Psicothema

Psicothema 2016, Vol. 28, No. 4, 465-470
doi: 10.7334/psicothema2016.113

ISSN 0214 - 9915 CODEN PSOTEG

www.psicothema.com

How response bias affects the factorial structure of personality self-reports

David Navarro-González, Urbano Lorenzo-Seva and Andreu Vigil-Colet
Universitat Rovira i Virgili

Abstract

Background: Various studies have shown that acquiescence can distort the factor structure of personality questionnaires based on the five-factor model. In the present study, we analysed how acquiescence and social desirability affect the factor structure of a measure based on this personality model and a measure of aggression. **Method:** We analysed the factor structures of both tests before and after removing both biases in a sample of 532 adolescents aged between 11 and 18 (M= 14.75, SD= 2.1). **Results:** The factor structure of both tests presented a worse fit to the expected model when response bias was not controlled, and the congruence indexes for the personality and aggression measures showed a moderate (from C= .948 to C= .872) or great (from C= .931 to C= .475) decrease, respectively. Furthermore, acquiescence was largely responsible for these effects, and social desirability effects were only shown on the aggression measure. **Conclusions:** Response bias, and especially acquiescence, should be controlled during the development of personality measures to avoid distorting them, especially with samples of people with a high level of acquiescence (for example, those with little education, the young or the elderly). Furthermore, the use of response bias loadings as a criterion for choosing the items minimizes those distortions.

Keywords: Response bias, personality, factor structure.

Resumen

Efectos de los sesgos de respuesta en la estructura factorial de los autoinformes de personalidad. **Antecedentes:** diversos estudios han mostrado que la aquiescencia genera distorsiones en la estructura factorial de los cuestionarios de personalidad. En este estudio analizamos los efectos tanto de la aquiescencia como de la deseabilidad social en la estructura factorial de dos cuestionarios. **Método:** se analizó la estructura factorial de ambos con y sin sesgos de respuesta en una muestra de 532 adolescentes con edades entre los 11 y los 18 años (M= 14.75 SD= 2.1). **Resultados:** cuando no se eliminó el efecto de los sesgos de respuesta, el ajuste de ambos tests en relación al modelo esperado empeoró, disminuyendo la congruencia factorial moderadamente (desde C= .948 hasta C= .872) o notablemente (desde C= .931 a C= .475) para las medidas de personalidad y agresividad, respectivamente. Además, la aquiescencia fue la principal responsable de estos efectos, mientras que la deseabilidad social tan solo afectó la medida de agresividad. **Conclusiones:** es necesario controlar los sesgos de respuesta para evitar estructuras factoriales distorsionadas, especialmente en muestras con elevados niveles de aquiescencia, como poblaciones con bajo nivel educativo, adolescentes o en la tercera edad. Además, la minimización de los sesgos de respuesta durante el proceso de elección de ítems parece reducir dichas distorsiones.

Palabras clave: sesgos de respuesta, personalidad, estructura factorial.

The impact of response bias on typical response measures is an issue that has generated a great deal of research in recent decades. Although most of the research has focused on questionnaires within the framework of the five-factor model (FFM) of personality (i.e., Holden & Passey, 2010; Ones, Viswesvaran, & Reiss, 1996), the impact of response biases has also been assessed in other typical response measures such as impulsivity (Vigil-Colet, Ruiz-Pamies, Anguiano-Carrasco, & Lorenzo-Seva, 2012), aggression and violence (Becker, 2007; Bell & Naugle, 2007), psychological maturity (Morales-Vives, Camps, & Lorenzo-Seva, 2013), mood states (Soubelet & Salthouse, 2011) or well-being (Kozma & Stones, 2012).

Received: April 12, 2016 • Accepted: July 28, 2016
Corresponding author: Andreu Vigil-Colet
Research Center for Behavior Assessment
Universitat Rovira i Virgili
43007 Tarragona (Spain)
e-mail: andreu.vigil@urv.cat

A glance at the scientific literature on the issue reveals that the two most important response biases are social desirability (SD), defined as the tendency for people to present themselves in a generally favourable fashion (Holden, 2010), and acquiescence (AC), defined as the tendency of respondents to agree with statements without regard to their content (Paulhus & Vazire, 2005).

Most of the research in this field has focused on how response bias affects the validity of self-reports. For instance, a great deal of research has analysed the effects of SD on test scores in employment selection processes (Ones, Dilchert, Viswesvaran, & Judge, 2007; Ones et al., 1996; Salgado, 2005). Other issues, however, such as the effects of response bias on the factor structure of questionnaires, especially in the case of SD, have received less attention.

Response bias can affect the factor structure of questionnaires because it distorts the inter-item correlation matrix pattern (Bentler, Jackson, & Messick, 1971; Rammstedt, Goldberg, & Borg, 2010; Soto, John, Gosling, & Potter, 2008). For instance, in the presence

of AC, items worded in the same direction tend to show a positive relationship that is not due to the content measured, while items worded in different directions will tend to show a negative relationship. In this case, AC will lead to the overestimation or underestimation of correlations in terms of the items' direction. A similar effect may be expected in the case of SD: that is, items most affected by SD will show a positive correlation independently of their content. As a consequence, the distortion of the inter-item correlation matrix may have a considerable impact on the resulting factor structure.

Some studies have shown these effects for AC. In this respect, tests administered to samples of low educational levels, low intelligence or adolescents and pre-adolescents have the worst fits to the five factor model of personality (Meisenberg & Williams, 2008; Rammstedt et al., 2010; Rammstedt & Kemper, 2011; Soto et al., 2008). Therefore, the validity of personality measures may be affected in subpopulations with high levels of AC such as those described above or others who also have high levels of AC, such as the elderly (Ross & Mirowsky, 1984; Vigil-Colet, Lorenzo-Seva, & Morales-Vives, 2015).

These studies have analysed the effects of AC because such methods as ipsatizing allow the effects of AC to be removed from the inter-item correlation matrix (Ten Berge, 1999). Nevertheless, there are fewer methods available for removing the distortions due to SD.

Ferrando, Lorenzo-Seva, & Chico (2009) developed a general method for controlling both biases simultaneously. The first step in the method identifies a factor related to SD by using items that are taken as markers of SD.

Then the loadings of the content items on this SD factor are used to compute a residual inter-item correlation matrix free of SD. Subsequently the residual correlation matrix is analysed by applying the method developed by Lorenzo-Seva & Ferrando (2009), which removes from the content those items of the variance that are due to acquiescent responding. This process makes it possible to analyse a residual inter-item correlation matrix that is free of the distortions caused by SD and AC, and can be used in classical exploratory factor analysis (EFA) to determine the factor structure of the questionnaire.

As can be seen the main objective of the method is to remove the effects of both biases and compare the factor structures obtained with and without bias. Nevertheless, an EFA can be run on each residual matrix and the factor structures can be compared without controlling for bias, controlling only for AC, controlling only for SD and controlling for both biases. Therefore, the impact of the distortion caused by AC and SD on the factor structure of questionnaires can be assessed.

This is the main objective of the present research, which focuses on the conjoint and individual effects of response bias on the factor structure. Although previous research has shown the effects of AC, less is known about the effects of SD and the conjoint effect of both biases. This is relevant because their effects seem to depend upon the nature of the item. In this regard, the effects of AC seem to be weak and the effects of SD stronger when the items reflect highly desirable or undesirable behaviours. In neutral items, however, the effects of AC are stronger (Ferrando & Anguiano-Carrasco, 2010). From this viewpoint, the effects of both kinds of response bias on the structure of questionnaires may depend upon the level of social acceptation or rejection of the content measured. Taking this into account, we will analyse the impact of biases on a personality questionnaire based on the FFM and an aggression questionnaire because these kinds of measure are highly affected by SD (Becker, 2007; Morren & Meesters, 2002; Vigil-Colet et al., 2012). Furthermore, the tests were administered to adolescents, who usually show higher levels of AC than adults (Soto et al., 2008).

Method

Participants

A total of 532 volunteer students (252 men and 280 women) from 4 different high schools from the Tarragona province with ages ranging from 11 to 18 years old (M=14.75 SD=2.1) participated. A total of 29.2% of the sample was aged between 11 and 13 years old, 52.5% between 14 and 16, and 18.3% between 17 and 18. Two high schools were in small cities and two in Tarragona.

Instruments

Overall Personality Assessment Scale –OPERAS– (Vigil-Colet, Morales-Vives, Camps, Tous, & Lorenzo-Seva, 2013). This is a 40-item instrument which gives scores for the factors: Extraversion (EX), Agreeableness (AG), Conscientiousness (CO), Emotional stability (ES), and Openness to experience (OE). Item responses are produced using a 5-point Likert scale. The test has suitable psychometric properties, with the following factorial consistencies: $r_{\theta\theta}$=.86, $r_{\theta\theta}$=.71, $r_{\theta\theta}$=.77, $r_{\theta\theta}$=.86, and $r_{\theta\theta}$=.81 for EX, AG, CO, ES and OE, respectively. This questionnaire contains four items of SD, the aim of which is to control this response bias. In addition, some of the items are content balanced so that the acquiescent responding bias can be controlled.

The indirect-direct aggression questionnaire –IDAQ– (Ruiz-Pamies, Lorenzo-Seva, Morales-Vives, Cosi, & Vigil-Colet, 2014). This test gives scores for the factors Physical aggression (PA), Verbal aggression (VA) and Indirect aggression (IA), as well as SD and AC scores for each individual. The factors measured by I-DAQ have appropriate factor reliabilities: $r_{\theta\theta}$=.83, $r_{\theta\theta}$=.77 and $r_{\theta\theta}$=.78 for PA, VA and IA, respectively.

Procedure

School approval and parental written informed consent were obtained before participation in the study. Children's participation was voluntary, and no incentives were given for their participation. About 96% of the children who were invited to participate in the study eventually did so.

A professional psychologist administered the tests collectively. The participants were asked to volunteer to answer the inventories in their classroom. The questionnaires were anonymous, and respondents had to provide only their gender and age.

Data analysis

We computed four EFA for each questionnaire, which took into account its three- (IDAQ) or five- (OPERAS) factor structure. These EFA were performed on the polychoric inter-item correlation matrix and on the residual matrixes obtained after SD, AC and both biases had been removed, using the method developed by Ferrando et al. (2009). To assess the fit of each loading matrix to the

expected factorial solutions, the congruence index developed by Tucker (1951) was computed between the rotated loading matrix and the ideal loading matrix. Data was analysed using MatLab (MathWorks, 2012).

Results

Table 1 shows the loading matrix for IDAQ with and without removing bias. As can be seen, when both SD and AC were removed, all the items except item 10 had their salient loading on the expected factor. On the other hand, when they were not removed, 9 of the 23 content items did not load on the expected factor. Congruence indexes showed clear improvement when response biases were removed. It should be taken into account that indexes higher than .85 imply a fair congruence between the rotated loading matrix and the ideal loading matrix, while indexes of .95 or higher imply that the rotated loading matrix and the ideal loading matrix are essentially equal (Lorenzo-Seva & Berge, 2006). When response biases were removed, all the congruence indexes were over .85, and most of them were around .95. But when they were not removed, most of the indexes showed a bad fit. It is worth mentioning that when biases were not removed, item number ten seems to load on the IA scale but, when they were, the loading seems to reflect the distortion caused by bias.

As can be seen, the questionnaire items showed that both biases had a considerable effect, the mean loading on SD (excluding the items used as markers) and AC being λ_m=.232 and λ_m=.222, respectively.

Table 2 shows the improvement when only one of the two response biases was removed. When SD was controlled, there was an improvement in the factor structure. Only three items did not show their salient loading on the expected factor, and the congruencies were clearly better than the ones obtained without any control, although not good enough. We found a greater improvement when AC was controlled. In this case, the results were quite similar to those obtained when both biases were removed, and the congruence indexes were similar.

Table 1
Loading matrix obtained with and without controlling response bias and factorial congruence with expected solution for I-DAQ. In bold face salient loadings on content factors

		Controlling bias					With bias		
	Item	SD	AC	PHY	IND	VER	PHY	IND	VER
PHYSICAL	13	.608							
	2	.647							
	8	.391							
	21	.747							
	1	-.209	.235	**-.496**	-.117	-.076	**.594**	.040	.180
	6	.294	.217	**.688**	-.090	-.153	**-.495**	.101	.128
	17	-.195	.314	**-.729**	-.081	-.052	**.794**	.156	.209
	19	-.213	.195	**-.785**	.077	.056	**.692**	.092	.013
	20	.324	.254	**.639**	.018	.026	**-.548**	.214	.091
	25	.272	.078	**.466**	.033	.287	**-.612**	.121	.012
INDIRECT	3	.386	.269	.036	**.463**	-.075	.101	**.691**	-.152
	4	.281	.283	.194	**.234**	-.071	-.066	**.453**	.018
	10	.212	.305	-.038	.020	-.033	.106	**.360**	.229
	11	.401	.343	.030	**.396**	.008	.038	**.645**	-.055
	14	-.001	.296	.142	**-.597**	-.076	-.047	-.145	**.573**
	16	-.119	.328	.057	**-.546**	-.076	.067	-.149	**.516**
	18	.114	.197	-.030	**.721**	-.022	.091	**.506**	-.333
	23	.330	.175	.078	**.323**	.020	-.081	**.402**	-.113
	24	.101	.310	-.027	**-.153**	.023	.073	.213	**.326**
	26	-.034	.325	-.085	**-.504**	.095	.081	-.025	**.593**
VERBAL	5	-.278	.117	.020	.010	**-.660**	**.410**	-.136	-.015
	7	.056	.198	.021	-.083	**.381**	**-.210**	.028	.073
	9	.459	.183	.010	.101	**.489**	-.320	**.374**	.051
	12	-.138	.135	-.128	-.045	**-.450**	**.419**	-.034	.069
	15	.331	.028	-.016	-.104	**.583**	**-.367**	.142	.101
	22	-.062	.197	-.047	-.080	**-.224**	**.219**	.037	.152
	27	.513	.105	.118	.035	**.278**	-.305	**.403**	.142
	λ_m	.232	.222						
	Congruence:			.953	.884	.955	.851	.534	.003
	Overall congruence:				.931			.475	

Note: SD: social desirability, AC: acquiescence PHY: physical aggression, VER: Verbal aggression, IND: Indirect aggression. $\lambda_{m:}$ mean of loadings

Table 2
Loading matrix controlling social desirability or acquiescence and factorial congruence with expected solution. In bold face salient loadings on content factors

		Controlling SD			Controlling AC		
	Item	PHY	IND	VER	PHY	IND	VER
PHYSICAL	1	**.528**	.178	-.028	**-.508**	-.126	-.146
	6	**-.602**	.201	-.023	**.747**	-.084	-.155
	17	**.774**	.187	.066	**-.750**	-.082	-.112
	19	**.782**	-.040	.093	**-.794**	.069	-.024
	20	**-.560**	.108	.185	**.654**	.016	.035
	25	**-.464**	-.017	.310	**.456**	.021	.314
INDIRECT	3	.066	**-.240**	.147	.035	**.524**	-.028
	4	-.091	-.036	**.146**	.199	**.256**	-.049
	10	.138	**.180**	.167	-.070	.035	.052
	11	.077	-.158	**.273**	.039	**.417**	.047
	14	-.076	**.674**	-.033	.158	**-.630**	-.053
	16	.005	**.619**	.015	.055	**-.577**	-.101
	18	.123	**-.529**	.203	-.046	**.697**	-.084
	23	-.026	**-.184**	.117	.078	**.347**	.083
	24	.100	**.296**	.146	-.056	**-.162**	.073
	26	.154	**.601**	.148	-.101	**-.514**	.098
VERBAL	5	.053	.057	**-.547**	.049	.021	**-.747**
	7	-.038	.076	**.367**	.018	-.123	**.312**
	9	-.019	-.039	**.585**	.036	.116	**.517**
	12	.183	.116	**-.363**	-.108	-.030	**-.470**
	15	-.050	.059	**.522**	-.043	-.112	**.695**
	22	.103	**.166**	-.152	-.045	-.080	**-.255**
	27	-.092	.068	**.386**	.128	.072	**.417**
Congruence		.952	.808	.837	.944	.884	.940
Overall Congruence			.864			.922	

Note: SD: social desirability, AC: acquiescence PHY: physical aggression, VER: Verbal aggression, IND: Indirect aggression

We performed the same analysis for OPERAS. Table 3 showed that, when biases were removed, all the items loaded on the expected dimension, with appropriate or excellent congruence indexes. When biases were not removed, the fit was worse, especially in the case of AG, and the congruence indexes decreased. Furthermore, five items did not load on the expected dimension. EX and ES were the least affected dimensions and, to reduce the size of the tables, their loadings are not included in the table. As can be seen, the impact of response biases was lower than for IDAQ, especially for SD, which showed a mean loading of λ_m=.177.

Table 4 shows the results when only one bias was controlled. When SD effects were removed, the improvement in the congruencies was almost negligible but, when only AC was controlled, we obtained nearly the same structure as the one reported in Table 3 when both biases were removed.

Discussion

Various authors have pointed out that AC is a source of bias in typical response measures which may distort their factor structure, but less is known about the effects of SD. The present study analysed the impact of both response biases on the factor structure of personality questionnaires and showed that both AC and SD have effects on factor structures, but of different magnitudes, and apparently related to the content measured.

The results discussed above seem to show that response bias has a considerable impact on the factor structure of questionnaires, and can cause a great deal of distortion. In fact, in the case of IDAQ, the structure obtained when bias is not controlled is absolutely incongruent with the expected structure, while in the case of OPERAS, the fit is worse. It is worth mentioning that during the development of OPERAS, the items were chosen by taking their loadings on AC and SD into account so as to minimize their impact on the test. It is logical, then, that controlling for bias had less effect on its structure than in the case of IDAQ. The results reported above seem to show the importance of controlling response bias when the factor structure is assessed.

Our results also show that the distortion due to AC is clearly bigger than the distortion due to SD. This result has been reported previously by several authors who assessed the effects of AC on the

Table 3

Loading matrix obtained with and without controlling response bias. In bold face salient loadings on content factors for OPERAS' scales conscientiousness, agreeableness and openness to experience

		Controlling bias							With bias				
	item	SD	AC	EX	ES	CO	AG	OP	EX	ES	CO	AG	OP
	5	-.310											
	11	.731											
	19	.722											
	26	.792											
CONSCIENTIOUSNESS	4	-.251	.228	.074	-.296	**-.327**	.174	.095	-.183	.101	-.300	**.429**	.042
	10	.311	.063	.064	.155	**.599**	-.170	-.120	.166	.044	**.463**	-.335	-.116
	16	.325	.059	-.080	.115	**.588**	-.065	-.151	.153	-.054	**.556**	-.238	-.157
	22	.271	.155	-.089	.016	**.432**	-.139	.135	.156	.019	**.441**	-.082	.079
	28	-.009	.188	-.010	.106	**-.342**	-.267	.198	.209	.099	**-.236**	.017	.230
	33	.379	.183	-.175	.172	**.526**	-.018	-.060	.262	-.081	**.628**	-.118	-.072
	38	-.128	.357	.098	-.144	**-.481**	-.061	.036	.027	.213	-.250	**.342**	.042
AGREEABLENESS	6	-.272	.319	.003	-.095	-.157	**.350**	.095	.004	-.022	-.195	**.513**	-.004
	12	-.371	.301	.006	-.110	-.220	**.512**	.120	-.042	-.060	-.275	**.637**	-.003
	17	-.207	.204	.168	-.188	.060	**.332**	-.017	-.131	.096	-.076	**.330**	-.091
	23	.390	.065	-.092	.036	.018	**-.468**	.097	.168	.103	.216	**-.303**	.157
	29	.486	.126	-.042	.173	.130	**-.390**	-.070	.291	.150	**.364**	-.351	.019
	34	-.276	.362	.043	-.207	-.120	**.426**	.248	-.103	.026	-.061	**.687**	.119
	39	.444	.159	.106	-.112	-.078	-.176	**-.223**	.051	**.311**	.273	-.099	-.159
OPENNESS	7	.026	.080	.105	-.101	.018	-.058	**-.640**	-.035	.124	-.011	-.100	**-.599**
	13	-.125	.296	-.015	-.128	-.179	-.107	**.462**	.043	.080	-.141	.282	**.382**
	18	.041	.077	.046	.064	-.037	.102	**-.240**	.065	.041	-.029	.003	**-.207**
	24	.013	.061	-.151	.088	-.041	.045	**.694**	.087	-.127	.001	.140	**.686**
	30	-.114	.263	.208	-.056	-.302	.282	**.373**	.021	.234	-.147	**.508**	.280
	35	-.070	.451	.198	-.167	-.109	.140	**.510**	.033	.287	.084	**.569**	.362
	40	.109	.161	-.026	-.070	.047	.016	**-.761**	.026	.039	.148	-.042	**-.689**
	λ_m	.177	.206										
	Congruence			.968	.965	.959	.902	.942	.930	.892	.864	.760	.914
	Overall Congruence					.948					.872		

Note: SD: social desirability AC: Acquiescence EX: extraversion; AG: agreeableness; CO: conscientiousness; ES: emotional stability; OE: openness to experience. λ_m: mean of loadings

structure of personality measures based on the FFM (Meisenberg & Williams, 2008; Rammstedt et al., 2010; Soto et al., 2008) but there have been no studies about its effects on other personality dimensions. In this regard, our results show that AC has a similar effect on aggression measures.

Two important questions are raised by the results reported above. The first is that AC has a greater effect on factor structures than SD does, and the second is the different effect that controlling SD has on the two tests analysed.

To answer the first question, we hypothesise that the effects of SD and AC on the inter-item correlation matrix are slightly different. When a pair of items is affected by SD, their correlation may increase in two circumstances. If they are measuring the same content, SD overestimates their relationship and, if their contents are not related, SD may generate a correlation where none is expected. It is noted that both distortions always give rise to a positive relationship. The effect of AC is more complex because, if the items share the same content, AC may over- or underestimate their relationship depending on whether the items are keyed in the same or in opposite directions. On the other hand, if the contents are not related, AC generates positive or negative relationships between items where no relationship is expected. Therefore, AC affects not only the magnitude of the relationships but also their sign; it involves a greater distortion of the inter-item correlation matrix and has a deeper impact on the factor structure.

The second question refers to the differences observed after controlling SD in both tests. In the case of OPERAS, removing SD effects led to negligible improvement on the congruence of the factor solution (from C=.872 to C=.874) while in the case of IDAQ, the improvement was considerable (from C=.475 to C=.864). These differences may be because aggression measures are usually highly affected by SD, because this kind of behaviour is socially rejected (Vigil-Colet et al., 2012). In fact, the mean loadings on SD for IDAQ (λ_m=.232) were greater than for OPERAS (λ_m=.177). Therefore, it is possible that SD distorts aggression measures more than overall personality measures, which may explain the differences observed.

To summarize, the main conclusion of the present study is that response bias, and especially AC, should be controlled when assessing the dimensionality of a personality measure because, if it is not, the researcher may find a distorted dimensionality or a dimensionality that is not the one expected from the theoretical model underlying the measure. This point is even more important when the sample comprises a group of people with a high level of acquiescence such as people with a low level of education, the young or the elderly (Meisenberg & Williams, 2008; Soto et al., 2008; Vigil-Colet, Morales-Vives, & Lorenzo-Seva, 2013). Taking into account that even low levels of AC have effects on item covariation (Rammstedt & Farmer, 2013), further research should apply the method used in the present study in samples with moderate or low levels of AC to assess the degree of distortion of the factor structures in those populations.

Table 4
Loading matrix controlling social desirability or acquiescence. In bold face salient loadings on content factors for conscientiousness, agreeableness and openness to experience

		Controlling SD					Controlling AC				
	item	EX	ES	CO	AG	OP	EX	ES	CO	AG	OP
CONSCIENTIOUSNESS	4	.083	-.194	-.278	**.393**	-.013	-.279	.040	**-.373**	.310	.092
	10	.092	.186	**.543**	-.259	.063	.170	.093	**.604**	-.221	-.138
	16	-.067	.151	**.569**	-.149	.138	.113	-.040	**.627**	-.194	-.153
	22	-.039	.147	**.464**	.055	-.092	.016	-.058	**.433**	-.236	.121
	28	.060	.217	**-.345**	.034	-.211	.123	-.020	**-.400**	-.203	.186
	33	-.161	.260	**.545**	.011	.090	.143	-.130	**.601**	-.239	-.043
	38	.158	.051	**-.422**	.357	.046	-.140	.072	**-.543**	.034	.034
AGREEABLENESS	6	-.026	-.001	-.099	**.489**	.019	-.079	-.037	-.195	**.490**	.089
	12	-.056	-.050	-.147	**.596**	.020	-.098	-.046	-.257	**.658**	.119
	17	.134	-.133	.090	**.351**	.103	-.173	.128	.003	**.413**	-.021
	23	.012	.154	.002	**-.199**	-.145	.034	-.042	.057	**-.598**	.089
	29	.037	.282	.114	-.206	.029	.153	.019	.211	**-.598**	-.062
	34	.011	-.093	-.034	**.631**	-.100	-.208	.003	-.158	**.497**	.259
	39	.160	.025	-.036	.062	**.264**	-.147	.164	.023	**-.441**	-.201
OPENNESS	7	.126	-.039	.002	-.108	**.635**	-.090	.107	.007	-.019	**-.632**
	13	.054	.045	-.137	.315	**-.402**	-.103	-.041	-.246	.052	**.442**
	18	.026	.062	-.049	.016	**.242**	.064	.054	-.015	.063	**-.245**
	24	-.150	.081	-.037	.172	**-.689**	.091	-.148	-.029	.022	**.684**
	30	.179	.023	-.217	**.536**	-.247	-.072	.188	-.294	.274	**.393**
	35	.228	.057	.015	**.660**	-.338	-.187	.176	-.123	.186	**.544**
	40	-.014	.020	.056	-.032	**.785**	-.087	-.016	.075	-.052	**-.756**
Congruence		.965	.890	.945	.658	.909	.936	.853	.856	.858	.846
Overall Cong.				.874					.944		

Note: EX: extraversion; AG: agreeableness; CO: conscientiousness; ES: emotional stability; OE: openness to experience

One possible limitation of the study is that the factorial invariance across the age groups was not tested. Although the age range is small, the quick development of personality during adolescence suggests that further research with larger samples should analyse the factorial invariance across age groups for both tests.

Acknowledgments

The research was partially supported by a grant from the Catalan Ministry of Universities, Research and the Information Society (2014SGR73) and by a grant from the Spanish Ministry of Economy and Competitivity (PSI2014-52884-P).

References

Becker, G. (2007). The Buss-Perry aggression questionnaire: Some unfinished business. *Journal of Research in Personality*, *41*, 434-452.

Bell, K. M., & Naugle, A. E. (2007). Students' self-Reporting of partner abuse perpetration and victimization. *Violence and Victims*, *22*, 243-257.

Bentler, P., Jackson, D., & Messick, S. (1971). Identification of content and style: A two-dimensional interpretation of acquiescence. *Psychological Bulletin*, *76*, 186-204.

Ferrando, P. J., & Anguiano-Carrasco, C. (2010). Acquiescence and social desirability as item response determinants: An IRT-based study with the Marlowe-Crowne and the EPQ Lie scales. *Personality and Individual Differences*, *48*, 596-600.

Ferrando, P. J., Lorenzo-Seva, U., & Chico, E. (2009). A General factor-Analytic procedure for assessing response bias in questionnaire measures. *Structural Equation Modeling, 16*, 364-381.

Holden, R. (2010). Social desirability. In *Corsini Encyclopedia of Psychology*. London: John Wiley & Sons, Inc.

Holden, R. R., & Passey, J. (2010). Socially desirable responding in personality assessment: Not necessarily faking and not necessarily substance. *Personality and Individual Differences*, *49*, 446-450.

Kozma, A., Stones, M. J. (1988). Social desirability in measures of subjective well-being: Age comparisons. *Social Indicators Research*, *20*, 1-14.

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

Lorenzo-Seva, U., & Ferrando, P. J. (2009). Acquiescent responding in partially balanced multidimensional scales. *The British Journal of Mathematical and Statistical Psychology*, *62*, 319-326.

MathWorks, M. (2012). *The language of technical computing*. Mathlab Incorporated. Natick, Massachusetts, USA.

Meisenberg, G., & Williams, A. (2008). Are acquiescent and extreme response styles related to low intelligence and education? *Personality and Individual Differences*, *44*, 1539-1550.

Morales-Vives, F., Camps, E., & Lorenzo-Seva, U. (2013). Development and validation of the psychological maturity assessment scale (PSYMAS). *European Journal of Psychological Assessment*, *29*, 12-18.

Morren, M., & Meesters, C. (2002). Validation of the Dutch version of the aggression questionnaire in adolescent male offenders. *Aggressive Behavior*, *28*, 87-96.

Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. (2007). In support of personality assessment in organizational settings. *Personnel Psychology*, *60*, 995-1027.

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.

Paulhus, D. L., & Vazire, S. (2005). The Self-Report Method. In R. W. Robins, R. Fraley & R. F. Krueger (Eds.). *Handbook of research methods in personality psychology* (pp. 224-239). New York: Guilford Press.

Rammstedt, B., & Farmer, R. F. (2013). The impact of acquiescence on the evaluation of personality structure. *Psychological Assessment*, *25*, 1137-1145.

Rammstedt, B., Goldberg, L. R., & Borg, I. (2010). The measurement equivalence of Big-Five factor markers for persons with different levels of education. *Journal of Research in Personality*, *44*, 53-61.

Rammstedt, B, & Kemper, C. J. (2011). Measurement equivalence of the Big Five: Shedding further light on potential causes of the educational bias. *Journal of Research in Personality*, *45*, 121-125.

Ross, C. E., & Mirowsky, J. (1984). Socially-desirable response and acquiescence in a cross-cultural survey of mental health, *25*, 189-197.

Ruiz-Pamies, M., Lorenzo-Seva, U., Morales-Vives, F., Cosi, S., & Vigil-Colet, A. (2014). I-DAQ: A new test to assess direct and indirect aggression free of response bias. *The Spanish Journal of Psychology*, *17*, E41.

Salgado, J. F. (2005). Personalidad y deseabilidad social en contextos organizacionales: implicaciones para la práctica de la psicología del trabajo y las organizaciones [Personality and social desirability in organizational contexts: Implications for practice of psychology of work and organizations]. *Papeles del Psicólogo*, *26*, 115-128.

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of big five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, *94*, 718-737.

Soubelet, A., & Salthouse, T. (2011). Influence of social desirability on age differences in self-reports of mood and personality. *Journal of Personality*, *79*, 741-762.

Ten Berge, J. M. (1999). A legitimate case of component analysis of ipsative measures, and partialling the mean as an alternative to ipsatization. *Multivariate Behavioral Research*, *34*, 89-102.

Tucker, L. (1951). A method for synthesis of factor analysis studies (Personnel Research Section Report No. 984). *Washington D.C. Department of the Army.*

Vigil-Colet, A., Lorenzo-Seva, U., & Morales-Vives, F. (2015). The effects of ageing on self-reported aggression measures are partly explained by response bias. *Psicothema*, *27*, 209-215.

Vigil-Colet, A., Morales-Vives, F., Camps, E., Tous, J., & Lorenzo-Seva, U. (2013). Development and validation of the Overall Personality Assessment Scale (OPERAS). *Psicothema*, *25*, 100-106.

Vigil-Colet, A., Morales-Vives, F., & Lorenzo-Seva, U. (2013). How social desirability and acquiescence affect the age-personality relationship. *Psicothema*, *25*, 342-348.

Vigil-Colet, A., Ruiz-Pamies, M., Anguiano-Carrasco, C., & Lorenzo-Seva, U. (2012). The impact of social desirability on psychometric measures of aggression. *Psicothema*, *24*, 310-315.