Psicologia: Reflexão e Crítica

RESEARCH

Open Access

Self-report personality tests and medical school selection

Isabel Lourinho[1,2*], Maria Amélia Ferreira[1] and Milton Severo[1,2]

Abstract

Background: There has been a growing interest on the assessment of personality when selecting medical students. However, how faking may affect its usefulness has been poorly addressed. Therefore, we aimed to assess the faking effect on self-report personality tests in the selection process of graduates to a medical school.

Methods: Sixty-seven graduates admitted as medical students completed the 60-item NEO Five-Factor Inventory (NEO-FFI) and the Marlowe-Crowne Social Desirability short-form scale at the stage of applying (baseline assessment) and after they had already been admitted (follow-up assessment). Reliability was assessed by the intra-class correlation coefficient and means of the personality traits compared by two paired sample t tests.

Results: At baseline assessment, the participants showed higher scores on the conscientiousness and lower scores on the neuroticism traits, respectively, 40.3 vs. 38.5 ($p = 0.026$) and 17.0 vs. 18.5 ($p = 0.089$). Also, at the follow-up assessment, the participants with low social desirability scored higher on the traits of openness to experience (-1.63 vs. 1.12, $p = 0.036$), conscientiousness (-3.09 vs. 0.03, $p = 0.022$), and neuroticism (3.88 vs. -0.69, $p = 0.012$).

Conclusions: Our study does not suggest the use of self-report personality assessment in medical student's selection as it can be faked particularly among applicants with low social desirability. Research is required to evaluate the faking effect on indirect personality assessment, namely through the tools that aim to select non-academic characteristics.

Keywords: Medical school selection, Personality assessment, Self-report personality tests, Faking effect, Graduate entry model, Longitudinal study

Background

Medical schools aim to select persons who besides becoming competent physicians in the future also express other competencies such as behaviour skills (Mahon et al. 2013). The importance of these competencies has been widespread in the medical education field. In our country, the "Medical Graduate in Portugal" is a document that defined 112 competences organised in five domains (knowledge, professional attitudes and behaviour, clinical skills and practical procedures, communication skills and general skills (Victorino et al. 2005).

Although for years medical students have mainly been selected based on academic achievement, there has been a recent and growing interest in the assessment based on personal attributes with particular emphasis on personality (Ferguson et al. 2003; Hojat et al. 2013; Lumsden et al. 2005). Some of the available medical selection tools which seek to choose other non-academic characteristics are interviews, mini multiple interviews (MMI) and situational judgement tests (Patterson et al. 2016).

In a historic perspective, admission to a medical school in Portugal has been dominated by young school-leavers, typically aged 18–19 years, and selection is based solely on their previous scholar achievement. However, since the 2007/2008 academic year, a graduate entry mode has been in force and each of the eight Portuguese medical school

* Correspondence: i.lourinho@med.up.pt
[1]Department of Medical Education and Simulation, Faculty of Medicine of the University of Porto, Alameda Prof. Hernâni Monteiro, Porto 4200-319, Portugal
[2]EPIUnit, Institute of Public Health, University of Porto, Porto, Portugal

Lourinho *et al. Psicologia: Reflexão e Crítica* (2016) 29:48

Page 2 of 5

has its own criteria for admitting graduates. For this quota, some medical schools have a written examination followed by a MMI as selection process while others have chosen the combination of previous achievement and admission interview.

Few studies that relate personality traits with the existent medical selection tools show that different selection processes call upon different personality traits (Azman et al. 2014; Griffin and Wilson 2012; Jerant et al. 2012; Schripsema et al. 2014; Schripsema et al. 2016). For instance, if MMI performance is associated with extraversion (Griffin and Wilson 2012; Jerant et al. 2012), it has been shown that the admitted medical students with higher top pre-university grades have higher conscientiousness scores when compared to the lottery-admitted group (Schripsema et al. 2014; Schripsema et al. 2016). Nevertheless, the research on the selected personality traits with regard to the application of direct personality assessment by self-report personality tests is practically non-existent.

It has already been shown that under guidance, individuals can fake personality tests (Viswesvaran and Ones 1999). Faking consists of the deliberate false presentation of one's self that may be favourable (fake good) or unfavourable (fake bad) (Hayes et al. 2012). Various theories exist to explain the faking behaviour that can occur due to personal characteristics of an individual or as a result of contextual variables (McFarland and Ryan 2000; Snell et al. 1999; Tett and Simonet 2011). The majority of the faking research is cross-sectional, and participants receive instructions either to answer honestly ("honest conditions") or to make a good impression or to make a specific impression of themselves ("faking conditions") (Shoss and Strube 2011; Tett et al. 2012; Topping and O'Gorman 1997). In the endeavour to identify and avoid the faking behaviour when answering to personality tests, some strategies were devised such as the use of social desirability scales or the use of response times (Donovan et al. 2003; Holden and Lambert 2015). However, how faking may affect personality assessment usefulness in the medical selection field has been poorly addressed. With this study, we aimed to assess the faking effect on self-report personality tests in a real-life medical school selection process.

Methods

Context
The Faculty of Medicine of University of Porto (FMUP) has had the highest access ratings for secondary school-leavers in Portugal in the last decades. This population has 245 places available per year while the graduate entry approach has 37 places.

In addition, FMUP's graduate admission scheme is a two-phase selection process that comprises previous achievement and an admission interview.

Flow of the participants
The baseline assessment occurred at the stage of applying to the FMUP (April 2012 and April 2013). Applicants were asked to complete a personality measure (as part of a larger battery of psychometric tests) and were informed that their collaboration would not have any consequences for the selection process and that the selection committee would not have access to their information.

The follow-up assessment occurred after the selection process was complete and after applicants had already been admitted and enrolled in the FMUP (October 2012 and October 2013). The participants voluntarily completed the same psychometric tests again, but this time, a social desirability scale was included.
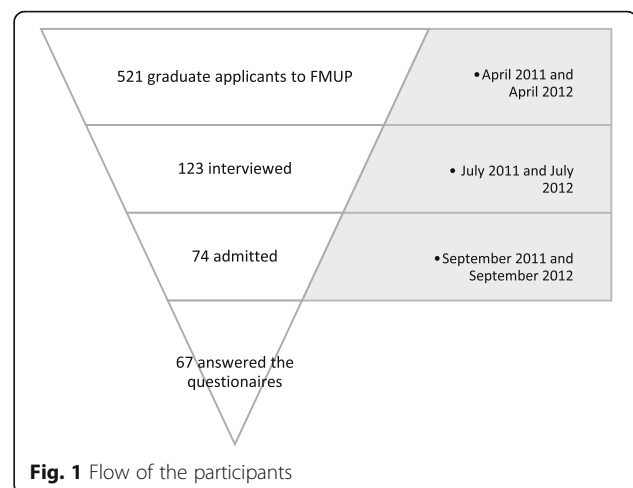
At baseline and on the follow-up assessments, questionnaires were sent by e-mail and the participants had 2 weeks to answer in their own time, unsupervised.

The study was reviewed and had the approval of the Ethics Committee of the Faculty of Medicine of the University of Porto, and the participants provided their written informed consent.

Although our participants were asked to respond honestly on both assessments and no manipulated instructions were given, we assume that our baseline and follow-up assessments correspond to the faking conditions and to the honest conditions settings, respectively.

Participants
The inclusion criteria for this study were (1) being admitted to FMUP; (2) having completed the questionnaires at baseline assessment; and (3) repeating the questionnaires at follow-up conditions. Of the 74 graduates admitted as medical students at the FMUP, 67 (90.5%) were eligible for this study (Fig. 1). Majority were women (80%) with a mean age of 27 years.

Fig. 1 Flow of the participants

Lourinho *et al. Psicologia: Reflexão e Crítica* (2016) 29:48

Page 3 of 5

Measures

Personality traits were assessed through the short version of the NEO Personality Inventory (NEO-PI-R), which is called the NEO Five-Factor Inventory (NEO-FFI). This 60-item, multiple-choice questionnaire evaluates five main dimensions of personality: openness to experience, conscientiousness, extraversion, agreeableness and neuroticism in a five-point Likert scale that ranges from 0 (strongly disagree) to 4 (strongly agree). Additionally, the NEO-FFI had already been validated for the Portuguese population (Magalhães et al. 2014).

The Marlowe-Crowne Social Desirability Scale is a measure of the influence of social desirability on test responses (Crowne and Marlowe 1960) that has been extensively used in personality research (Reynolds 1982). It consists of 33 items in a true-false response format. The items were chosen on the basis that they describe culturally approved behaviours that have a low incidence of occurrence and that response to items in the keyed or non-keyed direction have minimal implication of psychopathology (Crowne and Marlowe 1960). Short forms of this scale were developed and for this, study the 13-item form with a five-point Likert scale was used (Reynolds 1982).

Data analysis

The intra-class correlation coefficient (ICC) for a single measure was used to measure the reliability between baseline and follow-up assessments. Two paired sample t tests were used to compare the means of the personality traits between baseline and follow-up assessments. The participants were classified as high and low in social desirability according to the median of the final score on the Marlowe-Crowne Social Desirability Scale: the personality traits' mean difference between baseline and follow-up assessments was compared using two independent sample t tests. Pearson's correlation between personality traits and social desirability were estimated.

Results

The reliability between times of assessment was low for every personality trait (Table 1).

Significant differences were found for the conscientiousness and the neuroticism traits (Table 2) between times of assessment. More specifically, if the conscientiousness

Table 1 Reliability between times of assessment

Personality traits	ICC[a]
Openness to experience	0.254
Conscientiousness	0.459
Extraversion	0.318
Agreeableness	0.366
Neuroticism	0.411

[a]Intra-class correlation coefficient for a single measure

Table 2 Differences between times of assessment on personality traits.

Total

Personality traits	Baseline assessment Mean (SD)	Follow-up assessment Mean (SD)	Cohen's d	p
Openness to experience	33.1 (3.89)	32.9 (4.38)	−0.036	0.781
Conscientiousness	40.3 (5.06)	38.5 (5.48)	−0.297	0.026
Extraversion	34.2 (4.88)	33.6 (4.52)	−0.123	0.338
Agreeableness	36.76 (4.42)	36.87 (3.97)	−0.007	0.953
Neuroticism	17.0 (6.93)	18.5 (6.90)	0.216	0.088

trait decreased significantly at the follow-up assessment ($p = 0.044$), the neuroticism trait increased significantly at that time of assessment ($p = 0.035$). However, the mean scores of the openness to experience, extraversion and agreeableness traits did not show significant differences between times of assessment.

There were significant differences between times of assessment for the traits of openness to experience (−1.63 vs. 1.12, $p = 0.036$), conscientiousness (−3.09 vs. 0.03, $p = 0.022$) and neuroticism (3.88 vs. −0.69, $p = 0.012$), for the group with low desirability than for the high desirability group (Table 3).

Regarding the correlation between personality traits and social desirability, it was stronger for all personality traits at the follow-up assessment except for extraversion that was stronger at the baseline assessment (Table 4).

Discussion

This study has shown that participants faked on the personality traits of conscientiousness and neuroticism at the baseline assessment. Moreover, it was also found that participants with higher desirability levels were more honest at the follow-up conditions.

Table 3 Differences between times of assessment on personality traits stratified by desirability score (follow-up assessment-baseline assessment)

Personality traits	Low desirability Mean dif (SD)	High desirability Mean dif (SD)	Cohen's d	p
Openness to experience	−1.63 (4.99)	1.12 (4.81)	0.545	0.036
Conscientiousness	−3.09 (4.52)*	0.03 (5.55)	0.594	0.022
Extraversion	−0.33 (4.55)	−0.97 (6.07)	−0.121	0.638
Agreeableness	0.13 (5.26)	−0.18 (4.16)	−0.067	0.798
Neuroticism	3.88 (7.16)[a]	−0.69 (6.94)	−0.620	0.012

[a]Significant differences within the group
*Significant differences between follow-up and baseline assessments within this group
Italicized indicates that there were significant differences for the significant level of 0.005

Lourinho *et al. Psicologia: Reflexão e Crítica* (2016) 29:48

Page 4 of 5

Table 4 Pearson correlation between personality traits and social desirability

Personality traits	Baseline assessment	Follow-up assessment
Openness to experience	R (95CI %)	R (95CI %)
Conscientiousness	0.039 (−0.204, 0.278)	0.249 (−0.005, 0.473)
Extraversion	0.197 (−0.047, 0.419)	*0.506 (0.289, 0.673)*
Agreeableness	*0.348 (0.115, 0.544)*	*0.284 (0.038, 0.497)*
Neuroticism	*0.320 (0.085, 0.522)*	*0.450 (0.229, 0.628)*
Personality traits	*−0.294 (−0.500, −0.056)*	*−0.617 (−0.744, −0.431)*

Italicized indicates that there were significant differences for the significant level of 0.005

We believe that these results are due to the existence of some individual and contextual components of the faking process (Tett et al. 2012). It is possible that our participants could have not only the natural ability to fake (Tett and Simonet 2011) but also lower scores under honest conditions setting on the above-mentioned personality traits, having greater opportunity to fake at the faking conditions setting (Tett and Simonet 2011; Tett et al. 2012). We cannot ignore the fact that in our country like in so many others, there are much more applicants than available places in medicine (Patterson et al. 2016). In particular, the FMUP has had the highest access ratings for secondary school-leavers in Portugal over the last decades. Since our participants already hold a degree, and most of them wanted to study medicine since they were younger, we are in the position to assume that motivation was very high at this selection process. In addition, our participants came from a two-stage selection process in which stage 1 was based solely on previous achievement. It has been suggested that achievement is related to g (McManus et al. 2003) and also that cognitive ability (higher g) may facilitate the faking behaviour as brighter students seem to better identify which traits are job-relevant and therefore they fake accordingly (Tett et al. 2012).

However, the fact that had been already pre-selected by their previous achievement when they completed the personality measure is simultaneously a limitation of this study. As far as other study limitations, like other studies about medical student selection and personality testing, this is a single-centre study, which can lead to a selection bias because applicants usually apply to particular medical schools based on their personal preferences (Abbiati et al. 2016). In addition, unlike the personality measure, the social desirability scale was only administered at the follow-up assessment but we did not assess if desirability changed and if the association between social desirability and personality traits at baseline assessment was stronger or weaker when compared to the follow-up assessment. Furthermore, although important contributions to the faking research have been made with similar or even lower n (Shoss and Strube 2011; Robie et al. 2007), our small sample may have hidden

results for other personality traits. Finally, our sample only comprises graduate participants whose average age is closer to the thirties whereas the high-school entrants are closer to the twenties which calls the generalisation of these findings to high-school entrants into question.

Nevertheless, our study also has important strengths as it is one of the first studies on medical selection that assesses the faking effect on self-report personality tests. Also, it was carried out in a real selection process setting and not in an imaginary faking setting with manipulated faking instructions (Shoss and Strube 2011; Tett et al. 2012; Topping and O'Gorman 1997). The most used self-report personality test was applied (Hojat et al. 2013), one that is already validated for the Portuguese population (Magalhães et al. 2014). Moreover, it is a longitudinal study while most of the published research about faking is cross-sectional (Shoss and Strube 2011; Donovan et al. 2003).

Conclusions

In conclusion, our study shows some evidence that the incorporation of personality self-report tests in medical student's selection it is not advisable and should be confirmed in other settings with larger samples and using different personality instruments.

We agree that selection processes must be rigorous and publicly defensible (Prideaux et al. 2011) and that personality assessment may play an important role to the selection of medical students. But we also believe that faking it is a demanding and complex task for the combination of all the existent components (Tett et al. 2012).

If medical schools select skilled applicants who are able to present a desirable image on personality assessment, they will be in danger of admitting someone low in the future physician-relevant traits (Tett et al. 2012). Moreover, they may especially get away with negative behaviours as medical students and as physicians.

Research is required to evaluate the faking effect on indirect personality assessment, namely through the tools that aim to select non-academic characteristics.

Lourinho *et al. Psicologia: Reflexão e Crítica* (2016) 29:48

Page 5 of 5

References
Mahon, K. E., Henderson, M. K., & Kirch, D. G. (2013). Selecting tomorrow's physicians: the key to the future health care workforce. *Academic Medicine, 88*(12), 1806–11.

Victorino R, Jollie C, McKim J. Licenciado Médico em Portugal. Core graduates learning outcomes project..Faculdade de Medicina de Lisboa 2005.

Ferguson E, James D, O'Hehir F, Saunders A. A pilot study of the roles of personality, references and personal statements in relation to performance over the five years of a medical degree. British Medical Journal. 2003;326(7386):429-32.

Hojat, M., Erdmann, J. B., & Gonnella, J. S. (2013). Personality assessments and outcomes in medical education and the practice of medicine: AMEE Guide No. 79. *Medical Teacher, 35*(7), e1267–e301.

Lumsden, M. A., Bore, M., Millar, K., Jack, R., & Powis, D. (2005). Assessment of personal qualities in relation to admission to medical school. *Medical Education, 39*(3), 258–65.

Patterson, F., Knight, A., Dowell, J., Nicholson, S., Cousans, F., & Cleland, J. (2016). How effective are selection methods in medical education? A systematic review. *Medical Education, 50*(1), 36–60.

Azman, M. A.-Z., Yaacob, N. A., Yusoff, M. S. B., & Noor, S. H. (2014). Comparative study on medical student personality traits between interview and non-interview admission method in University Sains Malaysia. *Procedia - Social and Behavioral Sciences, 116*, 2281–5.

Griffin, B., & Wilson, I. (2012). Associations between the big five personality factors and multiple mini-interviews. *Advances in Health Sciences Education: Theory and Practice, 17*(3), 377–88.

Jerant, A., Griffin, E., Rainwater, J., Henderson, M., Sousa, F., Bertakis, K. D., et al. (2012). Does applicant personality influence multiple mini-interview performance and medical school acceptance offers? *Academic Medicine, 87*(9), 1250–9.

Schripsema, N. R., van Trigt, A. M., Borleffs, J. C., & Cohen-Schotanus, J. (2014). Selection and study performance: comparing three admission processes within one medical school. *Medical Education, 48*(12), 1201–10.

Schripsema, N. R., van Trigt, A. M., van der Wal, M. A., & Cohen-Schotanus, J. (2016). How different medical school selection processes call upon different personality characteristics. *PLoS One, 11*(3), e0150645.

Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: implications for personality measurement. *Educational and Psychological Measurement, 59*(2), 197–210.

Hayes TL. Matthias Ziegler, Carolyn MacCann, and Richard D. Roberts. (Eds.). New perspectives on faking in personality assessment. Oxford, UK: PB - Cambridge University Press, 2012, 364 pages, $59.95, hardcover. Personnel Psychology. 2013;66(3):798–801.

McFarland, L. A., & Ryan, A. M. (2000a). Variance in faking across noncognitive measures. *Journal of Applied Psychology, 85*(5), 812–21.

Snell, A. F., Sydell, E. J., & Lueke, S. B. (1999). Towards a theory of applicant faking: integrating studies of deception. *Human Resource Management Review, 9*(2), 219–42.

Tett, R. P., & Simonet, D. V. (2011). Faking in personality a: a "multisaturation" perspective on faking as performance. *Human Performance, 24*(4), 302–21.

Shoss, M. K., & Strube, M. J. (2011). How do you fake a personality test? An investigation of cognitive models of impression-managed responding. *Organizational Behavior and Human Decision Processes, 116*(1), 163–71.

Tett, R. P., Freund, K. A., Christiansen, N. D., Fox, K. E., & Coaster, J. (2012). Faking on self-report emotional intelligence and personality tests: effects of faking opportunity, cognitive ability, and job type. *Personality and Individual Differences, 52*(2), 195–201.

Topping, G. D., & O'Gorman, J. G. (1997). Effects of faking set on validity of the NEO-FFI. *Personality and Individual Differences, 23*(1), 117–24.

Donovan, J. J., Dwight, S. A., & Hurtz, G. M. (2003). An assessment of the prevalence, severity, and verifiability of entry-level applicant faking using the randomized response technique. *Human Performance, 16*(1), 81–106.

Holden, R. R., & Lambert, C. E. (2015). Response latencies are alive and well for identifying fakers on a self-report personality inventory: a reconsideration of van Hooft and Born (2012). *Behavior Research Methods, 47*(4), 1436–42.

Magalhães, E., Salgueira, A., Gonzalez, A.-J., Costa, J. J., Costa, M. J., Costa, P., et al. (2014). NEO-FFI: psychometric properties of a short personality inventory in Portuguese context. *Psicologia Reflexão e Crítica., 27*, 642–57.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349–54.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne social desirability scale. *Journal of Clinical Psychology, 38*(1), 119–25.

McManus, I., Smithers, E., Partridge, P., Keeling, A., & Fleming, P. (2003). A levels and intelligence as predictors of medical careers in UK doctors: 20 year prospective study. *British Medical Journal, 327*, 139–42.

Abbiati M, Baroffio A, Gerbase MW. Personal profile of medical students selected through a knowledge-based exam only: are we missing suitable students? Medical Education Online. 2016;21:10.3402/meo.v21.29705.

Robie, C., Brown, D. J., & Beaty, J. C. (2007). Do people fake on personality inventories? a verbal protocol analysis. *Journal of Business and Psychology, 21*(4), 489–509.

Prideaux, D., Roberts, C., Eva, K., Centeno, A., McCrorie, P., McManus, C., et al. (2011). Assessment for selection for the health care professions and specialty training: consensus statement and recommendations from the Ottawa 2010 Conference. *Medical Teacher, 33*(3), 215–23.