Personality and Individual Differences xxx (2018) xxx-xxx

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com

ELSEVIER

Is general intelligence responsible for differences in individual reliability in personality measures?

David Navarro-González, Pere Joan Ferrando, Andreu Vigil-Colet*

Universitat Rovira i Virgili, Research Center for Behavior Assessment, Spain

ARTICLE INFO

Article history:
Received 31 January 2018
Received in revised form 16 March 2018
Accepted 17 March 2018
Available online xxx

Keywords:
Person reliability
Response bias
Personality differentiation hypotheses

ABSTRACT

One possible hypothesis for personality differentiation is the higher reliability of high-ability individuals in typical response measures. This differential reliability has been explained as resulting from different verbal abilities as a consequence of the difficulties that low-ability individuals have in understanding items, or as the effect of response bias, or due to higher precision in the answers of high-ability individuals. The lack of an estimation of individual reliability has made it difficult to test these hypotheses. However, recent psychometric advances have made it possible to measure person reliability and thus address the issue. The present study analyses the relationships between person reliability measures and the response bias of different personality measures in measurements of intelligence in a sample of 532 adolescents. The results show that person reliability is more closely related to general intelligence than to specific abilities and that the results for low-ability individuals cannot be explained by verbal deficits or by higher levels of acquiescence or social desirability. The differential reliability of measures across ability levels therefore seems to be related to higher levels of traitedness in high-ability individuals, i.e. traits are represented in them with greater strength and clarity.

1. Introduction

The potential interactions between intelligence and personality measures are a subject that has generated considerable controversy for many decades. These interactions do not refer directly to the relationships between personality and intelligence, but rather to a series of problems related to (a) the extent to which intelligence levels affect the factorial structure of personality measures or the relationships between personality dimensions, and (b) the possibility that the level of differentiation of abilities may depend on certain personality dimensions.

The issue summarized above was first reported by Shure and Rogers (1963), who found that the factor structure of personality scales differed as a function of individual levels of intelligence, and Eysenck and White (1964), who found a different factor structure of intelligence depending on individual levels of neuroticism. These types of result were later integrated into the personality differentiation hypothesis (PDH) framework developed by Brand, Egan, and Deary (1994). The PDH suggests that people with a higher level of ability have a more differentiated personality structure because they have more freedom to develop their personality, and this, results in greater distinction between them. If this hypothesis is true, then certain outcomes can be predicted when analysing the interactions between measures of personality and measures of ability. First we can expect a lack of factorial invariance when assessing the structure of personality measures across different intelligence levels, insofar as fewer dimensions will be needed to describe the personality structure of less intelligent individuals. Second, high-ability individuals will show greater variability in personality measures than low-ability individuals. Finally, we can expect a lack of invariance of ability measures across levels of personality due to different relationships between ability measures across levels of different personality dimensions such as neuroticism.

The above predictions have generated a considerable amount of research over the last 30 years, but so far the evidence in favour of the PDH is inconsistent. With respect to the first issue mentioned, certain studies have detected a lack of invariance in personality measures across intelligence levels (Allik, Laidra, Realo, & Pullmann, 2004; Mclarnon & Carswell, 2013) or different correlations between personality measures across ability levels (Austin et al., 2002). Others, however, have reported that personality remains essentially invariant (De Fruyt, Aluja, García, Rolland, & Jung, 2006; Waiyavutti, Johnson, & Deary, 2012) or that the correlations between personality measures were equal across ability levels (Austin, Deary, & Gibson, 1997).

With regard to the second prediction, different authors have reported an increased variance of personality scores among high-ability individuals, but only for some of the personality dimensions analysed. Austin et al. (1997), for instance, reported this effect only for openness and neuroticism, while Harris, Vernon, and Jang (2005) found an increased variance for three of the twenty dimensions of personality and De Fruyt et al. (2006) found increased variance only for neuroticism and extraversion. However, other studies have reported no differences in any dimension (Allik et al., 2004; Escorial, García, Cuevas, & Juan-Espinosa, 2006).

* Corresponding author at: Universitat Rovira i Virgili, Departament de Psicologia, Crtra. Valls s/n, 43007 Tarragona, Spain.

Email address: andreu.vigil@urv.cat (A. Vigil-Colet)

https://doi.org/10.1016/j.paid.2018.03.034
0191-8869/

Finally, regarding the lack of invariance of ability measures across levels of personality, Austin et al. (1997) and Austin, Hofer, Deary, and Eber (2000) found that the correlation between two intelligence measures increased as neuroticism increased, while Austin et al. (2002) found that the correlation between fluid and crystallized intelligence increased with the level of neuroticism. Nevertheless, Escorial et al. (2006) found no difference between the eigenvalues of the *g* factor across different levels of personality dimensions, and Bonaccio and Reeve (2006) reported that the structure of cognitive abilities remained invariant across neuroticism levels.

Overall, the results so far summarized suggest that, despite the inconsistencies, there is partial support for the predictions deriving from the PDH. However, a clear and univocal rationale for the results obtained is still lacking. Although the PDH suggests that more intelligent individuals have more differentiated personalities, there are other explanations that may account for these results. Different authors have reported that personality measures have varying amounts of reliability depending upon individual levels of ability and education, with high-ability groups showing higher reliability (Allik et al., 2004; Austin et al., 1997; McFarland & Sparks, 1985). This increase in turn is expected to result in both higher variability (because of the increase in true variance) and higher score correlations (because they become less attenuated by measurement error). In the end these stronger correlations are expected to impact the factor structure of the measures analysed. Taking this alternative explanation into account, Austin et al. (1997) suggested that the results associated with the PDH may be reflecting (a) a true personality differentiation, (b) a simple effect of differential reliability, or (c) a mixture of the two.

Different explanations have been put forward regarding the differential reliability associated with ability levels (DRAAL). These explanations mainly derive from the fact that the process of answering items requires a considerable amount of cognitive processing, and have therefore focused on issues such as the difficulties that low-ability individuals have in understanding certain items, differences at verbal ability level, and the presence of a “highly calibrated ruler” in high-ability subjects that enables them to give more meaningful responses (Austin et al., 1997, 2000). Other authors have suggested that the DRAAL may ultimately be related to differences in response styles between high and low-ability groups, i.e. groups may show different levels of faking, self-enhancement and/or acquiescence which may be the cause of differential reliability (Allik et al., 2004; Austin et al., 2000). So far, however, there has been little research relating response styles and intelligence. De Fruyt et al. (2006) found no relationship between intelligence and self-enhancement. Meanwhile acquiescence has been related to intelligence and low levels of education (Meisenberg & Williams, 2008) and has been proved to have a considerable impact on the factor structures of personality insofar as the number of factors extracted in a personality test varies depending upon whether or not acquiescence effects are removed (Navarro-González, Lorenzo-Seva, & Vigil-Colet, 2016; Rammstedt & Farmer, 2013; Soto, John, Gosling, & Potter, 2008). These results may partly explain the effects described in the PDH because, if low ability individuals have higher levels of acquiescence and these effects are not removed, then different factor structures for these individuals are expected to arise.

Overall, as pointed out by Austin et al. (2000), the main problem is that it is difficult to disentangle which of the effects associated with the PDH are due to changes in personality structure across ability levels and which are due to other problems such as differential reliabilities on the sole basis of self-report results. At a group level, it is quite straightforward to assess whether the marginal reliability of personality scores is lower for the low-ability groups. However, assessing (a) the individual contributions to reliability, and (b) further potential relations to response bias indexes, verbal ability measures, etc. is not so simple. It is submitted here that a more finely-graded analysis that would enable points (a) and (b) above to be assessed would, in turn, enable the different explanations given for the DRAAL to be better investigated. This type of analysis, which is based on the concept of person reliability, is already feasible and is summarized below.

1.1. Person reliability

Conventional psychometric models for personality consider only a single parameter for each respondent: his/her level of the trait being measured. Implicitly, therefore, this modelling assumes that all individuals respond to the test with the same degree of consistency and accuracy. This view has been challenged for over 70 years (Coombs, 1948; Mosier, 1942) and the evidence in personality is also against it; some individuals respond to personality items with very high consistency, almost deterministically, whereas the responses of others are much more random. This differential degree of consistency has been labelled “person fluctuation”, “person reliability” or “person discrimination” (Ferrando, 2007, 2009). Person reliability is the term we shall use here.

Ferrando (2007, 2008, 2013) proposes a comprehensive item response theory (IRT) model for assessing person reliability under a variety of response formats. Essentially, the proposal consists of a series of extended conventional IRT models with an extra parameter that functions as an individual slope or discrimination index, and which models the degree of response consistency. This parameter is bounded below by zero and has no upper bound. Values near zero imply that the way the individual responds is almost random, i.e. totally insensitive to the normative item ordering, whereas very high values imply an almost deterministic, Guttman-type responding.

Following Tellegen (1988), Ferrando (2007, 2009, 2013) conceptualized person reliability as a relevant individual-differences dimension to partly explain the behaviour of the individual responding to a test. Furthermore, and also in line with previous proposals (Markus, 1977; Tellegen, 1988), Ferrando hypothesized that this dimension was related to the degree of clarity and strength with which the trait was organized in the individual. Recent empirical evidence suggests that this interpretation is tenable, with person reliability measures being indicators of traitedness (LaHuis, Barnes, Hakoyama, Blackmore, & Hartman, 2017). More generally, applied research results suggest that person reliability estimates have certain relevance in personality assessments. They are directly related to measures of conscientiousness and impulsivity (Austin, Deary, Gibson, McGregor, & Dent, 1998; Ferrando, 2007; Ferrando, 2009) and they have also been shown to function as moderator variables in validity assessments, in the sense that stronger relevant validity relations have been found for the most reliable individuals (Ferrando, 2015).

1.2. Aims of the study

The feasibility of obtaining reliability estimates at individual level might enable us to answer some of the questions discussed above. Thus if the low marginal (i.e. mean) reliability values found in low-ability groups is due to a poor understanding of the item content by low-ability individuals, then we can expect the person reliability estimates to be more closely related to measures of verbal ability than to measures of fluid or general intelligence. In the present research we shall use different personality measures, some of which have been developed using a method proposed by Ferrando, Lorenzo-Seva, and Chico (2009), enabling not only content but also acquiescence (AQ) and social desirability (SD) scores to be obtained for each individual. Hence relationships between intelligence measures and response bias measures can also be directly assessed. If, as authors such as Allik et al. (2004) and Austin et al. (2000) have suggested, response biases are responsible for the DRAAL results, then substantial relations between intelligence measures and response biases should be expected.

2. Method

2.1. Participants

The sample consisted of 532 student volunteers (252 men and 280 women) from 8 different high schools in the province of Tarragona, with ages ranging from 11 to 18 years old ($M = 14.75$ $SD = 2.1$). The same sample was the community sample used as a control in a study comparing the personality and abilities of juvenile offenders and community adolescents (Duran-Bonavila, Vigil-Colet, Cosi, & Morales-Vives, 2017).

2.2. Measures

The Indirect-Direct Aggression Questionnaire (IDAQ) (Ruiz-Pamies, Lorenzo-Seva, Morales-Vives, Cosi, & Vigil-Colet, 2014). This test gives scores for physical aggression (PA), verbal aggression (VA) and indirect aggression (IA) factors and an overall aggression score. The items were chosen from an initial pool selected by a panel of judges from the best of existing aggression measures refined after two studies using exploratory factor analysis. The test was developed using a method to control social desirability and acquiescence and has a considerable effect on the scores and factor structure of aggressive behaviour self-reports (Navarro-González et al., 2016; Vigil-Colet, Ruiz-Pamies, Anguiano-Carrasco, & Lorenzo-Seva, 2012). In addition to the three content factors, therefore, the test also gives scores for SD and AQ. The reliabilities of the factor score estimates derived from the IDAQ are appropriate: $r_{\theta\theta} = 0.83$, $r_{\theta\theta} = 0.77$ and $r_{\theta\theta} = 0.78$ for PA, VA and IA respectively.

The Barratt Impulsiveness Scale-11 for children (Chahin, Cosi, Lorenzo-Seva, & Vigil-Colet, 2010; Cosi, Vigil-Colet, Canals, & Lorenzo-Seva, 2008). This is a self-report questionnaire for assessing impulsivity that is specifically designed for children and adolescents. The test gives scores for motor impulsivity (MI), non-planning impulsivity (N-PI) and cognitive impulsivity (CI). MI is related to a lack of inhibition and delay, N-PI is related to planning abilities and CI to the tendency to make quick cognitive decisions.

The Psychological Maturity Assessment Scale (PSYMAS; Morales-Vives, Camps, & Lorenzo-Seva, 2013). This questionnaire consists of three scales: work-orientation (WO), self-reliance (SR) and identity (ID). It is made up of 25 items: seven items for each scale and four social desirability items. The reliability of the factor score estimates derived from the total scale is $r_{\theta\theta} = 0.82$, while the reliability of the subscale score estimates are $r_{\theta\theta} = 0.71$ for WO, $r_{\theta\theta} = 0.78$ for SR and $r_{\theta\theta} = 0.77$ for ID. In addition to these content factors, the test also gives factor score estimates for SD and AQ.

The Inventory of Callous Unemotional Traits (ICU; Frick, 2004). This is a questionnaire specifically designed to evaluate the precursors of psychopathy in youth populations. We use the Spanish adaptation developed by López-Romero, Gómez-Fraguela, and Romero (2015), which consists of 24 items with a 4-point response format (0 = never/almost never; 3 = always/almost always) with reliabilities of $\alpha = 0.76$, $\alpha = 0.82$ and $\alpha = 0.78$ for CA, UC and UE respectively.

Thurstone's Primary Mental Abilities Test (Cordero, Seisdedos, González, & de la Cruz, 1989). The subscales of Thurstone's test were verbal (PMA-V, $\alpha = 0.91$), spatial (PMA-S, $\alpha = 0.73$), numerical (PMA-N, $\alpha = 0.95$), reasoning (PMA-R, $\alpha = 0.92$) and word fluency (PMA-WF, $\alpha = 0.73$). The test comprises fluid and crystallized intelligence scales.

Raven's Progressive Matrices Test (Raven, 1996). This can be regarded as a measure of fluid intelligence free of cultural bias. The test has a reliability of $\alpha = 0.91$.

The information scale of the WAIS intelligence test for adults (Wechsler, 2003). This scale is an indicator of crystallized intelligence. The test has a reliability of $\alpha = 0.84$.

2.3. Procedure

School approval and parental written informed consent were obtained before participation in the study. Participation was voluntary and no incentives were given. About 96% of the participants who were invited to take part 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.

2.4. Data analysis

Individual scores on general intelligence were maximum likelihood (ML) factor score estimates obtained from the first principal factor based on all the intelligence measures. As for the personality measures, person reliability score estimates for each individual were obtained separately for each measure. These were ML estimates obtained as proposed in Ferrando (2013). We should note that standard errors and confidence intervals for these estimates can be obtained analytically and are important when the aim of the study is individual assessment. Person reliability estimates were not computed for each scale but only for the overall measures, since they need a minimum of 20 items to reach stability. However, it should be pointed out that all the personality measures used here enable an overall score to be used, which implies that the assumption of a general factor running through all the items is tenable. SD and AC were computed using the program Psychological Test Toolbox (Navarro-González, Vigil-Colet, Ferrando, & Lorenzo-Seva, in press). Finally, the relationships between intelligence and personality measures were analysed using product-moment correlations.

3. Results

Scores for the ability measures were factor-analysed using maximum-likelihood extraction. Sampling adequacy was good (KMO = 0.83). Only one factor had an eigenvalue greater than one, accounting for 43% of the variance. (See Table 1.)

Table 2 shows the relations between levels of ability and person reliability on the one hand, and the reliabilities of the personality measures on the other. In both 'g' factor and person reliability, high and low groups were obtained by using a median cut-off value. It can be seen that there is a slight tendency of high-ability individuals to show greater reliabilities, in the same sense as reported by Austin et al. (1997) and Allik et al. (2004). We observe a more noticeable ef-

Table 1
Loadings of ability scales on the first factor extracted by maximum likelihood used to estimate "g" factor scores.

Scale	Loading
WAIS (information)	0.601
PMA verbal	0.581
PMA spatial	0.524
PMA reasoning	0.698
PMA numeric	0.518
PMA word fluency	0.605
Raven	0.564

Table 2
Reliability of typical response measures for high and low-ability groups and person consistency groups. Cronbach's alpha for BIS and CFI and factorial reliability for IDAQ and PSYMAS. Between brackets 95% confidence interval for reliabilities.

	Ability		Person consistency	
	Low	High	Low	High
IDAQ	0.705 (0.652–0.753)	0.806 (0.772–0.838)	0.721 (0.671–0.767)	0.797 (0.761–0.830)
BIS	0.814 (0.780–0.844)	0.825 (0.794–0.854)	0.789 (0.751–0.823)	0.848 (0.820–0.873)
PSYMAS	0.725 (0.676–0.770)	0.750 (0.705–0.791)	0.693 (0.638–0.743)	0.786 (0.747–0.821)
ICU	0.709 (0.657–0.753)	0.739 (0.692–0.781)	0.705 (0.652–0.753)	0.747 (0.701–0.788)

fect when we compare the high and low person-reliability groups. In this case, as expected, the more inconsistent individuals have lower reliability coefficients.

Table 3 shows the product-moment correlations between ability measures and response bias as derived from the IDAQ and PSYMAS measures. As can be seen, SD seems to be mostly unrelated to ability measures in the sense reported by De Fruyt et al. (2006). However, there seems to be a noticeable inverse relationship between AC and ability, with about half of the relationships between ability measures and AC being significant, although the magnitude of the relationships is quite low. Despite the low magnitude of the correlation coefficients, this result may partly support the hypothesis that AC may be responsible for the lower reliability of the low-abilities group.

Person reliability was calculated from each of the personality measures. Table 4 shows the relationships between these person reliabilities and the various ability measures. There is a clear positive pattern of relationship between person reliability and ability found in all personality measures. This consistent result suggests that individuals with higher ability are more consistent when answering items in those questionnaires. The same table shows the mean correlation of personality measures and the different ability measures, ranking ability measures from the lowest correlation to the highest. It can be seen that the lowest correlations between person reliability and abilities were found for specific abilities such as verbal, word fluency, etc., while the highest correlations were found for overall measures of ability such as Raven's test or the "g" factor estimate. Finally, taking into account the potential relationships between AC and abilities as discussed above, we also computed the correlations shown in Table 4 controlling for AC only for IDAQ and PSYMAS (i.e. the tests that allow for this control). The magnitude of the correlation coefficients, however, remained mainly unaffected.

Table 3
Correlations of response bias in IDAQ and PSYMAS with ability measures.

	Acquiescence		Social desirability	
	IDAQ	PSYMAS	IDAQ	PSYMAS
WAIS (information)	**−0.157**	**−0.145**	**0.112**	−0.044
PMA verbal	**−0.160**	**−0.116**	0.077	−0.015
PMA spatial	0.040	−0.048	0.028	−0.035
PMA reasoning	−0.088	0.008	0.065	−0.004
PMA numeric	0.043	0.024	−0.035	−0.063
PMA word fluency	**−0.151**	**−0.123**	0.039	0.023
PMA total	**−0.145**	*−0.098*	0.055	−0.022
Raven	**−0.121**	*−0.087*	0.067	−0.034
G estimate	**−0.164**	**−0.114**	*0.091*	−0.023

p < .01 *p < .05.*

Table 4
Correlations between ability measures and person reliabilities computed from each personality measure ranked by the average correlation across personality measures.

	IDAQ	PSYMAS	BIS	ICU	Average
PMA verbal	**0.172**	**0.106**	0.04	**0.124**	**0.110**
PMA word fluency	**0.166**	**0.147**	0.068	0.08	**0.115**
PMA spatial	**0.143**	**0.147**	**0.104**	**0.154**	**0.137**
PMA numeric	**0.218**	**0.176**	**0.144**	**0.18**	**0.180**
WISC (information)	**0.284**	**0.202**	*0.095*	**0.144**	**0.181**
PMA reasoning	**0.202**	**0.167**	**0.166**	**0.218**	**0.188**
PMA total	**0.253**	**0.211**	**0.147**	**0.216**	**0.207**
Raven	**0.266**	**0.210**	**0.178**	**0.25**	**0.226**
G estimate	**0.308**	**0.245**	**0.171**	**0.246**	**0.243**

p < .01 *p < .05.*

4. Discussion

Since Austin et al. (1997) suggested that results relating to the PDH may be partly or totally due to a differential reliability effect, little research has been conducted to test the possible reasons for this effect. Three possible causes that may either individually or jointly explain the effect have been proposed: differences in social desirability and/or acquiescence related to ability, difficulties for low-ability individuals at the verbal processing level, and the possibility that high-ability individuals are more accurate when self-assessing.

The results reported in this article tend to discard some of these "a priori" explanations and basically suggest that high-ability individuals tend to provide self-assessments that are more accurate. As for the response bias conjectures, almost none of the SD measures showed relationships with either specific ability measures or general intelligence, confirming the results found by De Fruyt et al. (2006). However, a low but consistent relation between AC and intelligence measures was obtained. This is consistent with results reported by Meisenberg and Williams (2008) and suggests that AC may, at least in part, be a candidate for explaining the DRAAL.

Perhaps the most relevant results are those relating to the person reliability estimates. As Austin et al. (2000) pointed out, the main problem in testing the possible causes of DRAAL is the difficulty in determining the contribution of each individual to the overall reliability of a personality measure. The person reliability indexes used here, however, enable the response consistency to be estimated at the individual level and the relations between reliability and relevant individual-difference measures to be assessed. In this study the results point to a clear relationship between individual consistency and ability, with high-ability individuals showing greater person response consistency. Furthermore, these relations were found for both specific abilities and general intelligence. An inspection of their magnitude suggests that consistency is more closely related to general intelligence than to specific abilities. The lowest relationships were found for verbal abilities, which would seem to discard the conjecture that DRAAL is reflecting difficulties in understanding the items at this level. Overall, our results seem to favour the hypothesis by Austin et al. (1997) in that high-ability individuals are able to give more meaningful responses to the items and provide more accurate self-assessment. More generally, high-ability individuals appear to have higher levels of traitedness, i.e. it seems that personality traits are more meaningful for them, with the traits being represented with greater strength and clarity.

The results reported here therefore also have consequences in the domain of traitedness. Previous studies have shown that individual differences in traitedness were related to personality dimensions such as conscientiousness or impulsivity (Austin et al., 1998; Ferrando, 2007; Ferrando, 2009). Our results suggest that traitedness is related not only to personality but also to ability levels, which have consequences in the development of personality measures. Perhaps the most relevant consequence is the importance of avoiding biased samples such as university students when developing these types of measures because their higher intelligence, and therefore their higher person reliability, may falsely inflate the overall test reliability.

Although the results reported above may help to clarify the causes of DRAAL, they cannot answer the question as to whether the PDH is mainly due to the DRAAL or whether it is only partly explained by this phenomenon. Further research is needed to answer this. One possibility may be to apply the same approach used in this paper to bigger samples and then analyse the structure of personality at different levels, equating the individuals within each ability level in person reliability. The expected result, if the PDH can indeed be explained by differential reliability effects, is that we should find a lack of factorial invariance of personality across ability levels when individuals are not equated, and factorial invariance when the analysis is performed with individuals equated in person reliability.

Acknowledgements

This work was supported by a grant from the Spanish Ministry of Economy and Competitivity (PSI2017-82307-P).

References

Allik, J., Laidra, K., Realo, A., Pullmann, H., 2004. Personality development from 12 to 18 years of age: Changes in mean levels and structure of traits. European Journal of Personality 18, 445–462.
Austin, E.J., Deary, I.J., Gibson, G.J., 1997. Relationships between ability and personality: Three hypotheses tested. Intelligence 25, 49–70.
Austin, E.J., Deary, I.J., Gibson, G.J., McGregor, M.J., Dent, J.B., 1998. Individual response spread in self-report scales: Personality correlations and consequences. Personality and Individual Differences 24, 421–438.
Austin, E.J., Deary, I.J., Whiteman, M.C., Fowkes, F.G.R., Pedersen, N.L., Rabbitt, P., … McInnes, L., 2002. Relationships between ability and personality: Does intelligence contribute positively to personal and social adjustment?. Personality and Individual Differences 32, 1391–1411.
Austin, E.J., Hofer, S.M., Deary, I.J., Eber, H.W., 2000. Interactions between intelligence and personality: Results from two large samples. Personality and Individual Differences 29, 405–427.
Bonaccio, S., Reeve, C.L., 2006. Differentiation of cognitive abilities as a function of neuroticism level: A measurement equivalence/invariance analysis. Intelligence 34, 403–417.
Brand, C., Egan, V., Deary, I.J., 1994. Intelligence, personality, and society: "Constructivist" versus "essentialist" possibilities. In: Detterman, D.K. (Ed.), Current topics in human intelligence. Vol. 4, Ablex, New Jersey, pp. 29–42.
Chahin, N., Cosi, S., Lorenzo-Seva, U., Vigil-Colet, A., 2010. Stability of the factor structure of Barrat's Impulsivity Scales for children across cultures: A comparison of Spain and Colombia. Psicothema 22, 983–989.
Coombs, C.H., 1948. A rationale for the measurement of traits in individuals. Psychometrika 13, 59–68.
Cordero, A., Seisdedos, N., González, M., de la Cruz, V., 1989. PMA. Primary mental abilities. TEA Ediciones, Madrid.
Cosi, S., Vigil-Colet, A., Canals, J., Lorenzo-Seva, U., 2008. Psychometric properties of the Spanish adaptation of the Barrat mpulsiveness scale 11-a for children. Psychological Reports 103, 336–346.
De Fruyt, F., Aluja, A., García, L.F., Rolland, J.P., Jung, S.C., 2006. Positive presentation management and intelligence and the personality differentiation by intelligence hypothesis in job applicants. International Journal of Selection and Assessment 14, 101–112.
Duran-Bonavila, S., Vigil-Colet, A., Cosi, S., Morales-Vives, F., 2017. How individual and contextual factors affects antisocial and delinquent behaviors: A comparison between young offenders, adolescents at risk of social exclusion, and a community sample. Frontiers in Psychology 8, 1–12.
Escorial, S., García, L.F., Cuevas, L., Juan-Espinosa, M., 2006. Personality level on the big five and the structure of intelligence. Personality and Individual Differences 40, 909–917.
Eysenck, H.J., White, P.O., 1964. Personality and the measurement of intelligence. British Journal of Educational Psychology 24, 197–201.
Ferrando, P.J., 2007. A Pearson-type-VII item response model for assessing person fluctuation. Psychometrika 72, 25–41.
Ferrando, P.J., 2009. A graded response model for measuring person reliability. The British Journal of Mathematical and Statistical Psychology 62, 641–662.
Ferrando, P.J., 2013. A general linear framework for modeling continuous responses with error in persons and items. Methodology 9, 150–161.
Ferrando, P.J., 2015. The relevance of measuring individual discrimination in personality: A study with the test anxiety scale for children. Personality and Individual Differences 82, 253–257.
Ferrando, P.J., Lorenzo-Seva, U., Chico, E., 2009. A general factor-analytic procedure for assessing response bias in questionnaire measures. Structural Equation Modeling: A Multidisciplinary Journal 16, 364–381.
Frick, P.J., 2004. The inventory of callous-unemotional traits. University of New Orleans, New Orleans, LA, Unpublished rating scale.
Harris, J.A., Vernon, P.A., Jang, K.L., 2005. Testing the differentiation of personality by intelligence hypothesis. Personality and Individual Differences 38, 277–286.
LaHuis, D.M., Barnes, T., Hakoyama, S., Blackmore, C., Hartman, M.J., 2017. Measuring traitedness with person reliabilities parameters. Personality and Individual Differences 109, 111–116.
López-Romero, L., Gómez-Fraguela, J.A., Romero, E., 2015. Assessing callous-unemotional traits in a Spanish sample of institutionalized youths: The inventory of callous-unemotional traits. Journal of Psychopathology and Behavioral Assessment 37, 392–406.
Markus, H., 1977. Self-schemata and processing information about the self. Journal of Personality and Social Psychology 35, 63–78.
McFarland, S.G., Sparks, C.M., 1985. Age, education, and the internal consistency of personality scales. Journal of Personality and Social Psychology 49, 1692–1702.
Mclarnon, M.J.W., Carswell, J.J., 2013. The personality differentiation by intelligence hypothesis: A measurement invariance investigation. Personality and Individual Differences 54, 557–561.
Meisenberg, G., Williams, A., 2008. Are acquiescent and extreme response styles related to low intelligence and education?. Personality and Individual Differences 44, 1539–1550.
Mosier, C.I., 1942. Psychophysics and mental test theory II: The constant process. Psychological Review 48, 235–249.
Navarro-González, D., Lorenzo-Seva, U., Vigil-Colet, A., 2016. How response bias affects the factorial structure of personality self-reports. 28, 465–470.
Navarro-González, D., Vigil-Colet, A., Ferrando, P.J., Lorenzo-Seva, U., 2018. Psychological test toolbox: A new tool to compute factor analysis controlling response bias. Journal of Statistical Software(in press).
Rammstedt, B., Farmer, R.F., 2013. The impact of acquiescence on the evaluation of personality structure. Psychological Assessment 25, 1137–1145.
Raven, J.C., 1996. Matrices progresivas. Escalas CPM Color y SPM General. TEA Ediciones, Madrid.
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. Spanish Journal of Psychology 17.
Shure, G.H., Rogers, M.S., 1963. Personality factor stability for three ability levels. The Journal of Psychology 55, 445–456.
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.
Tellegen, A., 1988. The analysis of consistency in personality assessment. Journal of Personality 56, 622–663.
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.
Waiyavutti, C., Johnson, W., Deary, I.J., 2012. Do personality scale items function differently in people with high and low IQ?. Psychological Assessment 24, 545–555.
Wechsler, D., 2003. Wechsler intelligence scale for children-WISC-IV. Psychological Corporation, New York, NY.