Faking Resistance of a Quasi-ipsative Forced-Choice Personality Inventory without Algebraic Dependence

Jesús F. Salgado and Mario Lado

University of Santiago de Compostela, Spain

ABSTRACT

This short research note reports on a study on the robustness of a quasi-ipsative forced-choice (FC) personality inventory for controlling the effects of faking. A sample of 126 active managers was randomly divided in three independent groups, with 42 individuals per group. We used an experimental three-group design in which the independent variable was the instructional set (faking, neutral, and honest), and the dependent variables were the scores in the Big Five personality dimensions. The results showed that the average effect sizes were .01, -.02, and 0 for the comparisons among faking-honest, faking-neutral, and neutral-honest groups. These findings showed that the quasi-ipsative FC format with algebraic non-dependence among the scales is a very robust way of controlling faking. We recommend practitioners to use this technology for making personnel selection decisions.

Resistencia al falseamiento de un inventario de personalidad cuasi-ipsativo de elección forzada

RESUMEN

Este breve nota de investigación se centra en un estudio sobre la fortaleza de un inventario de personalidad cuasi-ipsativo de elección forzosa para controlar los efectos del falseamiento. Se dividió aleatoriamente la muestra de 126 ejecutivos en activo en tres grupos independientes de 42 sujetos cada uno. Se utilizó un diseño experimental de tres grupos en el que la variable independiente era el conjunto de instrucciones (falseamiento, neutralidad u honestidad) y las variables independientes estaban constituidas por las puntuaciones de las dimensiones de personalidad de los cinco grandes. Los resultados mostraron que el tamaño medio del efecto era .01, -.02 y 0 para las comparaciones entre los grupos falseamiento-honestidad, falseamiento-neutralidad y neutralidad-honestidad. Estos resultados manifiestan que el formato cuasi-ipsativo de elección forzada sin dependencia algebraica entre las escalas es un modo muy sólido de controlar el falseamiento. Recomendamos a los profesionales que utilicen esta tecnología a la hora de tomar decisiones en selección de personal.

A serious problem of personality inventories in applied settings is their weaknesses to resist the effects of faking among respondents, particularly when single stimulus (SS) personality inventories are used for making decisions (Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, 2007). Faking has been defined as any voluntary distortion of responses to personality inventories (Birkeland, Manson, Kisamore, Brannick, & Smith, 2006; Grieve & Hayes, 2016; Levashina & Campion, 2007; Viswesvaran & Ones, 1999). Three particularly serious effects of faking are the shrinkage of reliability and validity of measures as a consequence of the artefactual homogenization of the scores (Salgado, 2016) and the alteration of the rank of the individuals. These effects are

most relevant in the area of personnel selection as the hiring decisions can be altered by faking. According to the psychometric theory of faking (Salgado, 2016), in order to control for faking, the method should control the average scores and simultaneously maintain the standard deviation of the distribution of scores.

Along the years, several methods have been developed to reduce the effects of faking on personality inventories. Among these methods, the forced-choice (FC) format was suggested as a relatively strong method to reduce the effect of faking (Nguyen & McDaniel, 2000; Jackson, Wroblewski, & Ashton 2000; Christiansen, Burns, & Montgomery, 2005).

Typically, FC format gives the respondent several alternatives along with instructions to choice the ones he or she most, or in some cases least, likes when it is applied to the respondent. The number of alternatives varies among FC inventories, but the more frequent case is using pairs, triads, or tetrads, which are paired in terms of their degree of social desirability. Contrarily to the most frequently used SS formats (such as Likert, agree-indecisive-disagree, and true-false formats in which the respondent rates each alternative), the individual has to make a choice among alternatives in case of the FC formats.

FC formats can result in several different scoring methods with specific statistical and psychometric particularities. Hicks (1970) suggested that three different types of FC measures can be distinguished: (a) purely ipsative measures, (b) quasi-ipsative (or partially ipsative) measures, and (c) normative FC measures. Recent meta-analytic findings showed that the quasi-ipsative FC inventories are more valid predictors of job performance than SS, ipsative, and normative FC personality inventories (Salgado, Anderson, & Tauriz, 2015).

Some degree of dependence can be generally found among the scales of quasi-ipsative FC inventories, which is a characteristic of ipsative scores (Heggestad, Morrison, Reeve, & McCloy, 2006). It is important to remark that the algebraic dependence affects the average correlation among the scales and, in the case of the purely ipsative questionnaires, the average correlation among m scales is bound below by $-1/(m-1)$ and above by $(m-4)/m$, whereas for algebraically independent measures the average correlation ranges between $-1/(m-1)$ and 1 (Gleser, 1972). A second effect of the algebraic dependence is that many of the typical statistical procedures used by researchers (e.g., ANOVA, factor analysis, multiple regression), which are based on the assumption that variables and samples are independent, cannot be logically used. A third effect of the algebraic dependence is that reliability coefficients are, generally speaking, smaller (Bartram, 1996). A fourth effect is that, because variances and covariances are generally reduced due to algebraic dependence among the variables, the multiple correlation between several algebraically dependent variables and an external criterion (e.g., job performance) is typically smaller that the multiple correlation found with the same variables if they are algebraically independent.

However, as Horn (1971) pointed out, quasi-ipsative scoring does not always introduce algebraic dependence. Over forty-five years ago, Horn (1971) suggested a method or strategy for avoiding algebraic dependence among personality scales but to the best of our knowledge it has not been implemented until now. Horn's strategy is based on the idea that if the items used to assess a specific scale are not used for assessing another scale, the two scales would be algebraically non-dependent, even though the nature of the score would remain quasi-ipsative.

Horn's method, by assuring independence among scales, also assures that observed item covariances are accurate estimates of true item covariances. Consequently, practically all the limitations of ipsative and quasi-ipsative FC measures are resolved if Horn's method is used for developing the FC questionnaires. For example, Cronbach's alpha is an appropriate coefficient for estimating the reliability of the scores from this measure, which is not true for typical quasi-ipsative FC questionnaires (Heggestad et al., 2006), and ANOVA and factor analyses can be conducted without limitations.

According to Hicks (1970; see also Bartram, 1996) the justification for the use of quasi-ipsative inventories would require them to fulfil three conditions: (a) that a significant faking degree exists in SS personality inventories; (b) that faking reduces the validity of SS inventories; and (c) that the quasi-ipsative format diminishes faking and increases criterion validity. With regard these three conditions, several meta-analytic studies have found that significant degrees of faking exist in personality assessment used for selection purposes (Birkeland et al., 2006; Levashina & Campion, 2007; Viswesvaran & Ones, 1999). In second place, Salgado (2016) demonstrated that faking reduces reliability and criterion validity of the personality

inventories. In third place, Salgado and his colleagues (Salgado, 2017; Salgado et al., 2015) demonstrated that the quasi-ipsative format increases criterion validity of personality inventories as compared with validity of SS and ipsative personality inventories. Therefore, the degree to which the quasi-ipsative format diminishes faking in the context of a real working sample (former studies used student samples) remains non-researched.

In this article, we report a study on the faking resistance of the QI5F (Salgado, 2014; 2017). This inventory combines three relevant characteristics: (1) it is based on the Five-Factor Model (FFM), (2) it uses a FC format, and (3) it was developed using the quasi-ipsative strategy suggested by Horn (1971) for producing scales for measuring that Big Five personality factors are independent, which avoids the metric dependence among personality scales and avoids ipsativization of scores. To the best of our knowledge, the QI5F is the only personality inventory that has used Horn's strategy of quasi-ipsativization until now and, consequently, no research has been developed to test faking resistance of quasi-ipsative FC personality inventories developed using Horn's strategy.

Method

Participants

The sample consisted of 126 managers working in small and medium size companies in Spain. There were 72 women (57.1%) and 54 men (42.9%). All the individuals participated voluntarily. Participants were recruited during 2018 through Linkedin social network of authors.

Measures

QI5F (Salgado, 2014). This 140-item personality questionnaire was designed to assess Big Five personality factors using a quasi-ipsative FC format. Each Big Five factor is assessed by 28 items. Each item consists of three alternatives, and the respondent must indicate which alternative he or she most likes and which alternative he or she least likes. In the current sample, internal consistency coefficients for emotional stability (ES), extraversion (E), openness to experience (OE), agreeableness (A), and conscientiousness (C) were .71, .73, .80, .66, and .80, respectively. The technical manual of the QI5F (Salgado, 2014) reports four-week test-retest reliabilities of .91, .90, .79, .65, and .72 for ES, E, O, A, and C, respectively. The technical manual also reports convergent validity evidence using an SS personality inventory. Observed correlations between scales measuring the same factors were .65, .73, .60, .70, and .70 for ES, E, O, A, and C, respectively. Correlations corrected for measurement error were .84, .91, .73, .98, and .85 for ES, E, O, A, and C, respectively. Exploratory factor analyses confirmed the five-factor structure of the QI5F (Salgado, 2014).

Experimental Design

We used a between-group experimental design in which the independent variable was the response condition. The dependent variables were their Big Five scores. Participants were randomly assigned to one of the three possible conditions: honest, neutral, and faking. Participants in the honest condition received instructions to respond the QI5F in the most honest manner. Participants in the faking condition were asked to respond in order to maximize their chances of being selected for a managerial position. Participants in the neutral condition did not receive specific instructions of being honest or of faking the personality inventory. Each group consisted of 42 individuals. In addition, the participant received the following instruction: "Assume you are responding the following inventory as a candidate for a top management position. The scores in this

inventory will serve for making a decision about your suitability for the job."

Procedure

The experiment lasted approximately 45 minutes to be completed in a single session. It was administered using the online software provided by Google Drive. Participants completed the questionnaire in the place they decided (e.g., job, home, and so on) and when they preferred (e.g., morning, afternoon, etc.). Participants were candid about the true objective of the experiment. The survey had three sections. The first section consisted of the instructions to participants and a brief socio-demographic questionnaire which asked about sex, age, education level, length of job experience, industry in which the participant worked, and number of supervised employees. The second section was the QI5F. The third section included several items of no interest for the current study.

Results

Table 1 reports the descriptive statistics for the Big Five in the three experimental conditions. As can be seen, both means and standard deviations are very similar across experimental conditions. These results suggest that faking effects were controlled by the quasi-ipsative FC format. In order to examine if differences among groups were statistically significant, we compared the honest condition with the faking condition, the honest condition with the neutral condition, and the faking condition with the neutral condition for the Big Five. We found that there was only a significant comparison out of fifteen comparisons. It was for extroversion when the honest and the neutral conditions were compared. In the neutral condition, individuals showed statistically significant higher scores ($F = 5.3$, $p = .024$, $df = 83$).

Table 1. Descriptive Statistics of the Big Five in the Three Experimental Conditions

	EE		EX		OE		A		C	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Honest	25.5	7.6	26.0	8.1	28.5	9.3	30.8	7.5	30.8	9.1
Neutral	24.0	7.0	29.9	7.3	27.6	7.7	30.9	5.6	28.3	9.0
Faking	24.6	6.8	28.2	7.3	28.2	9.2	30.1	6.4	28.8	7.8

Note. EE = emotional stability; EX = extroversion; OE = openness to experience; A = agreeableness; C = conscientiousness.

We calculated Cohen's *d* in order to estimate the magnitude of the faking effect. Table 2 reports these estimates. As can be seen, all effect sizes are very small when for the Big Five when we compared responses under faking instructions with honest and neutral conditions. The *d* values ranged from -0.09 to 0.28 (average *d* = 0.01) for comparisons among faking and honest conditions. The *d* values ranged from -0.24 to 0.13 (average *d* = -0.02) for comparisons among faking and neutral conditions. Finally, *d* values ranged from -0.27 to 0.50 (average *d* = 0) for comparisons among neutral and honest conditions. Based on these results we can conclude that differences are of no practical relevance and that the quasi-ipsative FC format with non-dependence among scores controlled the effects of faking in this study.

Table 2. Cohen's *d* for the Comparisons among the Three Experimental Conditions

	ES	EE	OE	A	C
F-H	-0.09	0.28	0.01	-0.09	-0.07
F-N	0.06	-0.24	0.13	-0.12	0.06
N-H	-0.14	0.50	-0.11	0.01	-0.27

Note. EE = emotional stability; EX = extroversion; OE = openness to experience; A = agreeableness; C = conscientiousness; H = honest condition; F = faking condition; N = neutral condition.

Discussion

Faking is one of the most serious concerns for the use of personality inventories as procedures for making personnel selection decisions because it might change the rank-order of applicants, which means that different applicants might be hired due to faking (Morgeson et al., 2007). For this reason, it is important that if personality inventories are used for selection purposes, they combine two features: (a) criterion validity and (b) faking resistance. Quasi-ipsative FC personality inventories showed to be the most valid predictors of job performance among all varieties of personality inventories (see, for instance, Lado & Alonso, 2017; Salgado, 2017; Salgado et al., 2015). However, the resistance to faking of quasi-ipsative FC inventories developed for assessing the FFM of personality remained relatively unexplored. Some previous research suggested that FC inventories showed some degree of robustness against faking, but the specific robustness of quasi-ipsative FC inventories remained unknown.

The findings of the current study suggest that faking can be confidently controlled, at least in the case of the personality inventory used here, which was developed following Horn's strategy and, therefore, can be characterized as a quasi-ipsative FC inventory with algebraic non-dependence among scales. Therefore, we can conclude that faking can be avoided.

Based on current results and previous meta-analytic findings on the prediction of job performance by the technology examined in this study, we strongly endorse quasi-ipsative FC personality inventories as the tool of election for making selection decisions. We also recommend practitioners to use this kind of personality inventories in their practical work in personnel selection. Finally, we suggest that additional studies be conducted with different samples and jobs in order to have a clear picture of this technique robustness against faking and also about the applicant perceptions (Aguado, Rico, Rubio, & Fernandez, 2016).

In summary, previous meta-analytic research has shown that quasi-ipsative FC personality inventories are the most valid personality predictors of job performance. Now, this study shows that the quasi-ipsative FC format with no algebraic dependence among scales is also robust against faking effects.

Conflict of Interest

The authors of this paper declare no conflict of interest. Both authors contributed equally, and the order is arbitrary.

References

Aguado, D., Rico, R., Rubio, V. J., & Fernández, L. (2016). Applicant reactions to social network web use in personnel selection. *Journal of Work and Organizational Psychology, 32*, 183-190. https://doi.org/10.1016/j.rpto.2016.09.001

Bartram, D. (1996). The relationship between ipsatized and normative measures of personality. *Journal of Occupational and Organizational Psychology, 69*, 25-39. https://doi.org/10.1111/j.2044-8325.1996.tb00597.x

Birkeland, S., Manson, T., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment, 14,* 317-335. https://doi.org/10.1111/j.1468-2389.2006.00354.x

Christiansen, N. D., Burns, G. N., & Montgomery, G. E. (2005). Reconsidering forced-choice item formats for applicant personality assessment. *Human Performance, 18*, 267-307. https://doi.org/10.1207/s15327043hup1803_4

Gleser, L. J. (1972). On bounds for the average correlation between subtest scores in ipsatively scored tests. *Educational and Psychological Measurement, 32*, 759-765. https://doi.org/10.1177/001316447203200314

Grieve, R., & Hayes, J. (2016). Employment testing online, offline, and over the phone: Implications for e-assessment. *Journal of Work and Organizational Psychology, 32*, 95-101. https://doi.org/10.1016/j.rpto.2016.04.001

Heggestad, E. D., Morrison, M., Reeve, C. L., & McCloy, R. A. (2006). Forced-choice assessment of personality for selection: Evaluating issues of normative assessment and faking resistance. *Journal of Applied Psychology, 91*, 9-24. https://doi.org/10.1037/0021-9010.91.1.9

Hicks, L. E. (1970). Some properties of ipsative, normative and forced-choice normative measures. *Psychological Bulletin, 74,* 167-184. https://doi.org/10.1037/h0029780

Horn, J. L. (1971). Motivation and dynamic calculus concepts from multivariate experiment. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology* (2nd printing, pp. 611-641). Chicago, Il: Tand McNally.

Jackson, D. N., Wroblewski, V. R., & Ashton, M. C. (2000). The impact of faking on employment tests: Does forced choice offer a solution? *Human Performance, 13*, 371-388. https://doi.org/10.1207/S15327043HUP1304_3

Lado, M., & Alonso, P. (2017). The Five-Factor Model and job performance in low complexity jobs: A quantitative synthesis. *Journal of Work and Organizational Psychology, 33*, 175-182. https://doi.org/10.1016/j.rpto.2017.07.004

Levashina, J., & Campion, M. A. (2007). Measuring faking in employment interview: Development and validation of an employment faking behavior scale. *Journal of Applied Psychology, 92*, 1638-1656. https://doi.org/10.1037/0021-9010.92.6.1638

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel contexts. *Personnel Psychology, 60*, 683-729. https://doi.org/10.1111/j.1744-6570.2007.00089.x

Nguyen, N. T., & McDaniel, M. A. (2000). *Faking and forced choice scales in applicant screening: A meta-analysis.* Paper presented at 15th Annual Meeting of the Society for Industrial and Organizational Psychology.

Salgado, J. F. (2014). *Technical manual for the QI5F.* Santiago de Compostela, Spain: Metis.

Salgado, J. F. (2016). A theoretical model of psychometric effects of faking on assessment procedures: Empirical findings and implications for personality at work. *International Journal of Selection and Assessment, 24*, 209-228. https://doi.org/10.1111/ijsa.12142

Salgado, J. F. (2017). Moderator effects of job complexity on the validity of forced-choice personality inventories for predicting job performance. *Journal of Work and Organizational Psychology, 33*, 229-238. https://doi.org/10.1016/j.rpto.2017.07.001

Salgado, J. F., Anderson, N. & Tauriz, G. (2015). The validity of ipsative and quasi-ipsative forced-choice personality inventories for different occupational groups: A comprehensive meta-analysis. *Journal of Occupational and Organizational Psychology, 88*, 797-834. https://doi.org/10.1111/joop.12098

Visweswaran, C., & Ones, D. S. (1999). Meta-analysis of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement, 60*, 197-210. https://doi.org/10.1177/00131649921969802