Personality and Individual Differences xxx (2018) xxx-xxx

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com

Is individual reliability responsible for the differences in personality differentiation across ability levels?

Sergio Escorial[a], David Navarro-González[b], Pere Joan Ferrando[b], Andreu Vigil-Colet[b,] *

[a] Universidad Complutense de Madrid, Spain
[b] Universitat Rovira i Virgili, Research Center for Behavior Assessment, Spain

ARTICLE INFO

Article history:
Received 27 September 2018
Received in revised form 29 November 2018
Accepted 2 December 2018
Available online xxx

Keywords:
Person reliability
Personality differentiation hypotheses
Intelligence

ABSTRACT

Differential reliability associated with ability level has been proposed as a possible (partial or total) explanation of personality differentiation across ability levels: that is to say, high ability individuals have more differentiated personalities. Recent studies have shown that person reliability is related to general intelligence, which is necessary to support this hypothesis. The present study analysed the factorial structure and variability of a personality measure (Revised NEO Personality Inventory) in a sample of 7988 adult job applicants to test whether the effects of personality differentiation hypotheses are present and whether they are modified depending on person reliability. The results showed that the dimensionality reflected by the factorial structure of the test depended on ability levels, but this effect was not found when the analysis was performed using individuals with higher person reliability. As regards the variability of scores, as in previous research the ability effects were almost negligible. The results relating to the factorial structure were coherent with the hypothesis that personality differentiation across ability levels may be reflecting an issue of differential reliability.

1. Introduction

There have been a great many contradictory results since Brand, Egan, and Deary (1994) first put forward their personality differentiation hypothesis (PDH), which states that people with higher levels of ability have a more differentiated personality because they have more choices or freedom as regards its development, and this is reflected in a more differentiated personality structure (Austin, Deary, & Gibson, 1997). If this hypothesis is correct, then certain consequences can be expected at a psychometric level, these being mainly that a) the personality of low ability individuals will be explained by fewer factors than that of high ability individuals, and hence personality measures will exhibit a lack of invariance across ability levels, and b) high ability individuals will show greater variability in personality measures (however, see the discussion below). If the PDH is true it has important implications at various levels. First, from a personality theory perspective, research should aim to study the causes of this differentiation. Second, at a psychometric level, the dimensionality of personality measures must be adjusted as a function of the ability of the individuals assessed.

A great deal of research has aimed to test the predictions of the PDH, but the results have often been contradictory. Thus various studies have found a lack of invariance across ability levels (Allik, Laidra, Realo, & Pullmann, 2004; Austin et al., 2002; Mclarnon & Carswell, 2013) while others have not (De Fruyt, Aluja, García, Rolland, & Jung, 2006; Waiyavutti, Johnson, & Deary, 2012). A similar pattern of results has been found for the increased variance prediction, with some studies supporting this for only a few personality dimensions in high ability individuals (Austin et al., 1997; De Fruyt et al., 2006; Harris, Vernon, & Jang, 2005) and others finding no differences at all (Allik & McCrae, 2004; Escorial, García, Cuevas, & Juan-Espinosa, 2006). It should be noted that there are no differences in the origin or age of the samples, which may explain why in some cases the results support the PDH while in others the predicted results are not found.

An alternative to the PDH that may explain the results it predicted was proposed by Austin et al. (1997). These authors realized that personality scores in high ability individuals showed higher reliabilities than in low ability individuals and suggested that there was a differential reliability effect associated with ability levels (DRAAL). This effect has also been reported for different samples and personality measures (Allik & McCrae, 2004; McFarland & Sparks, 1985; Navarro-González, Ferrando, & Vigil-Colet, 2018).

The most important element of the DRAAL is that measures in low ability individuals have more error variance than in high ability individuals. As a result, their measures have less true variance, and this is reflected in (a) less variability in the true scores of the trait measured, and (b) lower inter-item correlations, which in turn impact the factorial structure of the measures. The DRAAL may also therefore explain the results forecast by the PDH. Nevertheless, as Austin et al. (1997); Austin, Hofer, Deary, and Eber (2000) pointed out, it is difficult to disentangle whether the PDH effects are due only to these differential reliability phenomena or rather to a mixture of differential reliability and a more differentiated personality at high ability levels.

* 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.2018.12.004
0191-8869/ © 2018.

This difficulty mainly arises from the fact that it is straightforward to assess score reliability at a group level, but far more difficult to assess individual contributions to group reliability. If this individual assessment were feasible, the specific impact of differential reliability on the effects suggested by the PDH could be more clearly assessed.

Recent psychometric developments in person reliability assessment may help to overcome this limitation. Classic psychometric models consider that individuals answer the test with the same degree of accuracy. However, there is a great deal of evidence against this view, with some individuals being highly consistent while others answer almost randomly. These differences in consistency are measured by person reliability indexes (Ferrando, 2007, 2009; Tellegen, 1988) and seem to be related to the clarity and strength of the organization of the trait in the individual, and they may thus also be (perhaps indirectly) indicators of traitedness (Ferrando, 2009; LaHuis, Barnes, Hakoyama, Blackmore, & Hartman, 2017).

As regards intelligence, Navarro-González et al. (2018) recently showed that person reliability measures are related to general intelligence and not to poor verbal abilities or response biases. Some authors have suggested that the DRAAL was a result of misunderstanding items or differences in response biases related to ability (Allik et al., 2004; Austin et al., 1997; Austin et al., 2000). Nevertheless, it seems that this is not the case. The study by Navarro-González et al. (2018) shows that person reliability is related to general intelligence, which would suggest that the DRAAL is mostly due to differences in general intelligence. Because of sample size limitations, however, Navarro-González et al. (2018) were unable to establish whether the DRAAL may totally or only partly explain the effects deriving from the PDH. This is because a large sample is needed in order to (a) test for different factorial structures across ability levels, and (b) in a second step, remove individuals with low person reliabilities and then test again to see whether the differences in factorial structures across ability levels remain the same, diminish or disappear. In the first of these three possibilities, the DRAAL would be independent of the PDH, in the second it would partly explain it, and in the third the PDH will be reflecting only the DRAAL and not a more differentiated personality.

Taking into account the results and limitations discussed above, the main purpose of the present research is to test whether the differences in the factorial structure of personality across different ability levels (i.e. a greater number of factors for high ability individuals) diminishes or even disappears when controlling for person reliability. As a secondary objective we will also assess possible differences in the score variances of different personality dimensions across ability levels and the effects that person reliability has on them. As far as this second aim is concerned, it should be remembered that various studies have failed to find any differences, or only in a limited number of personality factors (Allik et al., 2004; Austin et al., 1997; Escorial et al., 2006; Harris et al., 2005).

2. Method

2.1. Participants

The study comprised 7988 adults who applied for jobs in a company connected to the transport industry in Spain. The sample was 53.7% male and 46.3% female with ages ranging from 17 to 51 ($M = 25.75$; $SD = 4.86$). Although the large sample size causes statistically significant sex differences in age, the effect size suggests that these are negligible ($t_{(4539)} = 4.410$, $p < 0.001$, $d = 0.132$).

The data were selected from the archives corresponding to several selection processes carried out between 2006 and 2008. The sample included people with a wide variety of education levels, ages and reasons for applying for a job. In addition, the jobs they were applying for varied greatly, from low level (e.g. maintenance) to high level (e.g. management).

All applicants completed a selection battery that among other measures included a personality inventory and a cognitive ability test.

2.2. Measures

The applicants completed the Spanish version of the Revised NEO Personality Inventory (NEO-PI-R; Costa & McCrae, 1999), which is a well-known measure of the five-factor model personality domains – neuroticism (N), extraversion (E), openness to experience (O), agreeableness (A) and conscientiousness (C) – and their 30 facets (six for each of the five factors). The 240 items of the questionnaire are answered according to a 5-point Likert-type scale ranging from 0 (strongly disagree) to 4 (strongly agree). The reliability estimates for the scores on the five NEO PI–R domains in the present sample were in the range of 0.77 to 0.89 [(N, $\alpha = 0.89$), (E, $\alpha = 0.81$), (O, $\alpha = 0.77$), (A, $\alpha = 0.80$) and (C, $\alpha = 0.86$)]. A great deal of validity evidence has been found for the NEO-PI-R scores in a variety of applied contexts (i.e. Costa, 1996; Detrick & Chibnall, 2017; Salgado & Fruyt, 2017).

To assess cognitive ability the applicants completed Raven's Standard Progressive Matrices test (SPM) (Raven, 1996), which is the most widely used of all culture-reduced tests (Raven, Court, & Raven, 1998). The total score is a very good measure of *g*, the general factor of intelligence (Jensen, 1980). The estimated reliability of the total scores in this sample was $\alpha = 0.88$. In recent years, the popular alpha reliability estimate we use here has been subject to considerable criticism, and theoretically superior alternatives have been proposed (e.g. Cho, 2016). However, for scale scores that conform to the single-factor model and are based on a large number of items with reasonable discriminating power, which is the present case, the use of this estimate is perfectly justifiable.

There is also a great deal of scientific literature on the usefulness of Raven's SPM validity and a considerable amount of validity evidence has been found for its scores in a variety of applied contexts (i.e. Balboni, Naglieri, & Cubelli, 2010; Ones, Dilchert, Viswesvaran, & Salgado, 2017; Raven, 1999; Salgado et al., 2003).

2.3. Procedure

Data were collected from different personnel selection processes between 2006 and 2008. For each selection process, the participants went to the company's training centre in Madrid. All applicants were informed and gave written consent that their results could be used for research purposes, over and above simply for selection purposes, and were assured that they would remain anonymous (with no personal identification data). Only 2.2% of the applicants ($N = 179$) did not want to give their written consent, so the company did not provide the researchers with their data. Likewise, no other demographic characteristics were provided by the employing organization in order to maintain anonymity. Professional psychologists administered the tests to groups of between 50 and 75 participants. The measures were administered in Spanish and the professionals followed the standardized procedures. First the applicants were evaluated with Raven's SPM and then with the NEO-PI-R.

2.4. Data analysis

All the data were analysed using the Factor (Lorenzo-Seva & Ferrando, 2013), SPSS 25.0 and Matlab R2017a statistical analysis programs.

Person reliability score estimates for each individual were obtained separately for each scale of the NEO-PI-R. Essentially, the person reliability parameter is a slope parameter that models the discriminating power or consistency of the individual when responding to the test items. The estimates of this parameter were maximum-likelihood (ML) estimates obtained as proposed in Ferrando (2013). Person reliability estimates were not computed for each facet but only for the five scales, since the method used to obtain them requires a minimum of 20 items to achieve stable and plausible estimates for all individuals (Ferrando, 2009, 2013).

To assess the dimensionality of the factorial structure for the different samples we performed several exploratory factor analyses on the polychoric inter-item correlation matrices, using the optimal implementation of parallel analysis (Timmerman & Lorenzo-Seva, 2011). In this approach, the dimensionality of the corresponding data is decided on the basis of the percentage of explained common variance. To obtain these percentages, minimum rank factor analysis (Ten Berge & Kiers, 1991) is computed for the sample inter-item polychoric correlation matrix and for a set of random polychoric correlation matrices. The number of advised dimensions corresponds to the number of dimensions that in the sample correlation matrix accounts for a larger percentage of common variance than the percentage in the random correlation matrices. We would like to stress that the issue of determining the most plausible dimensionality of the data is the most basic one in factor analysis and must be addressed without imposing further constraints on the structure. So, the choice of the exploratory model is clearly justified in the study.

The samples for which dimensionality was assessed were obtained by splitting the whole sample into three ability groups of equal size (high, medium and low), and parallel analysis was run three times for each ability level: one for the whole sample, a second for individuals with person reliabilities above the sample median, and a third for those below. This strategy enabled us to test whether the PDH predicts more dimensions in high ability individuals, and whether this effect disappears when individuals are equated in person reliability.

To obtain a single reliability estimate for each individual, we computed the average of the five person reliabilities corresponding to the personality scales. Taking into account that the subsamples obtained by combining ability and reliability levels were of different sizes and that this could affect the accuracy of parallel analysis, all the subsamples were equated at the same sample size ($N = 1086$).

Finally, we assessed differences in personality score variance between ability levels using Levene's mean-based test for the homogeneity of variance.

Table 1
Descriptive statistics for personality, intelligence and person reliability.

	Variable	Mean	S.D.
	Raven	44.19	5.97
	Neuroticism	70.32	19.35
	Extraversion	123.07	15.27
	Openness	110.82	14.38
	Agreeableness	131.01	14.77
	Conscientiousness	137.43	15.79
Person reliabilities	Neuroticism	0.528	0.158
	Extraversion	0.519	0.142
	Openness	0.446	0.167
	Agreeableness	0.448	0.166
	Conscientiousness	0.497	0.157

Table 2
Reliability estimates (Cronbach's alpha) for scores on the personality dimensions across ability levels.

	Intelligence		
	Low	Medium	High
Neuroticism	0.873	0.885	0.900
Extraversion	0.786	0.812	0.824
Openness	0.725	0.778	0.798
Agreeableness	0.778	0.803	0.820
Conscientiousness	0.841	0.856	0.874

3. Results

Table 1 shows the descriptive statistics for the scores on personality dimensions and intelligence, as well as for person reliability measures. Personality and intelligence scores were similar to the ones obtained in the calibration samples.

Table 2 shows the marginal reliability estimates for the personality scale scores across ability levels. It can be seen that, for all scales, the reliability coefficients increase with individuals' ability levels, although the magnitude of these increases might be considered rather small.

With regard to the correlations between person reliability measures and intelligence, Table 3 shows that in all the scales the person reliability estimates showed a significant but small correlation with Raven's score. They also showed large correlations between each other, which suggests that individual reliability is notably consistent across different measures.

Table 4 shows the score variances for each personality dimension across ability and reliability levels. It can be seen that most of the variances showed no significant differences in Levene's test. Furthermore, they were greater for low-reliability than for high-ability individuals ($t_{(28)} = 5.57$ $p < 0.01$), with this difference showing a considerable effect size $d = 2.1$.

Finally, we carried out a parallel analysis to determine the number of factors to be retained for each ability level. The outcomes of these

Table 3
Correlation coefficients between intelligence and person reliability estimates.

	Raven	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
Raven	–	**0.144**	**0.143**	**0.080**	**0.140**	**0.180**
Neuroticism		–	**0.655**	**0.594**	**0.620**	**0.624**
Extraversion			–	**0.599**	**0.591**	**0.610**
Openness				–	**0.557**	**0.529**
Agreeableness					–	**0.587**
Conscientiousness						–

***p* < 0.01.**

Table 4
Variance of the scores in the personality dimensions across ability and reliability levels. In bold, significant heterocedasticity across ability levels.

Dimension	Ability level	All sample	Low reliability	High reliability
Neuroticism	Low	19.04	**20.98**	16.15
	Medium	19.05	**21.67**	16.20
	High	19.68	**22.74**	16.90
Extraversion	Low	15.00	16.12	12.87
	Medium	15.34	16.76	13.47
	High	15.38	17.35	13.30
Openness	Low	**13.50**	**14.51**	**11.98**
	Medium	**14.53**	**15.69**	**13.21**
	High	**15.08**	**16.42**	**13.80**
Agreeableness	Low	14.59	15.84	**12.69**
	Medium	14.70	16.10	**13.01**
	High	14.98	16.52	**13.54**
Conscientiousness	Low	15.74	16.88	13.40
	Medium	15.51	16.84	13.21
	High	16.02	17.69	14.09

$p < 0.01$.

analyses advised that 14, 15 and 16 factors for low, medium and high ability levels respectively should be retained. In a second step we selected two subsamples of individuals comprising 50% lower and 50% higher person reliabilities and ran a second parallel analysis for each ability level. This recommended that 13, 14 and 15 factors should be retained for the lower reliability subsamples, but 16 for all ability levels in the higher reliability subsamples. Fig. 1 shows the results of these analyses and the differences between ability levels in the decision zone.

4. Discussion

Many of the above results in connection with the DRAAL and the relationships between person reliability and intelligence are in line with previous research. We see the same positive relationship between reliability and ability as reported in previous studies (Allik et al., 2004; Austin, Deary, Gibson, McGregor, & Dent, 1998; McFarland & Sparks, 1985). Nevertheless, our study, like the one by Navarro-González et al. (2018), seems to show that this effect is smaller than those reported before. This may explain the difficulties experienced with some of the results in the domain of the PDH because, if they are due to the DRAAL, the small effects of ability on person reliability may in turn generate effects on the factorial structure of personality measures that are difficult to detect.

As regards the relationships between person reliability and intelligence, our study shows a small to moderate positive relationship between the two, reflecting in adults the same relationship previously found in adolescents (Navarro-González et al., 2018). The estimates of person reliability computed for the five personality factors showed high relationships between each other, ranging from $r = 0.529$ to $r = 0.655$. This result supports the construct validity of traitedness, insofar as it appears to be quite stable across different personality mea-

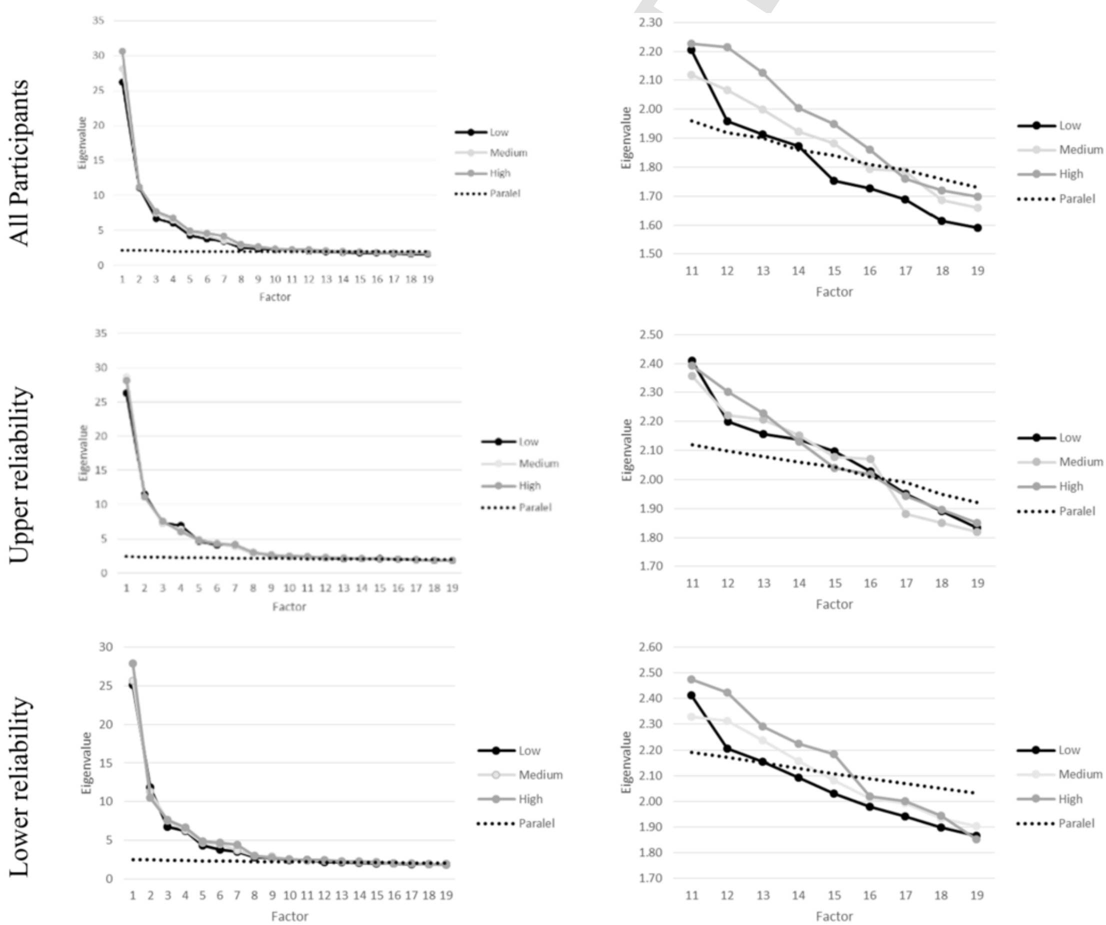

Fig. 1. Parallel analysis for all participants and for the 50% higher and lower person reliabilities by low, medium, and high intelligence group. The graphs at the right shows the decision zone magnified.

sures. In this respect, we used person reliability scores obtained for five personality dimensions to define high and low reliability individuals and then assessed their effects on the dimensionality of the first-order factors of the NEO-PI-R. The stability of the different estimates of person reliability allows us to assume that the effects of person reliability at second-order factor level of the NEO-PI-R may also generalize to lower-order facets.

As for the dimensionality at first-order factor level, the results are quite interesting. The PDH suggests that high ability individuals have a more differentiated personality, and as a result personality measures have more dimensions in high ability than in low ability individuals. Our results support this hypothesis, and parallel analysis advises 14, 15 and 16 dimensions for the three ability groups. Nevertheless, for low reliability individuals the number of dimensions recommended is 13, 14 and 15, showing that dimensionality is affected by person reliability. More interestingly, with high person-reliability individuals the analysis suggests the same number of dimensions (16) for all ability levels.

Our results are therefore coherent with the possibility that the PDH effects are reflecting a problem of different reliabilities linked to ability rather than a true effect of personality differentiation. Thus, at high person-reliability levels the PDH effects on dimensionality disappear. This kind of effect has implications for the development of personality measures because, depending on the kind of sample used, the reliability and dimensionality of the test may be over- (i.e. university student samples) or underestimated.

As far as the variance effects proposed by the PDH are concerned, the only personality dimension that showed this effect independently of individual person reliability was openness. Indeed many studies have failed to find the increases in variance in high ability individuals as suggested by the PDH, or have found them for only a few dimensions (Jüri Allik et al., 2004; Escorial et al., 2006; Harris et al., 2005). More interesting is the fact that variance seems to decrease depending on ability levels. However, Austin et al. (1997) argued that the increases in variance associated with ability levels might be due to increases in reliability, because when reliability increases the variability in the true scores also increases, and consequently so does the variability of the scores on the trait measured. Nevertheless, our results seem to show that the variance decreases at high reliability levels while it increases in the low reliability subsample. We believe that these apparently contradictory results can be explained by the interaction effects between the spread of item locations in the personality measure and the person reliability levels. To see this point more clearly, consider first a measure in which the item locations are all quite similar. In this case highly reliable individuals, who exhibit consistency in responding, are expected to respond similarly to all of the items, and their observed variances will therefore be small. Low-reliability individuals, however, would respond with a much higher error rate (i.e. more randomly) and their variances in this case would be expected to be larger. At the other end of the spectrum, when there is a wide spread of locations, the highly reliable individuals are expected to be sensitive to this and their true variance is expected to be high. Low-reliability individuals, however, who are far less sensitive to item locations, are expected to show lower variance here. Overall this is an issue that probably merits further research.

The present research has several limitations that need to be taken into account. The first is the fact that it uses a single measure of intelligence rather than various measures that might have allowed us to (a) obtain a “g” estimate, and (b) analyse the relationships of traitedness with different kinds of intelligence. Nevertheless, the study by Navarro-González et al. (2018) showed that traitedness is related more to general intelligence than to specific abilities and that the measurement most closely related to person reliability was the one used in the present study.

A second limitation is that many of the analyses were based on sub-groups obtained from an originally continuous variable (i.e. ability), a practice that always entails some loss of information. In theory, stronger results could be obtained by making more fine-grained splits, thus increasing the number of groups. However, this would have resulted in smaller groups, so the results would be potentially less stable and generalizable. We believe that for this data the three-split solution we used is a good compromise, but we also acknowledge that stronger analyses could be made in the future if large enough samples are available.

Finally a third limitation is the fact that the sample used was involved in a selection process, which may imply increased levels of social desirability and faking. Response biases may have effects on the factorial structure of personality measures, although the impact of social desirability on factorial structures is much lower that the impact of other response biases such as acquiescence (Morales-Vives, Lorenzo-Seva, & Vigil-Colet, 2017; Navarro-Gonzalez, Lorenzo-Seva, & Vigil-Colet, 2016).

Considering all the above, although the present research appears to show that the DRAAL may explain the effects of the PDH at factorial-structure level, in order for the present results to be confirmed, further research is needed using (a) more ability measures, (b) personality measures obtained in neutral testing situations, and (c) larger samples to assess dimensionality in high and low person-reliability individuals.

Acknowledgements

This work was supported by a grant from the Spanish Ministry of Economy and Competitiveness (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.

Allik, J., McCrae, R.R., 2004. Escapable conclusions: Toomela (2003) and the universality of trait structure. Journal of Personality and Social Psychology 87, 261–265.

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.

Balboni, G., Naglieri, J.A., Cubelli, R., 2010. Concurrent and predictive validity of the raven progressive matrices and the naglieri nonverbal ability test. Journal of Psychoeducational Assessment 28, 222–235.

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. (4). Ablex, New Jersey, pp. 29–42.

Cho, E., 2016. Making reliability reliable: A systematic approach to reliability coefficients. Organizational Research Methods 19, 651–682.

Costa, P.T., 1996. Work and personality: Use of the NEO-PI-R in industrial/organisational psychology. Applied Psychology 45, 225–241.

Costa, P.T., McCrae, R.R., 1999. Inventario de Personalidad NEO Revisado (NEO PI-R) e Inventario NEO Reducido de Cinco Factores (NEO FFI): Manual profesional [Revised NEO Personality Inventory (NEO PI–R) and Reduced NEO Five-Factor Inventory (NEO–FFI): Professional manual]. TEA Ediciones, Madrid.

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.

Detrick, P., Chibnall, J.T., 2017. A five-factor model inventory for use in screening police officer applicants: the revised NEO Personality Inventory (NEO PI-R). In: Mitchell, C.L., Dorian, E.H. (Eds.), Police psychology and its growing impact on modern law enforcement. IGI Global, Hershey, PA, pp. 79–92.
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.
Ferrando, P.J., 2007. A Pearson-type-VII item response model for assessing person fluctiation. 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.
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.
Jensen, A.R., 1980. Bias in mental testing. Free Press, New York.
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.
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, 497–498.
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.
Morales-Vives, F., Lorenzo-Seva, U., Vigil-Colet, A., 2017. How response biases affect the factor structure of big five personality questionnaires. Anales de Psicologia 33, 589–596.
Navarro-González, D., Ferrando, P.J., Vigil-Colet, A., 2018. Is general intelligence responsible for differences in individual reliability in personality measures?. Personality and Individual Differences 130 (1–5), .
Navarro-Gonzalez, D., Lorenzo-Seva, U., Vigil-Colet, A., 2016. How response bias affects the factorial structure of personality self-reports. Vol. 28, 465–470.
Ones, D.S., Dilchert, S., Viswesvaran, C., Salgado, J.F., 2017. Cognitive ability: Measurement and validity for employee selection. In: Farr, J.L., Tippings, N.T. (Eds.), Handbook of employee selection, 2nd ed. Taylor and Francis, Oxford, OX, pp. 251–276.
Raven, J., 1999. Psychometrics, cognitive ability, and occupational performance. In: Wechsler, S.M., Souza Lobo Guzzo, R. (Eds.), Avaliaçao Psicológica: Perspectiva Internacional. Casa do Psicólogo, Sao Paulo, pp. 299–344.
Raven, J.C., 1996. Manual del Test de Matrices Progresivas de Raven. Escalas CPM, SPM y APM. [Manual for Raven's proggressive matrices test. CPM, SPM and APM scales]. TEA Ediciones, Madrid.
Raven, J.C., Court, J.H., Raven, J., 1998. Manual for Raven's standard progressive matrices, 1998 ed. Oxford Psychologists Press, Oxford.
Salgado, J.F., Anderson, N., Moscoso, S., Bertua, C., De Fruyt, F., Rolland, J.P., 2003. A meta-analytic study of general mental ability validity for different occupations in the European community. Journal of Applied Psychology 88, 1068–1081.
Salgado, J. F., & De Fruyt, F. (2017). Personality in personnel selection. In Arne Evers, Neil Anderson & Olga Voskuijl (Eds.) The Blackwell handbook of personnel selection (pp. 174–198). Malden, MA: Blackwell Publishing.
Tellegen, A., 1988. The analysis of consistency in personality assessment. Journal of Personality 56, 621–663.
Ten Berge, J.M.F., Kiers, H.A.L., 1991. A numerical approach to the exact and the approximate minimum rank of a covariance matrix. Psychometrika 56, 309–315.
Timmerman, M.E., Lorenzo-Seva, U., 2011. Dimensionality assessment of ordered polytomous items with parallel analysis. Psychological Methods 16, 209–220.
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.