Elaboración de la versión breve del Cuestionario de Personalidad IPIP-Revisado: Control del sesgo de aquiescencia

MARCOS CUPANI*
Universidad Nacional de Córdoba, Córdoba, Argentina
ORCID: https://orcid.org/0000-0003-2132-5552
URBANO LORENZO-SEVA
Universitat Rovira i Virgili, Tarragona, España
ORCID: https://orcid.org/0000-0001-5369-3099
CELINA GRACIELA KORZENIOWSKI
Universidad del Aconcagua, Mendoza, Argentina
ORCID: https://orcid.org/0000-0001-6678-3472
ANA ESTEFANÍA AZPILICUETA
Universidad Nacional de Córdoba, Córdoba, Argentina
ORCID: https://orcid.org/0000-0002-6221-3162

How to quote this article: Cupani, M., Lorenzo-Seva, U., Korzeniowski, C.G., Azpilicueta, A.E. (2019). Elaboración de la Versión Breve del Cuestionario de Personalidad IPIP- Revisado: Control del sesgo de respuesta aquiescencia. *Acta Colombiana de Psicología,* 22(1), 261-272. doi: http://www.dx.doi.org/10.14718/ACP.2019.22.1.12

Resumen

En los últimos años se han desarrollado medidas breves de los cinco factores de personalidad, sin embargo, la ganancia práctica de tiempo provista por las formas breves puede implicar propiedades psicométricas más débiles de los instrumentos. En la construcción de escalas breves, para mantener propiedades psicométricas adecuadas se debe emplear criterios teóricos y empíricos en la selección de los ítems y minimizar los sesgos de respuesta, como el de la aquiescencia (AC), que hace referencia a la tendencia de las personas a estar de acuerdo con afirmaciones positivas independientemente del contenido de la afirmación. Teniendo esto en cuenta, el objetivo principal del presente estudio fue desarrollar un instrumento breve (30 ítems), de dominio público, para medir los cinco factores de personalidad en población latina, controlando el sesgo de respuesta AC. La muestra estuvo compuesta por 910 participantes, 543 de sexo femenino (59.6 %) y 367 de sexo masculino (40.3 %), con edades comprendidas entre los 15 y los 80 años (M = 29.52; DT = 12.25), pertenecientes a la ciudad de Córdoba, Argentina. Para el proceso de validación se propuso realizar un estudio de convergencia con las cinco escalas del NEO-FFI, un análisis de diferencia de grupos según el sexo y la edad de los participantes, y un estudio de validez predictiva respecto a algunas actividades recreacionales (uso de drogas, irresponsabilidad, amistad, erudición/creatividad y comunicación). Los resultados indican que el IPIP-R-30 presenta una estructura factorial de cinco factores, índices de confiabilidad adecuados tanto de consistencia interna como de estabilidad temporal, evidencia de validez convergente con las escalas del NEO-FFI, evidencia de diferencia de grupos según sexo y edad, y validez predictiva de la frecuencia de diferentes categorías de actividades específicas. De esta manera, se puede concluir que el IPIP-R-30 constituye una herramienta válida de evaluación de los rasgos de personalidad en nuestro medio, con puntuaciones libres del sesgo de AC.
Palabras clave: cinco grandes factores de la personalidad, inventario, IPIP, aquiescencia, sesgo de respuesta.

Development of a Brief Version of the Personality Inventory IPIP-Revised: Control of the Acquiescence Response Bias

Abstract

In recent years, brief measures of the five personality factors have been developed; however, the practical gain of time provided by the brief versions may imply weaker psychometric properties of the instruments. To maintain adequate psychometric properties in the construction of brief scales, theoretical and empirical criteria should be used in the selection of items, and response biases such as acquiescence (AC) should be minimized. The term AC refers to people's tendency to agree with positive statements, regardless of their content. The main purpose of the present study is to develop a brief public

* Instituto de Investigaciones Psicológicas (IIPsi-CONICET), Facultad de Psicología, Universidad Nacional de Córdoba, Ciudad Universitaria, Córdoba 5000, Argentina. Tel./Fax: +54-351-4334104. marcoscup@gmail.com

Recibido, julio 4/2017; Concepto de evaluación, noviembre 17/2017; Aceptado, enero 31/2018

domain instrument (30 items) to measure the five personality factors in the Latin American population, controlling the AC response bias. The sample consisted of 910 participants, 543 women (59.6 %) and 367 men (40.3 %), age range 15-80 years (M = 29.52; DT = 12.25) from the city of Córdoba, Argentina. For the validation process, a convergence study with the five NEO-FFI scales, an analysis of group differences according to the participants' sex and age, and a predictive validity study regarding recreational activities (drug use, irresponsibility, friendship, erudition/creativity, and communication) were performed. The results indicate that the IPIP-R-30 presents a five-factor factorial structure, adequate reliability indices of both internal consistency and temporal stability, evidence of convergent validity with the NEO-FFI scales, evidence of group differences as regards sex and age, and frequency predictive validity of different categories of specific activities. Thus, it can be concluded that the IPP-R-30 is a valid tool for assessing personality factors in our environment, with scores free of AC bias.
Keywords: big five personality factors, inventory, IPIP, acquiescence, response bias.

Elaboração da versão breve do Inventário de Personalidade IPIP-Revisado: controle do viés de aquiescência

Resumo

Nos últimos anos, têm sido desenvolvidas medidas breves dos cinco fatores de personalidade; contudo, o ganho prático de tempo previsto pelas formas breves pode implicar propriedades psicométricas mais fracas dos instrumentos. Na construção de escalas breves, para manter propriedades psicométricas adequadas, devem ser empregados critérios teóricos e empíricos na seleção dos itens, e devem ser minimizados os vieses de resposta, como o da aquiescência (AC). Esse conceito faz referência à tendência das pessoas que estão de acordo com afirmações positivas, independentemente do conteúdo da afirmação. Nesse sentido, o objetivo principal deste estudo foi desenvolver um instrumento breve (30 itens), de domínio público, para medir os cinco fatores de personalidade em população latina, controlando o viés de resposta AC. A amostra esteve composta por 910 participantes, 543 de sexo feminino (59.6 %) e 367 de sexo masculino (40.3 %), entre 15 e 80 anos de idade (M = 29.52; DP = 12.25), pertencentes à cidade de Córdoba, Argentina. Para o processo de validação, foi proposto realizar um estudo de convergência com as cinco escalas do NEO-FFI, uma análise de diferença de grupos segundo o sexo e idade dos participantes, e um estudo de validade preditiva a respeito de atividades recreativas (uso de drogas, irresponsabilidade, amizade, erudição/criatividade e comunicação). Os resultados indicam que o IPIP-R-30 apresenta uma estrutura fatorial de cinco fatores, índices de confiabilidade adequados tanto de consistência interna quanto de estabilidade temporal, evidência de validade convergente com as escalas do NEO-FFI, evidência de diferença de grupos segundo sexo e idade, e validade preditiva da frequência de diferentes categorias de atividades específicas. Dessa maneira, pode-se concluir que o IPIP-R-30 constitui uma ferramenta válida de avaliação dos traços de personalidade em nosso meio, com pontuações livros do viés de AC.
Palavras-chave: cinco grandes fatores da personalidade, inventário, IPIP, aquiescência, viés de resposta.

Introduction

Brief measures of the five personality factors have been developed along the past few years. The most recognized ones are the Ten-Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003), the Mini-IPIP (Donnellan, Oswald, Baird, & Lucas, 2006), the BFI-2-S and BFI-2-XS (Soto & John, 2017); the Abridged Big Five (Langford, 2003), the Five-Item Measure of the Big Five (Aronson, Reilly & Lynn, 2006), and the Big Five Inventory-10 questionnaire (BFI-10; Rammstedt & John, 2007). These scales have arisen with the aim of solving one of the practical disadvantages of the traditional personality questionnaires, that is, the time required to complete them (Sibley, 2012).

However, the practical gain of time provided by the brief versions may imply, in many cases, weaker psychometric properties of the instruments (Credé, Harms, Niehorster, & Gaye-Valentine, 2012). In comparison with a 60-item scale, a 30-item scale registers a loss in both reliability and validity estimated in 10%, whereas in a 15-item scale the loss rises to 20% (Soto & John, 2017). The internal consistency of the questionnaires is often the most affected item, although temporal stability and inter-examiner reliability also decreases (Baldasaro, Shanahan & Bauer, 2013; Sibley, 2012). The convergent, discriminant, and predictive validities decrease as well when compared with the extended versions (Baldaraso et al., 2013). As regards the construct validity, there is no conclusive evidence that the brief inventories fit the five-factor factorial structure. Some investigations report a fit from adequate to good (Baldasaro et al., 2013; Donellan et al., 2006; Soto & John, 2017), whereas some others report a lack of fit (Gosling et al., 2003).

In agreement with this, it has been documented that the use of very brief versions (one or two items per domain)

is associated with the increase of Type 1 and Type 2 errors (Credé et al., 2012; Milojev, Osborne, Greaves, Barlow, & Sibley 2013; Kruyen, Emons, & Sijtsma, 2013). For example, the TIPI provides scores with inadequate reliability values, and it has been criticized for its lack of amplitude in the construct evaluation (Baldaraso et al., 2013). With four items per domain, as in the case of the Mini IPIP (Donellan et al., 2011; Sibley, 2012), the internal consistency improves, although it remains relatively low; in contrast, the construct and convergent validities seem not to be significantly affected (De Vries, 2013). Researchers are mostly concerned about the internal consistency loss, because they have to decide between consistency and amplitude in the construct coverage (Baldaraso et al., 2013). In case of prioritizing the content, the result is a heterogeneous but unreliable scale. In the opposite way, when prioritizing consistency. there might be tautological problems due to the inclusion of strongly correlated items with the domain to be evaluated (Soto & John, 2017).

In sum, in the construction of brief scales there should be a balance between the eagerness to minimize the number of necessary items to evaluate each personality domain and the aim to maintain the coverage content with good psychometric properties (Milojev et al., 2013). One way of achieving this goal is to apply empirical and theoretical criteria to select the items and minimize the response bias. As regards the response bias, the concept of acquiescence (AC) has lately interested the specialists on the topic. This concept refers to people's tendency to agree with positive statements regardless of their content (Cronbach, 1942). Some research has evidenced that there are differences of sex, age, intelligence level, education level, and cultural variations in this response bias (Javeline, 1999; Johnson, Kulesa, Cho, & Shavitt, 2005; Vazsonyi, Ksinan, Mikuška & Jiskrova, 2015) and that the lack of control of such differences can distort the scale factorial structure and their associations with external criteria. In fact, some authors consider the AC an undesirable variation source that should be removed (e.g., Hofstee, ten Berge, & Hendriks, 1998), whereas others suggest that when the AC bias is controlled, the five-factor factorial structure fits correctly (Rammstedt, Kemper, & Borg, 2013). According to this statement, it is essential to control this bias when producing brief scales. To perform this kind of studies, it is vital to design a method at the time the inventory is being developed (Cupani & Lorenzo-Seva, 2016), such as the procedure proposed by Lorenzo-Seva and Ferrando (2009) for partially balanced scales.

As regards the evaluation of the five personality factors, in the last decades there has been an increase in the number of publications on the application, adaptation, and validation of the IPIP scales (Goldberg, 2001). The IPIP resulted from a worldwide collaborative effort to develop a set of public domain personality items that, accordingly, may be freely used by researchers from any place in the world for both scientific and commercial purposes. Currently, there is a resulting pool of 2000 IPIP items to be used (Goldberg, et al., 2005). The IPIP items have been translated into more than 10 languages, and there has been a sustained increase in the adaptation and construction of IPIP scales in different cultures and countries (Cupani & Lorenzo-Seva, 2016). For example, the 50-item and the 100-item IPIP scales have been validated in the USA (Goldberg et al., 2005), Scotland (Gow, Whiteman, Pattie, & Deary, 2005), New Zealand (Guenole & Chernyshenko, 2005), Croatia (Mlačić & Goldberg, 2007) and China (Zheng et al., 2008). In Latin America, there are a few works on the topic (Cupani, 2009; Gross, Zalazar Jaime, Piccolo & Cupani, 2012). In our context, the IPIP Personality Questionnaire – Revised (Cupani & Lorenzo-Seva, 2016), with its 50 items selected out of the 100 items proposed by Goldberg (1999) controlling the acquiescence effect, presents an orthogonal simple factorial structure and adequate psychometric properties (Cupani & Lorenzo-Seva, 2016).

Considering that the number of studies that have weighed the effect of acquiescence in the construction of brief personality scales is scarce, the present work aims to develop a brief public domain instrument (30 items) to measure the five factors in the Latin American population controlling the bias response and with representative items for the sociocultural context. To develop this instrument, theoretical and empirical criteria were used to select the items to obtain a more precise evaluation tool, free of the AC response bias and with a more stable factorial structure as regards the group differences (age and sex). For the construction of the inventory, the starting point was the set of 100 items proposed by Goldberg (1999) in the IPIP. When selecting the items, it was considered that each of them was related to a psychological aspect of each of the five personality dimensions evaluated by the test. The selection of six indicators per factor was proposed with the aim of both minimizing the Type1 and Type 2 errors associated with the brief scales and balancing the reduction of items with the content coverage and the good psychometric properties (Baldasaro et al., 2013). For the validation process, it was proposed to perform a convergence study with the five NEO-FFI scales (Costa & McCrae, 1992), which is considered a landmark in the five-factor model (FFM) evaluation (Chamorro-Premuzic & Furnham, 2009), as well as an analysis of the group differences regarding the participants' sex and age, and a study of predictive validity with respect to recreational activities, such as drug use, irresponsibility, friendship, erudition/creativity, and communication (Grucza & Goldberg, 2007), through the Frequency Inventory of Daily

Living Activities (Goldberg, 1999), which is also part of the IPIP. It was expected to find group differences regarding the participants' sex, such as higher levels of Neuroticism and Agreeableness in women (McCrae & Terracciano, 2005; Cupani & Lorenzo-Seva, 2016), and age, such as a decrease in the Extraversion, Neuroticism, and Openness factors and an increase of the Agreeableness and Conscientiousness factors as time passed by (Srivastava, John, Gosling, & Potter, 2003; Cupani & Lorenzo-Seva, 2016). As regards the predictive validity of recreational activities, it was expected to replicate what has been reported in previous studies (Grucza & Goldberg, 2007; Cupani & Lorenzo-Seva, 2016); that is, that Drug Use is correlated positively with Extraversion and negatively with Conscientiousness, and Irresponsibility is correlated negatively with Conscientiousness. Positive correlations were also expected in the Friendship activity with the Agreeableness/Extraversion factors, Erudition/Creativity with Openness, and Communication with Openness and Agreeableness.

Method

Type of Study

According to the classification performed by Montero and León (2002), the present work is considered an instrumental study, as this type of studies aims at the development of tests and apparatus including both their design and adaptation.

Participants

The sample consisted of 910 participants, 543 women (59.6%) and 367 men (40.03%), age range 15-80 years (M= 29.52; DT= 12.25), from the city of Córdoba, Argentina. As regards the participants' educational level, 6.6% were attending postgraduate courses, 19.6% had complete university education, 51.3% had incomplete university education, 12.6% had complete secondary education, 7.4% had incomplete secondary education, and the remaining 3% had complete primary education. Only one participant had incomplete primary education. As regards the current employment status, the sample composition was the following: 40% students (28.80% university students and 11.20% secondary level students), 32.3% employees, 17% professionals (10% full-time employees and 7% self-employed with no employees), 4% retired or pensioners and informal workers (2% of each category), and the remaining 4% were company owners or partners, full-time employee or self-employed technicians, skilled and unskilled self-employed workers. Only one person was unemployed at the moment of the data collection. The sampling was of accidental type (Kumar, 2005), as the study was carried out in institutions with the permission of their authorities. Concerning instrument administration, all the participants (n= 910) answered the IPIP Five-Factor Domain Scale (Goldberg, 1999). Some participants (n = 229) answered the inventory again a month later to evaluate the punctuation stability (Test-Retest). The NEO-FFI inventory was answered by 209 participants to evaluate the convergent validity. The six scales of the Frequency Inventory on Daily Living Activities were answered by 402 participants to perform the predictive validity study.

Instruments

IPIP Five-Factor Domain Scale (Goldberg, 1999). It consists of 100 items that define the five personality domains: Openness (O), Conscientiousness (C), Extraversion (E), Agreeableness (A), and Neuroticism (N). Each domain is measured by 20 items and is written in phrase form describing typical people's behavior. A subject is asked to evaluate the degree of precision with which each phrase describes him/her using a scale of five answer options (from "I strongly disagree with this description of myself" to "I agree a lot with this description of myself"). Goldberg reports Cronbach's alpha values from 0.88 to 0.91 for the IPIP scales and an average correlation between the IPIP scales and the 100 unipolar markers of 0.70 or 0.78 when the correlation for the attenuation was used (IPIP. (s.f.). Restored on 9th March, 2018, from https://tr.im/81f52).

NEO-FFI Personality Inventory (Costa & McCrae, 1992). It consists of 60 items that provide a fast and general measure of the Big Five personality factors. Each scale consists of 12 elements that describe a person's typical behavior and allows to measure each of the factors. The answer format is Likert type with five options from "total disagreement" to "total agreement" with a certain phrase (Costa & McCrae, 1992). The reliability indices were α = 0.82 for Extraversion, α = 0.70 for Agreeableness, α = 0.80 for Conscientiousness, α = 0.83 for Neuroticism, and α = 0.74 for Openness to Experience.

Frequency Inventory on Daily Living Activities (Goldberg, 1999). Six scales from the Frequency Inventory on Daily Living Activities were used in this study. These scales consist of 54 items that measure six types of recreational activities distributed as follows. Relatively undesirable activities: Drug Use (14 items) and Irresponsibility (seven items); desirable activities: Creativity (11 items) and Friendship (eight items); and relatively neutral activities: Communication (eight items) and Erudition (six items). A participant is asked to mention how frequently he/she has performed certain tasks (e.g., going to a public library) during the past year using a Likert-type scale of five answer

options from (1) never in my lifetime, (2) not in the past year, (3) once or twice during the past year, (4) three or more times, but no more than 15 times during the past year, and (5) more than 15 times in the past year. The reliability values of these scales vary from $\alpha = 0.67$ (Communication) to $\alpha = 0.87$ (Drug Use).

Procedure

The data collection was performed both collectively and individually. For the collective collection, secondary level students and university students from *Universidad Nacional de Córdoba* (UNC) completed the instruments in their classrooms with the corresponding authorization from teachers and authorities and the participants' prior informed consent. For the individual collection, the authors' friends, relatives, workmates and acquaintances were asked to participate in the study. All of them had previously been told about the aim of the work and the estimated time to complete the instruments, which might vary among one, two, or three weeks. They were also informed about the approximate time to answer each instrument and that they were expected to answer each test with no breaks in time. Then, they were given a closed envelope with the informed consent and the corresponding instruments.

Data Analysis

The Analysis of Missing Values routine from software SPSS version 19.0 was used to evaluate the pattern of missing values (Tabachnick & Fidell, 2007). Then, the sample was divided into two parts at random. The first half (n = 455) was chosen to identify the best markers of the five factors. To select the main 30 items, the 100 items from the IPIP Five-Factor Domain Scale (Goldberg, 1999) were analyzed using exploratory factor analysis (EFA), previously controlling the variance due to Acquiescence (Ferrando, Lorenzo-Seva, & Chico, 2009). This analysis was performed with the MATLAB program and FACTOR 9.3 (Lorenzo-Seva & Ferrando, 2013). As an external criterion, the correlation between the items of one dimension and the total scores of the same NEO-FFI dimension was considered. As a last criterion, it was tried to represent each scale by different contents. The second half of the sample was reserved to conduct the non-restricted confirmatory factor analysis of the 30 selected items (Ferrando & Lorenzo-Seva, 2000). The variance due to Acquiescence was biased, and the residual correlation matrix was factored through LISREL 8.5. The second half of the sample (n = 455) was reserved to conduct the non-restricted confirmatory factor analysis of the 30 selected items.

Then, an EFA with the total sample (N=910) was performed, from which the item factor saturations and the factor weights were estimated to then determine the factor scores. Moreover, the correlation corrected between the item response and the factor scores was calculated. When doing this calculation, the factor scores of the five factors plus the score in acquiescence was informed. To control the acquiescence, it is necessary a test with reversed items. It should be stated that the interpretation of the scores free of acquiescence is obtained by calculating the factor scores (instead of the scores obtained as the mere sum of item responses). In this way, the content scores are free of acquiescence. On the other hand, it can be interpreted that people with extreme scores in acquiescence have not answered the test seriously.

For the internal consistency study, the factor score reliability was calculated (Mislevy & Bock, 1990) and the test-retest method was used to determine the temporal stability.

With respect to the convergent validity estimation, the correlations between the factor scores and the direct scores of the five NEO-FFI scales, which are supposed to be similar, were analyzed. Then, an evidence study of concurrent validity was performed comparing the average scores of the scales between men and women, and youth and adults, through a multivariate variance analysis 2 (sex) x 2 (age) (MANOVA). With this analysis, it was intended to compare if the average factor scores of the groups theoretically agree with the personality profiles identified by the literature (Cupani et al. 2014). In accordance with what was proposed in previous studies (McCrae et al., 2000), participants younger than 30 years old were included in the young group. To estimate the effect size, the eta-squared coefficient (η^2) was calculated, following as a criterion what was suggested by Cohen (1992), that is, that the effect sizes (η^2) 1%, 10% and 25% are considered small, medium-size and big, respectively.

Finally, a multiple regression analysis (*enter* method) was performed with the aim of evaluating the predictive capacity of the five factors on the dependent variable *recreational activities*, considering sub-samples according to the participants' sex and age. Depending on the sample size, the statistical power of each model was calculated with the G*Power 3.1 program (Faul, Erdfelder, Buchner, & Lang, 2009).

Results

The results of (a) the missing case analysis, (b) the exploratory and confirmatory factor analysis, (c) reliability, and (d) evidence of convergent and predictive validity are presented below.

Missing Case Analysis

The percentage of missing cases was not over 5%; due to this, it was decided to impute the missing data by a central tendency measure (fashion) because this method provides an attractive balance between precision and conceptual simplicity (Shrive, Stuart, Quan, & Ghali, 2006). It was chosen to impute by fashion to try to count on the five (discrete) response options, proper of the scale, to estimate the polychoric correlations.

Development of the 30-item Inventory

Exploratory Factor Analysis (EFA). Firstly, the 100 items were analyzed through a polychoric correlation matrix. The estimation method used was the unweighted least square (ULS), which is a more robust estimator, with less biased estimates, and adequate when samples are not so big (Ferrando & Lorenzo-Seva, 2013). The Káiser-Mayer-Olkin (.883) sampling adequacy method indicated the feasibility of performing the Factor Analysis. Five factors were extracted and the Varimax rotation was used, since, on a theoretical level, the personality dimensions are expected to be independent in the population (i.e., orthogonal). The variance due to AC (Lorenzo-Seva & Ferrando, 2009) had been previously controlled. Because the IPIP 100-item scales are partially balanced, the procedure proposed by Lorenzo-Seva and Ferrando (2009) was used: a) the number of expected factors according to the content and an additional factor related to AC are retained; b) the first centroid is calculated and taken

Table 1.
Rotated factor structure and descriptive indices

	Ítem	M	SD	rc	AC	EX	NE	CO	AA	OP
1	IPIP21	3.57	0.93	0.49	-0.05	0.70	-0.18	0.11	0.14	-0.09
2	IPIP26	3.70	0.88	0.55	0.00	0.74	-0.05	0.08	0.17	0.16
3	IPIP66	2.83	0.88	0.40	0.10	-0.58	-0.06	-0.11	-0.10	-0.13
4	IPIP65	2.60	1.03	0.35	0.11	0.56	0.10	-0.15	-0.18	0.15
5	IPIP16	2.98	1.09	0.38	0.29	-0.56	-0.18	0.20	-0.05	-0.11
6	IPIP76	3.47	0.96	0.47	0.14	0.69	-0.13	0.09	0.16	-0.01
7	IPIP14	3.29	1.07	0.40	0.15	-0.03	-0.63	0.03	-0.06	0.05
8	IPIP53	2.74	1.06	0.56	0.26	-0.16	0.70	-0.04	-0.03	-0.03
9	IPIP29	2.71	1.05	0.47	0.21	-0.05	0.66	0.01	0.05	-0.01
10	IPIP69	2.66	1.13	0.60	0.12	0.00	0.78	0.00	-0.15	-0.04
11	IPIP89	2.72	1.10	0.36	0.20	-0.05	0.52	-0.19	0.29	0.07
12	IPIP99	2.98	1.14	0.45	0.13	0.01	0.62	-0.10	-0.08	-0.10
13	IPIP33	2.64	1.09	0.33	0.24	-0.01	0.22	-0.55	0.00	0.14
14	IPIP38	2.43	1.03	0.45	0.24	-0.10	0.09	-0.68	-0.10	-0.09
15	IPIP62	3.88	0.94	0.29	0.26	0.05	-0.02	0.47	0.13	-0.12
16	IPIP82	3.56	0.85	0.53	0.23	0.03	-0.09	0.72	0.05	0.10
17	IPIP88	3.26	0.94	0.44	0.14	-0.02	0.07	0.65	0.11	0.02
18	IPIP92	3.61	0.83	0.42	0.13	0.05	0.04	0.64	0.09	0.14
19	IPIP2	2.22	1.14	0.20	0.25	0.05	-0.04	-0.12	-0.43	-0.07
20	IPIP42	2.08	0.97	0.46	0.07	-0.11	0.06	-0.01	-0.75	-0.01
21	IPIP51	3.78	1.03	0.48	0.32	0.03	0.18	0.03	0.64	0.06
22	IPIP57	3.84	0.85	0.53	0.16	0.09	-0.11	0.06	0.78	0.12
23	IPIP81	3.51	0.98	0.59	0.13	0.08	0.03	0.07	0.74	0.02
24	IPIP87	3.79	0.85	0.55	0.12	0.13	-0.06	0.13	0.73	-0.01
25	IPIP5	3.64	0.88	0.32	0.18	0.10	-0.06	0.25	-0.09	0.58
26	IPIP15	2.37	1.02	0.39	0.28	-0.07	-0.02	-0.10	0.03	-0.67
27	IPIP25	3.83	0.91	0.35	0.18	0.22	-0.05	-0.11	0.15	0.60
28	IPIP70	2.19	0.98	0.32	0.20	-0.18	0.04	0.01	-0.14	-0.57
29	IPIP80	2.85	1.08	0.32	0.04	-0.04	-0.05	0.10	0.11	0.56
30	IPIP84	3.85	0.87	0.25	0.08	-0.16	-0.13	0.08	0.31	0.49

Note: M = Media; SD = Standard Deviation; rc = Corrected Correlation; AC = Acquiescence; EX = Extraversion; NE = Neuroticism; CO = Conscientiousness; AA = Agreeableness; OP = Openness. The saturations over 0.30 in absolute values are presented in bold.

as an estimate of the factor weight of each item on the AC factor; c) this set of estimates (one estimate per item) is used as an target in a congruent rotation to compute the factor loading value of each item content in the AC factor; d) once the factor related to AC is available, the variance explained by this factor is removed, and it was proceeded to identify the factors related to the instrument content on the residual matrix. From this study, it was observed that the factor loading values in the first AC factor varied between -0.03 and -0.71 for Agreeableness, between -0.04 and 0.63 for Extraversion, between 0.08 and 0.62 for Conscientiousness, between 0.14 and -0.84 for Neuroticism, and between 0.00 and -0.50 for Openness. Secondly, a correlation was performed a correlation (Spearman rank) among the 20 items in a dimension (e.g., Neuroticism) and the direct score of the factor theoretically similar to NEO-FFI (e.g., Neuroticism). The item-factor correlations varied between -0.50 and 0.50 for Extraversion, between -0.50 and 0.47 for Agreeableness, between -0.66 and 0.57 for Conscientiousness, between -0.44 and 0.73 for Neuroticism, and between -0.45 and 0.52 for Openness. Thirdly, it was tried to ensure that each dimension was represented by different personality trait contents. As a result of the combination of these criteria (items representing lower AC, higher factor loading values and item-factor correlations, and items representing the different personality traits), the best 30 items were selected. According to these results, it was observed that the IPIP-R-30, once the AC effect was controlled, presents a simple and orthogonal factor structure.

Confirmatory factor analysis (CFA). On the basis of the exploratory factor analysis, the 30 items that combined the three strategies mentioned before were selected; they were used for the unrestricted confirmatory factor analysis (Ferrando & Lorenzo-Seva, 2000) on the second sample. The variance due to AC was biased and the residual correlation matrix was factored using LISREL. To evaluate the model fit the already mentioned criteria were used. On the basis of the saturation factors obtained in the EFA, an item was selected as the marker for each factor. The obtained goodness-of-fit statistics indicated that the model fits adequately to data (GFI 0.96; CFI 0.94; RMSEA 0.05). Then, the total sample (N=910) was used to estimate the item saturation factors, and the factor weights were used to estimate the factor scores. The CFA suggests that the 30 items of the questionnaire present an adequate fit to data. Table 1 shows the factor loadings of the content factors.

Reliability. The discrimination indices for the 30 items varied between 0.20 and 0.60 (average 0.43), out of which three items presented values below the cut-off point (See Table 3). The reliability coefficients varied between 0.79 and 0.86. For temporal stability, the indices varied between $r = 0.71$ in the Extroversion factor and $r = 0.80$ in the Conscientiousness factor (See Table 2). Although the scale reliability was satisfactory as regards both the internal consistency and the temporal stability, it can be seen that, when compared with the 50-item version, the 30-item version registers and average loss of 4.82% in the internal consistency and of 5.75%.in the temporal stability.

Validity evidence. In the convergent validity study, the mean correlation was r = 0.61, with a range from r = 0.45 for Agreeableness to r = 0.71 for Conscientiousness (See Table 2). In the group difference study, an effect was observed from small to medium, significant for sex and age. Women presented higher values than men for Agreeableness ($\eta^2 = 0.09$), whereas men presented the highest values for Neuroticism ($\eta^2 = 0.01$). Regarding age, it was observed

Table 2.
Reliability indices, convergence studies, and differentiated studies according to sex and age

Cinco factores	Fiabilidad		Conv	Sexo						Edad					
				Femenino (n = 543)		Masculino (n = 367)				Jóvenes (n = 579)		Adultos (n = 331)			
	FCR	T–RT	r	M	DT	M	DT	F 1,906	η^2	M	DT	M	DT	F 1,906	η^2
IPIP–30															
Extraversión	.83	.71	.67	49.96	10.08	50.07	9.93	0.32	.00	50.74	10.40	48.72	9.17	7.88	.01
Amabilidad	.84	.76	.45	52.61	9.00	46.11	10.17	90.61	.09	50.40	9.80	49.27	10.34	0.73	.00
Responsabilidad	.83	.80	.71	49.65	9.82	50.52	10.26	1.46	.00	49.45	10.03	50.96	9.89	4.62	.01
Estabilidad Emocional	.86	.75	–.67	49.17	10.05	51.20	9.81	8.35	.01	49.79	10.11	50.34	9.80	0.36	.00
Intelecto	.79	.75	.61	49.50	9.74	50.72	10.34	3.51	.00	49.88	10.06	50.18	9.91	0.15	.00

Note: N = 910; FCR = Reliabilty with rotated component; T-RT = Test – retest; Conv. = Convergence with NEO-FFI scales. * $p \leq .05$; ** $p \leq .001$

that the youth presented higher values than adults for Extraversion (η^2 = 0.01), whereas adults presented the highest values for Conscientiousness (η^2 = 0.01). It can be observed that the five IPIP-R-30 markers evidenced a lower convergent validity with the NEO-FFI scales (5.85%) in contrast with the 50-item version.

For the test-criterion validity study, in the sub-sample Young Women the personality traits explain between 10% and 18% of the variance (See Table 3). The traits related to D*rug Use* are Conscientiousness (β= -0.19) and Extraversion (β= 0.24), to *Erudition* is Openness ($\beta = 0.32$), to *Communication* are Agreeableness ($\beta = 0.18$) and Openness ($\beta=0.28$), and to *Friendship* are Extraversion ($\beta = 0.26$), Agreeableness (β= 0.19), Conscientiousness ($\beta = 0.19$), and Openness (β= 0.19). In the sub-sample Adult Women, the personality traits explain between 13% and 18% of the variance. The traits related to *Drug Use* are Conscientiousness (β= -0.24) and Extraversion (β= 0.21), to *Irresponsibility* are Conscientiousness (β= -0.35), to *Creativity* and *Friendship* is Openness ($\beta = 0.30$ y $\beta = 0.30$, respectively). In the sub-sample Young Men, the traits explain between 19% y 29% of the variance. The trait related to *Erudition* is Openness ($\beta = 0.40$), to *Communication* are Extraversion (β = 0.20) and Openness (β = 0.29), and to *Friendship* are Extraversion ($\beta = 0.32$), Agreeableness (β= 0.25), and Openness (β= 0.25). In the sub-sample Adult Men, the personality traits explain between 15% and 38% of the variance. The traits related to *Drug Use* are Extraversion (β= 0.27) and Conscientiousness (β = -0.21), to *Irresponsibility* are Extraversion (β = 0.20) and Conscientiousness (β= -0.36), to *Erudition* are Extraversion (β = 0.23), Agreeableness ($\beta = 0.24$), Neuroticism (β = 0.28), and Openness ($\beta = 0.39$), to *Creativity* are Neuroticism ($\beta = 0.39$) and Openness (β = 0.24), and to *Friendship* are Extraversion ($\beta = 0.30$) and Agreeableness ($\beta = 0.38$).

Discussion

In accordance with the growing increase of personality brief scales (e.g., Mini-IPIP; Donnellan et al., 2006), this work examined the psychometric properties of the IPIP-R Personality Inventory in its 30-item brief version (IPIP-R30 Five-Factor Domain Scale) in a sample of Argentinean people. The results indicate that the IPIP-R-30, once the AC effect was controlled, presents a simple and orthogonal factor structure. The CFA suggests that the 30 items of the questionnaire present an adequate fit to data. These results agree with previous research demonstrating that brief inventories present a fit of the five factors to the factor structure (Baldasaro et al., 2013; Donellan et al., 2006; Soto & John, 2017). Similarly, it was observed that the factor weights calculated with EFA were, in average, 4.8% higher (between -0.43 and 0.78) than the ones observed for the 50-item version proposed by Cupani and Lorenzo-Seva (2016).

On the other hand, the scale reliability was satisfactory, as regards both the internal consistency and the temporal stability. However, in a 30-item version compared with a 50-item scale, a decrease in the internal consistency (4.82%) and the temporal stability (5.75%) was observed. These results agree with the specific literature on the topic, which reports that in the brief versions of the questionnaires the internal consistency is usually the most affected aspect, although both the temporal stability and the inter-examiner reliability decrease as well (Baldasaro, et al., 2013; Sibley, 2012).

The convergent, discriminant, and predictive validities also decrease in comparison with the extended versions (Baldaraso et al., 2013). In fact, the five IPIP-R-30 markers evidenced a convergent validity with the NEO-FFI scales a 5.85% lower in comparison with the 50-item version. Specifically, this work showed high associations among the Extraversion, Neuroticism, and Conscientiousness traits and, in a lower degree, between Openness and Agreeableness, similar results to the ones obtained in other investigations (e.g., Gow et al., 2005).

As regards the group differences according to sex, women presented higher levels in Agreeableness and lower levels in Neuroticism and Openness in comparison with men, which confirms what was proposed in previous studies (Cupani, et al., 2014; McCrae & Terracciano, 2005). On the other hand, when considering the group differences as regards age, it was observed that the average levels of Extraversion decrease with age, whereas the ones related to Conscientiousness increase. These small-sized effect changes are similar to the ones observed in previous studies on five-factor inventories (McCrae et al., 2000; Ledesma, Sánchez & Díaz-Lázaro, 2011; Srivastava et al., 2003), which provides evidence of the IPIP-R-30 concurrent validity.

In a final study, a multiple regression analysis was performed to estimate the degree in which each scale in the IPIP-R-30 contributes independently to explain how frequently individuals perform certain recreational activities. This study was carried out in different sub-samples considering the participants' sex and age. Indeed, the five scales allow to explain between 12% and 37% of the variance of recreational activities. In general, these results agree with the ones reported by Grucza and Goldberg (2007), where Drug Use is related to Extraversion, Irresponsibility is related to Conscientiousness, Erudition is related to Openness, Communication is related to Agreeableness and Openness, Creativity is related to Openness, and Friendship is related to Extraversion, Agreeableness, and Openness.

Table 3.

Prediction of recreational activities from the five factors.

Sample	Young Women			Adult Women			Young Men			Adult Men		
Dependent	Drug Use (α = 0.87)											
Predictor	*β*	*t*	*p*	*β*	*t*	*p*	*β*	*t*	*p*	*β*	*t*	*p*
Extraversion	0.24	2.74	0.01	0.21	2.22	0.03	0.04	0.35	0.72	0.27	2.41	0.02
Agreeableness	-0.04	-0.50	0.62	-0.17	-1.82	0.07	-0.05	-0.50	0.62	-0.18	-1.61	0.11
Conscientiousness	-0.19	-2.14	0.03	-0.24	-2.50	0.01	-0.22	-2.17	0.03	-0.21	-1.96	0.05
Neurotiscism	-0.04	-0.50	0.61	-0.10	-1.02	0.31	-0.15	-1.46	0.15	0.02	0.22	0.83
Openness	0.04	0.48	0.63	0.12	1.25	0.21	0.01	0.14	0.89	0.05	0.41	0.68
F	(5.122) 2.658*			(5.96) 4.691***			(5.94) 1.302			(5.71) 2.826*		
R^2	0.10			0.18			0.06			0.17		
1-β	0.68			0.88			0.35			0.67		
	Irresponsibility (α = 0.77)											
Extraversion	0.14	1.62	0.11	0.10	1.07	0.29	0.06	0.60	0.55	0.21	1.98	0.05
Agreeableness	-0.03	-0.30	0.76	-0.07	-0.72	0.47	-0.05	-0.50	0.62	0.13	1.17	0.25
Conscientiousness	-0.16	-1.83	0.07	-0.35	-3.64	0.00	-0.18	-1.74	0.09	-0.36	-3.41	0.00
Neurotiscism	-0.15	-1.74	0.08	-0.15	-1.54	0.13	0.01	0.11	0.91	-0.07	-0.63	0.53
Openness	-0.03	-0.39	0.70	0.06	0.58	0.56	0.16	1.58	0.12	0.18	1.68	0.10
F	(5.122) 1.914			(5.96) 4.212***			(5.94) 1.568			(5,71) 4.452**		
R^2	0.10			0.18			0.08			0.24		
1-β	0.68			0.88			0.46			0.88		
	Erudition (α = 0.71)											
Extraversion	-0.09	-1.08	0.28	0.07	0.72	0.47	-0.11	-1.22	0.23	0.23	2.36	0.02
Agreeableness	-0.04	-0.44	0.66	0.04	0.37	0.71	0.14	1.52	0.13	0.24	2.46	0.02
Conscientiousness	0.08	0.95	0.35	-0.14	-1.37	0.17	-0.07	-0.70	0.48	0.05	0.52	0.60
Neurotiscism	0.16	1.82	0.07	0.14	1.45	0.15	0.01	0.13	0.90	0.28	2.95	0.00
Openness	0.32	3.79	0.00	0.25	2.54	0.01	0.40	4.20	0.00	0.39	4.14	0.00
F	(5.122) 3.672**			(5.96) 2.254			(5.94) 5.077***			(5.71) 8.636***		
R^2	0.13			0.11			0.21			0.38		
1-β	0.83			0.59			0.93			0.99		
	Communication (α = 0.67)											
Extraversion	0.02	0.27	0.79	0.17	1.70	0.09	0.20	2.08	0.04	0.11	1.01	0.32
Agreeableness	0.18	2.07	0.04	0.05	0.49	0.62	0.18	1.93	0.06	0.36	3.11	0.00
Conscientiousness	0.07	0.78	0.44	-0.17	-1.68	0.10	0.00	-0.05	0.96	0.09	0.81	0.42
Neurotiscism	0.10	1.13	0.26	-0.05	-0.45	0.65	-0.07	-0.73	0.47	-0.08	-0.73	0.47
Openness	0.28	3.31	0.00	0.20	1.96	0.05	0.29	2.97	0.00	0.12	1.08	0.29
F	(5.122) 3.570**			(5.96) 2.100			(5.94) 4.431***			(5.71) 2.430*		
R^2	0.13			0.10			0.19			0.15		
1-β	0.83			0.56			0.89			0.6		
	Creativity (α = 0.76)											
Extraversion	0.13	1.45	0.15	0.13	1.37	0.17	0.17	1.68	0.10	0.19	1.79	0.08
Agreeableness	0.05	0.55	0.59	-0.10	-1.07	0.29	0.06	0.61	0.54	-0.14	-1.28	0.20
Conscientiousness	-0.01	-0.14	0.89	-0.12	-1.25	0.21	-0.09	-0.87	0.39	-0.20	-1.93	0.06
Neurotiscism	0.11	1.18	0.24	-0.08	-0.81	0.42	0.00	0.03	0.98	0.39	3.83	0.00
Openness	0.21	2.37	0.02	0.30	3.04	0.00	0.17	1.68	0.10	0.24	2.36	0.02
F	(5.122) 1.856			(5.96) 3.326**			(5.94) 1.770			(5.71) 5.560***		
R^2	0.07			0.15			0.09			0.28		
1-β	0.52			0.77			0.49			0.94		
	Friendship (α = 0.75)											
Extraversion	0.26	3.21	0.00	0.18	1.81	0.07	0.32	3.64	0.00	0.30	2.86	0.01
Agreeableness	0.19	2.26	0.03	0.14	1.40	0.16	0.25	2.81	0.01	0.38	3.49	0.00
Conscientiousness	0.19	2.36	0.02	-0.03	-0.29	0.77	0.02	0.21	0.84	-0.03	-0.33	0.74
Neurotiscism	0.04	0.45	0.66	-0.02	-0.20	0.84	-0.13	-1.49	0.14	0.12	1.14	0.26
Openness	0.19	2.26	0.03	0.30	3.03	0.00	0.25	2.71	0.01	0.15	1.46	0.15
F	(5.122) 5.293***			(5.96) 2.939*			(5.94) 7.498***			(5.71) 5.031***		
R^2	0.18			0.13			0.29			0.26		
1-β	0.96			0.71			0.99			0.92		

In general, the results obtained suggest that the IPIP-R-30 scales have adequate psychometric properties; however, some limitations should be mentioned. In the same way as in other brief personality instruments, there is a psychometric cost of the IPIP-R brief version. In comparison with the 50-item measures, the brief version is less reliable, correlates less strongly with the NEO-FFI scales and, at the level of content, it is less represented by the different personality traits. On the other hand, as regards the procedure, the extended period of time in which the questionnaire was administered might have increased the AC bias, although there is no evidence to support that in this study. These limitations, however, do not restrain its practical use (Donnellan et. al, 2006) because brief scales eliminate the topic redundancies and, therefore, at the same time they eliminate fatigue, frustration, and boredom associated with answering similar questions over and over again (Burisch, 1984), among other advantages (Gosling et al., 2003; Saucier & Goldberg, 2002).

Practical Implications

Although there is a slight decrease of certain psychometric properties in the IPIP-R-30 version (lower to 6%) when compared with the 50-item version, the time saving implied in administering a brief scale in certain studies, in which minimizing the evaluation time and the respondent fatigue are vital aspects, compensates this loss. Thus, the time to respond completely to an inventory is estimated between four and ten minutes for a 60-item one, between three and five for a 30-item one, and between two and three minutes for a 15-item one (Soto & John, 2017). This modest time saving can be relevant for certain types of studies. Some examples of studies in which it is really convenient to use brief scales include: the large-scale surveys designed to evaluate different constructs, longitudinal studies requiring that each participant completes the same personality measure in different times, laboratory studies in which part of the time should be destined to experimental manipulations and behavior observations, and pilot studies or research in which factors such as boredom, fatigue, and lack of interest might generate a careless or at random response pattern (Credé et al., 2012; Milojev et al., 2013; Soto & John, 2017; De Vries, 2013).

On the other hand, some studies have evidenced that personality tests administered to samples of individuals with low educational level, with lower to average intelligence, adolescents, preadolescents and older adults present the worst fit to the FFM (Meisenberg & Williams, 2008; Rammstedt, Goldberg, & Borg, 2010; Ross & Mirowsky, 1984; Soto, John, Gosling, & Potter, 2008; Vigil-Colet, Lorenzo-Seva, & Morales-Vives, 2015). In fact, Ramsteed et al. (2010) demonstrated that the FFM factor structure only remains the same throughout the different educational levels when the AC bias is removed. Therefore, the validity of the personality tests in these subpopulations might be compromised due to high levels of AC bias characteristic in these groups, explaining why it is relevant to use scales free of the AC bias, such as the IPIP-R-30, in the evaluation of those individuals.

According to what was previously stated, it was considered adequate to select six items per scale because recent research indicates that the use of a very brief form (e.g., two items) is associated with the increase of Type 1 and Type 2 errors (Credé et al., 2012; Kruyen et al., 2013). In this way, an instrument was designed with a simple orthogonal structure of five factors, satisfactory reliability indices (internal consistency and temporal stability), evidence of convergent validity with the NEO-FFI scales, group differences regarding sex and age, and predictive validity of recreational activities.

In summary, the results of the present study suggest that both versions represent a valid alternative to measure the personality traits in our population. The results also suggest that the IPIP-R-30 scales have adequate psychometric properties and might be used with relative reliability as an auxiliary tool in research and professional work (clinic, work-related, etc.). Moreover, one of the main advantages of this work is that the obtained scores are free from the AC bias.

References

Aronson, Z. H., Reilly, R. R., & Lynn, G. S. (2006). The impact of leader personality on new product development teamwork and performance: The moderating role of uncertainty. *Journal of Engineering and Technology Management*, *23*(3), 221–247.

Baldasaro, R. E., Shanahan, M. J., Bauer, D. J. (2013). Psychometric Properties of the Mini-IPIP in a Large, Nationally Representative Sample of Young Adults. *Journal of Personality Assessment*, *95*(1), 74–84. Doi: 10.1080/00223891.2012.700466

Burisch, M. (1984). Approaches to personality inventory construction: A comparison of merits. *American Psychologist*, *39*(3), 214.

Chamorro-Premuzic, T., & Furnham, A. (2009). Mainly Openness: The relationship between the Big Five personality traits and learning approaches. *Learning and Individual Differences*, 19(4), 524-529.

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155-159

Costa, P., & McCrae, R.R. (1992). NEO PI-R manual profesional. Odessa, FL: *Evaluación Psicológica Resources*, Inc.

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

Cronbach, L.J. (1942). Studies of acquiescence as a factor in the true–false test. *Journal of Educational Psychology*, *33*, 401–415.

Cupani, M. (2009). El Cuestionario de Personalidad IPIP-FFM: Resultados preliminares de una adaptación en una muestra de preadolescentes argentinos. *Perspectivas en Psicologia*, 6, 51–58.

Cupani, M. & Lorenzo-Seva, U. (2016). The development of an alternative IPIP inventory measuring the Big-Five factor markers in an Argentine sample. *Personality and Individual Differences*, *91*, 40–46. http://dx.doi.org/10.1016/j.paid.2015.11.051

Cupani, M., Pilatti, A., Urrizaga, A., Chincolla, A., & de Minzi, M. C. R. (2014). Inventario de Personalidad IPIP-NEO: estudios preliminares de adaptación al español en estudiantes argentinos. *Revista Mexicana de Investigación en Psicología*, *6*(1), 55-73.

De Vries, R.E. (2013). The 24-item Brief HEXACO Inventory (BHI). *Journal of Research in Personality*, *47*, 871–880. http://dx.doi.org/10.1016/j.jrp.2013.09.003

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The Mini-IPIP scales: Tiny-yet-effective measures of the big five factors of personality. *Psychological Assessment*, *18*(2), 192–203.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149-1160.

Ferrando, P. J., & Lorenzo Seva, U. (2000). Unrestricted versus restricted factor analysis of multidimensional test items: Some aspects of the problem and some suggestions. *Psicológica*, *21*(2), 301-323.

Ferrando, P. J., & Lorenzo-Seva, U. (2013). Unrestricted item factor analysis and some relations with item response theory. *Recuperado de http://psico.fcep. urv.es/utilitats/factor/[Links]*.

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*(2), 364-381.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality psychology in Europe*, *7*(1), 7-28.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., et al. (2005). The international personality item pool and the future of public domain personality measures. *Journal of Research in Personality*, 40, 84–96.

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

Gow, A.J., Whiteman, M.C., Pattie, A., & Deary, I.J. (2005). Goldberg's "IPIP" Big-Five factor markers: Internal consistency and concurrent validation in Scotland. *Personality and Individual Differences*, *39*, 317–329.

Guenole, N., & Chernyshenko, O. (2005). The suitability of Goldberg's Big-Five IPIP person- ality markers in New Zealand: A dimensionality, bias, and criterion validity evaluation. *New Zealand Journal of Psychology*, 34, 86–96.

Gross, M.N., Zalazar Jaime, M.F., Piccolo, N.V., & Cupani, M. (2012). Nuevos estudios de validación del cuestionario de personalidad IPIP-FFM. X Congreso Latinoamericano de Sociedades de Estadística, Córdoba, Argentina.

Grucza, R. A., & Goldberg, L R. (2007). The Comparative Validity of 11 Modern Personality Inventories: Predictions of Behavioral Acts, Informant Reports, and Clinical Indicators. *Journal of Personality Assessment, 89*(2), 167–18

Hofstee, W. K. B., ten Berge, J. M. F., & Hendriks, A.A.J. (1998). How to score questionnaires. *Personality and Individual Differences, 25*, 897–909.

Javeline, D. (1999). Response effects in polite cultures: a test of acquiescence in Kazakhstan. *Public Opinion Quarterly, 63*(1), 1–28.

Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The relationship between culture and response styles: Evidence from 19 countries. *Journal of Cross-Cultural Psychology, 36*, 264-277.

Kumar, R. (2005). Research Methodology: A Step-by-Step Guide for Beginners (Second Edition). Thousand Oaks, California: SAGE.

Kruyen, P. M., Emons, W. H. M., & Sijtsma, K. (2013). On the shortcomings of shortened tests: A literature review. *International Journal of Testing*, *13*, 223–2484.

Langford, P. H. (2003). A one-minute measure of the Big Five? Evaluating and abridging Shafer's (1999a) Big Five markers. *Personality and Individual Differences*, *35*(5), 1127–1140.

Ledesma, R. D., Sánchez, R., & Díaz-Lázaro, C. M. (2011). Adjective checklist to assess the big five personality factors in the Argentine population. *Journal of Personality Assessment*, *93*(1), 46-55.

Lorenzo-Seva, U., & Ferrando, P. J. (2009). Acquiescent responding in partially balanced multidimensional scales. *British Journal of Mathematical and Statistical Psychology*, *62*(2), 319-326.

Lorenzo-Seva, U., & Ferrando, P. J. (2013). Factor 9.2: A comprehensive program for fitting exploratory and semi-confirmatory factor analysis and IRT models. *Applied Psychological Measurement*, 37(2), 497-498.

McCrae, R. R., Costa Jr, P. T., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M. D., ... & Saunders, P. R. (2000).

Nature over nurture: Temperament, personality, and life span development. *Journal of personality and social psychology*, *78*(1), 173.

McCrae, R. R., & Terracciano, A. (2005). Universal features of personality traits from the observer's perspective: data from 50 cultures. *Journal of personality and social psychology*, *88*(3), 547.

Meisenberg, G., & Williams, A. (2008). Are acquiescent and extreme response styles related to low intelligence and education? *Personality and Individual Differences*, 44(7), 1539-1550.

Milojev, P., Osborne, D., Greaves, L.M., Barlow, F.K. & Sibley, C.G. (2013). The Mini-IPIP6: Tiny yet highly stable markers of Big Six personality. *Journal of Research in Personality*, *47*, 936–944. http://dx.doi.org/10.1016/j.jrp.2013.09.004

Mislevy, R. J., & Bock, R. D. (1990). *BILOG 3: Item analysis and test scoring with binary logistic models*. Scientific Software International.

Mlačić, B., & Goldberg, L.R. (2007). An analysis of a cross-cultural personality inventory: The IPIP Big-Five factor markers in Croatia. *Journal of Personality Assessment*, 88, 168–177.

Montero, I., & León, O. G. (2002). Clasificación y descripción de las metodologías de investigación en Psicología. *International journal of clinical and health psychology*, 2(3), 503-508

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212.

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., & Borg, I. (2013). Correcting Big Five personality measurements for acquiescence: An 18-country cross-cultural study. *European Journal of Personality*, *27*(1), 71-81.

Ross, C. E., & Mirowsky, J. (1984). Components of depressed mood in married men and women the center for epidemiologic studies' depression scale. *American Journal of Epidemiology*, 119(6), 997-1004.

Saucier, G., & Goldberg, L. R. (2002). Assessing the big five: Applications of 10 psychometric criteria to the development of marker scales. B. De Raad, M. Perugini (Eds.), Big five assessment, Hogrefe & Huber, Seattle, WA (2002), pp. 29-58.

Sibley, C. G. (2012). The Mini-IPIP6: Item Response Theory analysis of a short measure of the big-six factors of personality in New Zealand. *New Zealand Journal of Psychology, 41*(3), 21-31.

Shrive, F. M., Stuart, H., Quan, H., & Ghali, W. A. (2006). Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. *BMC medical research methodology*, *6*(1), 57.

Soto, C. J. & John. O.P. (2017). Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, *68*, 69–81. http://dx.doi.org/10.1016/j.jrp.2017.02.004

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.

Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of personality and social psychology*, *84*(5), 1041.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Allyn & Bacon/Pearson Education.

Vazsonyi, A. T., Ksinan, A. Mikuška, J. & Jiskrova, G. (2015). The Big Five and adolescent adjustment: An empirical test across six cultures. *Personality and Individual Differences*, *83*, 234–244. http://dx.doi.org/10.1016/j.paid.2015.03.049

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(3), 209-215.

Zheng, L., Goldberg, L.R., Zheng, Y., Zhao, Y., Tang, Y., & Liu, L. (2008). Reliability and concurrent validation of the IPIP Big-Five Factor markers in China: Consistencies in factor structure between internet-obtained heterosexual and homosexual samples. *Personality and Individual Differences*, 45(7), 649–654. http://dx.doi.org/10.1016/j.paid.2008.07.009.