness measurement. *Applied Psychological Measurement*, 20, 71–87. <http://dx.doi.org/10.1177/014662169602000107>

Procena efekata isključivanja varijanse društveno poželjnih odgovora iz skorova testova ličnosti u situacijama instruiranih ispitanika

Goran Opačić¹ & Tatjana Mentus²

¹*Filozofski fakultet, Univerzitet u Beogradu, Beograd, Srbija*

²*Fakultet za specijalnu edukaciju i rehabilitaciju, Univerzitet u Beogradu, Beograd, Srbija*

Cilj ove studije bilo je ispitivanje stepena u kom socijalno poželjno odgovaranje menja rezultate inventara ličnosti HEDONICA (akronim je zasnovan na osam dimenzija ovog inventara: Poštenje, Dezintegracija, Impulsivnost, Otvorenost, Ekstraverzija, Neuroticizam, Savesnost i Saradljivost). Inventar HEDONICA je spojen sa delovima Balansiranog inventara poželjnog odgovaranja (eng. Balanced Inventory of Desirable Responding, BIDR) koji je korišćen kao kontrolni inventar i zadat je uzorku od 227 studenata u dve eksperimentalne situacije, odnosno konteksta operacionalizovanog kroz dve instrukcije: standardnu (S; od učesnika u istraživanju je traženo da budu iskreni) i situaciju „lažnog dobrog“ (eng. fake good, FG) predstavljanja (učesnicima u istraživanju je data instrukcija da se prikažu u najpozitivnijem svetlu). Poređenje skorova dobijenih u S i FG situacijama korišćenjem MANOVA-e pokazalo je jasnu promenu skorova na svim osobinama ličnosti u društveno poželjnom smeru. Međutim, kada su BIDR subskele u FG situaciji dodate u MANOVA-u kao kovarijeteti, razlike između skorova osobina ličnosti u S i FG situacijama su se znatno smanjile, pa čak i postale statistički neznačajne na pet osobina ličnosti. Kada je varijansa dimenzija BIDR inventara iz FG situacije uklonjena iz varijanse HEDONICA crta ličnosti, interkorelacije dimenzija ličnosti u S i FG situacijama nisu se statistički značajno promenile. Ovo vodi ka zaključku da pristrasnost usled socijalno poželjnog odgovaranja, čak i ako utiče na rezultat testa (tj. popravlja skorove u FG situaciji) ne utiče na strukturu skala niti na prediktivnu validnost ispitivanog inventara ličnosti.

Ključne reči: kontrolna skala, korekcija za lažno predstavljanje, društveno poželjno odgovaranje, HEDONICA inventar

RECEIVED 07.10.2018.
REVISED 19.01.2019.
ACCEPTED 20.02.2019.

© 2019 by authors



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International license

Appendix

Table A1

Psychometric characteristics of instrument scales/sample system used in this study for both the standard and the "fake good" situation, determined by using the software RTT10G¹⁴. The meaning of column labels is listed below the table.

	Standard												
	PSI1	PSI2	L	Alpha	L6	M11	Beta	Beta6	Gama	RHO	H1	H2	H5
H	.87	.71	.81	.84	.87	.82	.85	.88	.88	.88	.15	.46	.49
E	.86	.68	.80	.83	.86	.81	.84	.87	.87	.88	.14	.44	.54
D	.85	.70	.81	.83	.86	.81	.84	.87	.87	.87	.14	.48	.52
O	.88	.76	.81	.84	.86	.82	.85	.87	.87	.87	.15	.52	.47
N	.87	.71	.82	.85	.88	.82	.85	.88	.88	.88	.16	.47	.50
I	.81	.60	.79	.82	.85	.80	.82	.86	.86	.86	.13	.42	.47
C	.86	.69	.80	.83	.86	.81	.84	.87	.87	.88	.14	.45	.50
A	.87	.71	.81	.84	.87	.82	.84	.87	.87	.87	.15	.46	.49
SDE	.90	.78	.79	.83	.86	.80	.84	.87	.87	.87	.20	.60	.57
IM	.76	.63	.69	.73	.75	.71	.74	.77	.77	.77	.12	.56	.56
	FG												
H	.91	.79	.84	.87	.89	.84	.87	.89	.89	.90	.18	.54	.51
E	.89	.73	.82	.84	.87	.83	.86	.89	.89	.90	.15	.48	.53
D	.89	.75	.82	.85	.88	.83	.86	.89	.89	.89	.16	.48	.48
O	.84	.62	.79	.82	.85	.80	.83	.87	.87	.87	.13	.40	.44
N	.86	.64	.80	.83	.87	.81	.84	.88	.88	.88	.14	.42	.49
I	.85	.72	.80	.82	.85	.80	.83	.86	.86	.86	.13	.48	.47
C	.84	.65	.79	.82	.86	.80	.83	.87	.87	.87	.13	.43	.50
A	.79	.55	.76	.79	.83	.77	.80	.84	.84	.84	.11	.37	.46
SDE	.82	.64	.75	.79	.81	.76	.80	.83	.83	.83	.15	.53	.52
IM	.81	.60	.72	.76	.79	.73	.77	.81	.81	.81	.13	.47	.50

Note. H = Amoral tendencies; E = Extraversion; D = Disintegration; O = Openness; N = Neuroticism; I = Impulsivity; C = Conscientiousness; A = Agreeableness; SDE = Selfdeception Enhancement; IM = Impression Management.

The meaning of column labels:

Measures of sampling adequacy

PSI1-Kaiser-Mayer-Olkin coefficients^{2,3,4}

PSI2- Kaiser- Rice coefficients⁵

- Cerny, C. A., & Kaiser, H. F. (1977). A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multivariate Behavioral Research*, 12(1), 43–47. doi:10.1207/s15327906mbr1201_3
- Kaiser, H. (1974). An index of factor simplicity. *Psychometrika*, 39, 31–36. doi.org/10.1007/BF02291575
- Tenjović, L., & Radovanović, V., (1995). Interna metrijska svojstva skala Gilford-Zimmermanovog pregleda temperamenta [Internal metric characteristics of 'Guilford-Zimmerman temperament survey' scales]. *Psihologija*, 3-4, 363–374.
- Kaiser, H. F., & Rice J. (1974). Little jiffy, mark iv. *Educational and Psychological Measurement*, 34(1), 111–117. doi.org/10.1177/001316447403400115

Reliability measures, classic model of measurementL-(LAMBDA 1) – Guttman coefficient of reliability⁶Alfa – Spearman-Brown-Kuder-Richardson-Guttman-Cronbach coefficient of reliability⁷L6 – (LAMBDA 6) – Guttman coefficient of reliability⁴**Reliability measures of first main component**MI1 – Momirović-Knežević lower limit of reliability⁸Beta – Lord-Kaiser-Caffrey reliability of summary result^{9,10}Beta 6 – Momirović lower limit of reliability¹¹**Reliability measures, Guttman model of measurement**Gama – Momirovic, Knezevic lower reliability limit^{6,12}RHO – Guttman-Nicewander coefficient¹⁰**Measures of the scale homogeneity¹³**

H1 – average variable correlation

H2 – Momirović homogeneity measure

H5 – Knežević-Momirović homogeneity measure

-
- 6 Guttman, L. (1945). A basis for analyzing test-retest reliability, *Psychometrika*, 10(4), 255–282. doi.org/10.1007/BF02288892
- 7 Feldt, L. S. (1969). A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests, *Psychometrika*, 34(3), 363–373. dx.doi.org/10.1007/BF02289364
- 8 Momirović, K., Knežević, G., & Radović, B. (1995). Some relations between Hotelling's factors and Guttman's first lower bound to reliability. *Psihologija*, 28(Special Issue), 111–120.
- 9 Lord, F. M. (1953): The relation of test score to the trait underlying the test. *Educational and Psychological Measurements*, 15, 513–548.
- 10 Kaiser, H. F., & Caffrey, J. (1965): Alpha factor analysis. *Psychometrika*, 30, 1–14. dx.doi.org/10.1007/BF02289743
- 11 Momirović, K. (1975). Određivanje donje granice pouzdanosti kompozitnih testova [Determination of lower reliability limit of composite tests], *Materijali Petog kongresa psihologa SFRJ*, (str. 258–261), Skopje: Društvo psihologa SR Makedonije.
- 12 Nicewander, W. A. (1990) A latent-trait based reliability estimate and upperbound, *Psychometrika* 55(1), 65–74.
- 13 Knežević, G., & Momirović, K. (1996). RTT9G i RTT10G – dva nova programa za izračunavanje metrijskih karakteristika kompozitnih mernih instrumenata [RTT9G and RTT10G – two new programs for the computation of metric characteristics of composite measuring instruments], u: P. Kostić (Ur.), *Merenje u psihologiji*, 2 (str. 37–56). Beograd: Institut za kriminološka i sociološka istraživanja i Centar za primenjenu psihologiju – Društvo psihologa Srbije.