



## Evidências de validade na adaptação do Orpheus Business Personality Inventory no Brasil.

Jersica Assis Lozado<sup>1</sup>, John Neville Rust, Igor Gomes Menezes

<sup>1</sup> Universidade Federal da Bahia, Salvador, Bahia, Brasil.

### Palavras-Chave:

teste,  
personalidade,  
condições de trabalho,  
medida,  
validade

### Resumo

O propósito deste trabalho foi apresentar as evidências de validade do Orpheus Business Personality Inventory para a população brasileira. Cinco juízes contribuíram para a investigação das evidências de validade de conteúdo. Participaram deste estudo 938 indivíduos. Para investigar as evidências de validade de construto, utilizou-se a Full Information Factor Analysis. Os resultados deste estudo exploratório identificaram um conjunto de itens com cargas fatoriais baixas. O alfa de Cronbach e o coeficiente Omega foram calculados para medir a confiabilidade das escalas, seus valores variaram de 0,65 a 0,77 e de 0,53 a 0,77, respectivamente. Também foram avaliadas medidas de *misfit* usando o modelo Rasch, todos os itens apresentaram valores adequados. Estudos futuros são necessários para definir com clareza a sua adaptabilidade à população, entretanto foi possível demonstrar evidências de validade de conteúdo, estrutura interna e estabilidade teste-reteste para o OBPI, que poderá ser um importante instrumento de seleção nas organizações brasileiras.

Evidence of validity in the Orpheus Business Personality Inventory adaptation in Brazil.

### Keywords:

test,  
personality,  
work settings,  
measurement,  
validity

### Abstract

The purpose of this work was to present evidences of validity for Orpheus Business Personality Inventory to the Brazilian population. Five experts contributed to the investigation of evidence content validity. In total, 938 individuals responded the test. To examine the evidences of construct validity, a Full Information Factor Analysis was carried out. The results of this exploratory study identified a group of items with low factor loadings. Cronbach's alpha and omega coefficient were calculated to measure reliability in the scales, their values varied from .65 to .77, and from .53 to .77, respectively. Misfit measurements were also deployed using Rasch model all items presented adequate values. Future studies are needed to clearly define the adaptability to the population; however, the quality of the psychometric properties could be testified through evidences of content validity, internal structure and stability test-retest for OBPI, test-retest stability for OBPI, which may be an important selection test in Brazilian organizations.

Evidencias de validez en la adaptación del Orpheus Business Personality Inventory en Brasil.

### Palabras clave:

prueba,  
personalidad,  
condiciones de trabajo,  
medida,  
validez

### Resumen

El propósito de este trabajo fue presentar las evidencias de validez del Orpheus Business Personality Inventory para la población brasileña. Cinco jueces contribuyeron a la investigación de las evidencias de validez de contenido. Participaron de este estudio 938 individuos. Los datos fueron recolectados en la plataforma Concerto. Para investigar las evidencias de validez de construto, se utilizó el Full Information Factor Analysis. Los resultados de este estudio exploratorio identificaron un conjunto de ítems con cargas factoriales bajas. El alfa de Cronbach y el coeficiente Omega fueron calculados para medir la confiabilidad de las escalas, sus valores variaron de 0,65 a 0,77 y de 0,53 a 0,77, respectivamente. También se evaluaron medidas de *misfit* usando el modelo Rasch, todos los elementos presentaron valores adecuados. Los estudios futuros son necesarios para definir con claridad su adaptabilidad a la población, sin embargo fue posible demostrar evidencias de validez de contenido, estructura interna y estabilidad test-reteste para el OBPI, que podrá ser un importante instrumento de selección en las organizaciones brasileñas.

<sup>1</sup> Endereço para correspondência:

Instituto de Psicologia, Rua Aristides Novis, 197, Estrada de São Lázaro, CEP 40210-730, Salvador, Bahia. E-mail: jersicaassis@gmail.com

Como citar este artigo:

Lozado, J. A., Rust, J. N., & Menezes, I. G. (2019). Integração profissional de pessoas com deficiência visual: Evidências de validade na adaptação do Orpheus Business Personality Inventory no Brasil. *Revista Psicologia: Organizações e Trabalho*, 19(2), 624-630. doi: 10.17652/rpot/2019.2.15930

A internacionalização de construtos psicológicos é um processo gradativo que vem expandindo conhecimentos teóricos nos mais diversos campos de pesquisa. Van de Vijver (2013) descreve essa tendência quando diz que, tendo os anos 70 como ponto de partida, o número de publicações a respeito de estudos transculturais é quinze vezes maior atualmente. Salientando a importância destes estudos para o desenvolvimento da Psicologia, o autor ainda define que uma das etapas mais importantes desse processo, é o uso de instrumentos adaptados. A adaptação de um instrumento transcende barreiras geográficas ou culturais, possibilitando a comparação das características dos indivíduos inseridos em diferentes contextos, uma vez que existe equidade na avaliação através do uso da mesma medida (Hambleton, 2011; He & Van de Vijver, 2012). Considerando o uso de métodos transculturais, o modelo *Big Five* de personalidade, por exemplo, foi desenvolvido de maneira a acomodar certas diferenças culturais e se mostrou eficiente em descrever traços de personalidade em indivíduos das mais diversas culturas, sejam estas individualistas ou coletivistas (McCrae & John, 1998).

Alguns estudos dentro do campo organizacional utilizam o modelo *Big Five*, como Templer (2012) que demonstrou como as dimensões do modelo afetam a satisfação, assim como Syed, Saeed e Farrukh (2015) que fez o mesmo acerca do comprometimento. Wu (2016), por exemplo, descreveu como o aumento de estresse no trabalho está relacionado a um aumento no traço neuroticismo e em uma diminuição nos traços de extroversão e conscienciosidade. Choi, Oh e Colbert (2015) encontraram relações entre o traço de amabilidade e o comprometimento afetivo mesmo em diferentes culturas (coletivistas e individualistas). Mais recentemente, Cortez (2017) apresentou evidências de que os fatores do *Big Five* são capazes de prever até a atitude empreendedora, através de uma amostra de estudantes do estado de Minas Gerais. O uso do modelo *Big Five* permite que as organizações entendam e predigam não apenas o processo de desenvolvimento profissional ou do comprometimento dentre seus membros, mas também aspectos como desvios das normas do ambiente de trabalho (Grijalva & Newman, 2015), desde que tenham acesso aos diferentes níveis dos traços indicados como relevantes.

Outros processos organizacionais, como o de seleção de pessoas, poderiam ser realizados de maneira mais estratégica, tendo em vista a qualidade das informações que os profissionais teriam acesso através do uso de medidas psicométricas baseadas no *Big Five*. Contudo, apesar dos benefícios para os estudos organizacionais, são poucos os instrumentos que se propõem a medir a personalidade com itens contextualizados para o ambiente de trabalho. Ainda que a personalidade seja uma característica estável (McCrae & John, 1998), esta pode sofrer influências do contexto em que o indivíduo está inserido como descrito por Comensoli e MacCan (2013) e Wu (2016). Dessa maneira, um entendimento completo da personalidade deve levar em conta que o contexto profissional é complexo e envolve um conjunto de valores, crenças, normas e rotinas que contribuem para a definição da identidade dentro da organização (Fernandes & Zanelli, 2006); que por sua vez, afetam as expectativas, motivações e atitudes dos funcionários de maneira que os elementos ditos contextuais da personalidade são também influenciados.

Dentre as poucas medidas contextualizadas, pode-se citar o *Hogan Assessment System* (Hogan & Hogan, 2007) e testes que, apesar de não terem sido adaptados para a cultura brasileira, são utilizados por organizações em processos de seleção. Para além da utilização do modelo *Big Five*, o teste *Myers-Briggs Type Indicator* (MBTI) é bastante popular, principalmente por oferecer perfis

da personalidade baseados em uma tipologia. Contudo, Cooper, Knotts, McCord e Johnson (2013) identificam alguns problemas no uso dessa medida, como inconsistência nos resultados e a limitação dos descritores associados aos traços de personalidade apresentados. A necessidade de uma medida que seja contextualizada para o ambiente de trabalho, baseada no modelo *Big Five* e que apresente propriedades psicométricas sólidas, determinam a importância de um teste como o *Orpheus Business Personality Inventory* (OBPI).

O OBPI foi desenvolvido por Rust (2001) especialmente para o ambiente organizacional e pode ser usado sob diversos propósitos, como seleção, promoção e desenvolvimento de colaboradores. Os itens do OBPI são oriundos de um longo processo de testagem, sendo desenvolvido a partir dos itens de 12 instrumentos diferentes selecionados ao longo de 10 anos. Composto por 190 itens, o teste foi produzido para alcançar a solução de determinados problemas frequentemente encontrados nos instrumentos voltados a medir a personalidade, como a deseabilidade social. Os itens são distribuídos em 16 escalas: cinco escalas primárias, sete escalas secundárias e quatro escalas de checagem (Rust, 2001). Cada item do teste contém quatro categorias de resposta sob o formato de uma escala do tipo *Likert*: 1) Discordo totalmente, 2) Discordo, 3) Concordo, e 4) Concordo totalmente, representando o grau de concordância em relação a afirmação contida no item.

As escalas primárias, baseadas no modelo *Big Five* de personalidade, são: Companheirismo (*Fellowship*) com 22 itens, Autoridade (*Authority*) com 22 itens, Conformidade (*Conformity*) com 23 itens, Emoção (*Emotion*) com 18 itens e Detalhe (*Detail*) com 20 itens. As escalas secundárias, por sua vez, foram baseadas nos traços de Integridade apresentados no Teste de Integridade *Giotto* (Rust, 1999), são elas: Proficiência (*Proficiency*) com 14 itens, Orientação para o trabalho (*Work-orientation*) com 18 itens, Paciência (*Patience*) com 14 itens, Imparcialidade (*Fair-mindedness*) com 18 itens, Lealdade (*Loyalty*) com 14 itens, Transparência (*Disclosure*) com 16 itens e Iniciativa (*Initiative*) também com 16 itens. Por fim, as escalas desenvolvidas para checagem de respostas contaminadas são: Dissimulação (*Dissimulation*), Ambivalência (*Ambivalence*), Desânimo (*Despondency*) e Falta de atenção (*Inattention*).

A diferença do OBPI para o *Hogan Assessment System* é que o último está voltado a medir a personalidade através da reputação, tendo como plano de fundo a Teoria Socioanalítica (Hogan e Hogan, 2007). O teste OBPI, por sua vez, é uma medida de personalidade com escopo mais amplo, pois inclui escalas que também oferecem uma medida de integridade, englobando construtos como responsabilidade, ética de trabalho, sociabilidade, comportamentos agressivos e comportamentos desonestos (Rust, 1999). A avaliação da integridade em um teste voltado para o ambiente do trabalho, é importante devido as evidências de poder preditivo entre testes de integridade e variáveis como desobediência a regras, absenteísmo, desempenho no trabalho, comportamento contraproducente e *turnover* (Catano, O'Keefe, Francis & Owens, 2018; Nei, Foster, Ness & Nei, 2018).

Em adaptações para outros contextos culturais, o OBPI demonstrou fortes evidências de confiabilidade (Daouk-Oyry, 2008), indicando a estabilidade dos seus itens em diferentes idiomas. Além das suas propriedades psicométricas desejáveis, este é um teste computadorizado, que possui a vantagem de reduzir erros de correção e aumentar eficácia na produção dos relatórios, bem como suprir a lacuna de testes para seleção de pessoas que sejam modernos, confiáveis e baseados em sistemas automatizados. Adicionalmente, o OBPI oferece resultados que indicam não só as predominâncias dos traços de personalidade, como também aspectos associados a integridade do avaliado. Tendo em vista estas características e a

necessidade de testes específicos para o contexto organizacional, o propósito deste estudo foi apresentar evidências de validade do *Orpheus Business Personality Inventory* para a população brasileira.

## Método

### Adaptação do Instrumento

O processo de adaptação contou com fases de tradução direta e tradução reversa. A versão original foi traduzida por dois tradutores independentes, a partir disso, uma versão sintetizada foi desenvolvida. Esta versão foi encaminhada para a tradução reversa, que foi realizada por um nativo da língua inglesa com proficiência em português. A versão resultante desta tradução, foi comparada à versão original pelo autor do OBPI e as modificações sugeridas foram incluídas em uma nova síntese da versão em Português. Esta última versão foi avaliada na análise de juízes, e finalmente, quando as requeridas modificações foram realizadas, a versão final foi apresentada.

### Participantes

O critério de inclusão na amostra foi ter o Ensino Médio completo. Os participantes não precisavam estar empregados para responder ao teste, e quando este era o caso, eles foram instruídos a responder ao OBPI considerando suas atividades laborais anteriores ou suas atividades na universidade ou colégio. O processo de amostragem utilizado foi o *snowball sampling*, que é um procedimento não probabilístico, no qual é solicitado a cada participante que convide um amigo ou compartilhe a pesquisa na sua rede social (Johnson, 2014). A decisão de convidar alguém ou compartilhar a pesquisa foi completamente voluntária.

Ao fim, foram coletados 938 casos válidos. A maioria dos participantes (64,29%) era do sexo feminino e a média de idade foi de 31,88 anos, com desvio padrão de 10,39 anos. Em relação ao nível educacional, 41,04% relataram ser pós-graduados, aqueles com ensino superior completo foram 34,54% da amostra, e finalmente, os com ensino médio completo somaram 24,41%. Quanto ao estado civil, a maior parte dos participantes relataram ser solteiros (61,30%), ao passo que aqueles que são casados ou estão em uma união estável compuseram 33,05% da amostra; apenas 3 (0,32%) participantes relataram ser viúvos, e 5,33% são divorciados. Quanto à atuação profissional, mais da metade (55%) estavam enquadrados em atividades relacionadas às ciências ou artes (307 participantes) e serviços administrativos (207 participantes). Dentre os respondentes, apenas 9% declararam estar desempregados.

Todos os participantes foram solicitados a responder ao reteste, do total de 938 que responderam a primeira coleta, 99 se dispuseram refazer o teste, cerca de 10,5% da amostra total. Destes, 63% eram do sexo feminino. A média de idade do reteste foi de 32,68, com desvio padrão de 11,20 anos. Quanto ao nível educacional, 44,40% declararam ter Pós-Graduação, 31,30% relataram ter nível superior, e 24 (24,20%) participantes possuíam o Ensino Médio completo. Quanto ao Estado Civil, 60,60% eram solteiros, enquanto que os casados e divorciados eram 29,30% e 10,10%, respectivamente.

### Coleta de dados

Esta pesquisa foi submetida e aprovada pelo Comitê de Ética e todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido. O convite para participação na pesquisa foi enviado por diferentes plataformas virtuais. Por e-mail, o convite foi enviado utilizando os contatos dos pesquisadores. Para o Facebook,

que foi o principal meio de coleta, foram definidas algumas características fixas para o público alvo através de uma ferramenta disponibilizada por esta rede para divulgar as publicações. Ao clicar no convite da pesquisa, os participantes eram direcionados a plataforma Concerto, onde poderiam responder ao teste. Desenvolvida por Scalise e Allen (2015), esta plataforma é um software livre que permite que o usuário desenvolva testagem online e testagem adaptativa computadorizada.

Para incentivar as respostas dos participantes, e-mails periódicos foram enviados. Ao fim, foi solicitado para que os respondentes deixassem o seu e-mail para participação na retestagem. Aqueles que deixaram o endereço de e-mail foram contatados quatro meses após o fim da coleta inicial.

### Análise dos Dados

Para buscar evidências de validade baseadas no conteúdo, cinco juízes participaram do processo de avaliação da última versão do OBPI em Português. Os juízes responderam a um instrumento online em que os itens do OBPI foram avaliados individualmente. Para cada item, os juízes comentaram a respeito da estrutura, incluindo sugestões concernentes a gramática, formato ou outro problema. Adicionalmente, os juízes foram solicitados a avaliar a adaptação do item considerando duas questões retiradas do instrumento *Item Translation and Adaptation Review Form* (Hambleton & Zeniski, 2011, p. 71).

Os dados provenientes da análise dos juízes foram classificados em dois tipos: qualitativos (comentários) e quantitativos (respostas ao questionário). Os dados quantitativos foram examinados através da Análise de Concordância entre juízes (Gisev, Bell & Chen, 2013). Os dados qualitativos foram cuidadosamente examinados pelos pesquisadores. Cada comentário foi considerado individualmente e classificado em umas das três categorias a seguir: formato do item, utilização de versos ou ditados, gramática e estrutura.

Para estudar a estrutura interna do OBPI, uma *Full Information Factor Analysis* (FIFA) foi realizada através do uso das técnicas da Teoria de Resposta ao Item Multidimensional (MIRT). Esta técnica foi escolhida porque diferentemente da Análise Fatorial Exploratória, a FIFA não é primordialmente uma técnica de redução de dados. Isso significa que para a FIFA a melhor solução não é necessariamente aquela que explica mais variância nos dados com o menor número de fatores, mas sim aquela que melhor modela a interação entre o respondente e os itens (Reckase, 2009). Portanto, usar MIRT para o estudo da estrutura fatorial do OBPI foi a escolha mais adequada, considerando que não havia interesse em reduzir as variáveis (itens) estudadas.

Cada uma das escalas foi avaliada individualmente quanto às suas propriedades psicométricas. A razão para isso é que para o OBPI as escalas devem ser tratadas de maneira independente (Rust, 2001), e analisando-as separadamente permite que haja a certificação de que os itens não medem nada além do traço de Personalidade para o qual foram designados. Para as escalas avaliadas, uma FIFA foi realizada para extração de uma única dimensão com a utilização do algoritmo *Expectation Maximization* (EM) de estimação de parâmetros. Este algoritmo alterna entre a estimação das probabilidades de resposta dado o modelo estabelecido (conhecido como passo E), e a re-estimação dos parâmetros do modelo utilizando as informações do passo anterior, o passo M (Do & Batzoglou, 2008).

A presença de *misfit* também foi verificada através da análise de resíduos baseada no modelo Rasch. Bond e Fox (2007) indicam que os índices de *infit* e *outfit* devem estar entre os valores de referência 0,70 e 1,30. Dessa maneira, se o valor de um desses índices

for superior a 1,30, existe 30% mais variação nos dados observados do que o previsto pelo modelo; no caso dos índices estarem abaixo de 0,70, a interpretação é similar, existiria 30% menos variação nos dados observados do que o esperado.

Para verificar a qualidade da estrutura fatorial encontrada na FIFA, um estudo do erro absoluto de ajuste foi realizado através da Modelagem de Equações Estruturais. Para esta análise, todos os modelos foram especificados considerando os itens selecionados para cada dimensão. A técnica de estimação dos modelos estruturais foi a *Weighted Least Squares Means and Variance adjusted* (WLSMV). Os índices de ajuste utilizados foram o *Root Mean Square Error of Approximation* (RMSEA), cujo valor de referência é abaixo de 0,08, e o *Root Mean Square Residual* (RMR), que deve estar abaixo de 0,10 (Henseler, Hubona & Ray, 2016). Também foram reportados índices de qualidade de ajuste, como o *Comparative Fit Index* (CFI) e o *Tucker Lewis Index* (TLI), cujo valor de referência é acima de 0,90, como descrito por Henseler et al. (2016).

Para verificar a consistência interna do OBPI foram calculados o alfa de Cronbach ( $\alpha$ ) e o coeficiente ômega, que foi utilizado para suprir limitações do  $\alpha$ , como a subestimação dos valores devido a pequenas violações dos seus pressupostos (Dunn, Baguley & Brunson, 2013). Os valores do  $\alpha$  e do ômega variam de 0 a 1, sendo que a maioria dos autores, como descrito por Tavakol e Dennick (2011), concordam que o seu valor ideal deve estar entre 0.70 e 0.95. O estudo da confiabilidade das escalas também foi realizado através do reteste, tomando como estatística o coeficiente de correlação de Pearson. Todas as análises foram realizadas no software estatístico R, utilizando os pacotes *mirt* e *lavaan*.

## Resultados

### Evidências de validade baseadas no conteúdo

Os resultados da análise dos juízes indicaram uma concordância substancial; 74,7% das respostas dos juízes concordaram em direção a uma avaliação positiva dos itens. O coeficiente *Fleiss Kappa* foi calculado para examinar tanto a concordância quanto a confiabilidade das avaliações. De acordo com a tabela de valores de referência de Landis e Koch (1977), uma leve concordância foi encontrada entre os juizes,  $\kappa = 0,11$ ,  $p < 0,005$ .

Com o mesmo propósito, a correlação intra-idade (ICC) foi calculada individualmente para a concordância e para a consistência. Uma leve concordância (Landis & Koch, 1977) também foi encontrada, ICC (A,1) = 0,113,  $p < 0,005$ ; o intervalo de confiança de 95% variou de [0,075, 0,156]. Resultados muito similares foram encontrados para o cálculo do ICC para a consistência, ICC (C, 1) = 0,115,  $p < 0,005$ , e o intervalo de confiança de 95% variou de [0,076, 0,159].

A análise qualitativa culminou em 50 modificações pequenas no conteúdo dos itens. Estas modificações foram relacionadas a tradução e incluíram pequenas mudanças, como a inserção de um artigo definido ou o uso de uma palavra sinônima para melhorar a inteligibilidade de alguns itens. Apenas um dos itens sofreu uma mudança mais significativa no conteúdo (item 180). Os juízes apontaram que o item 180 (“Eu geralmente não gosto do tipo de pessoa que se dá bem como um vendedor de telemarketing insistente.”, cuja versão original é: “*I generally dislike the type of person who makes a successful door-to-door salesperson.*”) induzia a uma interpretação estereotipada de uma profissão em particular no Brasil. Para corrigir este aspecto, o item foi alterado e a versão final foi a seguinte: “Eu geralmente não gosto do tipo de pessoa que aborda os outros de forma insistente e invasiva (ex.: como um vendedor de porta em porta)”.

### Evidências de validade baseadas na estrutura interna

Os resultados da FIFA para a escala de Companheirismo demonstraram que apenas nove itens tiveram carga fatorial acima do ponto de corte de 0,30. A análise de resíduos calculados a partir do modelo Rasch, por sua vez, indicaram que os itens não apresentaram padrões de respostas inesperados, a média dos valores de *infit* foi de 0,90 (DP = 0,09), e a média de valores do *outfit* foi de 0,96 (DP = 0,11). Nenhum dos itens apresentou desajuste fora dos valores de referência. O alfa ( $\alpha$ ) de Cronbach da escala como um todo foi  $\alpha = 0,74$ . O valor do coeficiente *omega* foi 0,73 cujo intervalo de confiança de 95% foi [0,69, 0,76].

A escala de Autoridade apresentou seis itens com carga fatorial acima de 0,30. Nenhum dos itens apresentou desajuste de acordo com as análises de resíduo baseadas em Rasch, a média do *infit* foi de 0,96 (DP = 0,06); para o *outfit*, a média foi de 0,97 (DP = 0,07). O valor do  $\alpha$  para esta escala foi de 0,69. O *omega* calculado teve o valor de 0,67, com intervalo de confiança de 95% variando de [0,62, 0,71]. Apenas quatro itens da escala de Conformidade apresentaram cargas fatoriais acima de 0,30. A análise de resíduos revelou que a média do *infit* foi de 0,96 (DP = 0,05) e a média do *outfit* também foi de 0,96 (DP = 0,05), sem a existência de itens desajustados. Apesar das baixas cargas fatoriais para os itens, o valor do  $\alpha$  foi de 0,68, ainda dentro dos padrões aceitáveis (Clark & Watson, 1995). O valor do *omega*, por sua vez, foi de 0,53, com intervalo de confiança de 95% de [0,43, 0,62], indicando que esta escala apresentou problemas de consistência interna.

Os resultados da FIFA para a escala de Emoção demonstraram que dez itens apresentaram cargas fatoriais superiores a 0,30. Similarmente às escalas anteriores, as análises de resíduos não identificaram itens com padrões de respostas incomuns, a média de *infit* foi de 0,95 (DP = 0,12), a do *outfit*, por sua vez, foi 0,96 (DP = 0,12); mais uma vez, sem desajustes para os itens. A escala também apresentou evidências de consistência interna, com  $\alpha = 0,77$  e *omega* também igual a 0,77, cujo intervalo de confiança de 95% foi [0,75, 0,79]. A última escala a compor as escalas primárias do OBPI, Detalhe, apresentou seis itens com carga fatorial superior a 0,30. O estudo dos resíduos da escala através do modelo Rasch não identificou itens com problemas quanto ao ajuste, a média do *infit* foi 0,95 (DP = 0,13) e a média do *outfit* foi 0,96 (DP = 0,13). O valor do  $\alpha$  de Cronbach foi 0,70, assim como o do *omega*, cujo intervalo de confiança de 95% variou entre [0,67, 0,73].

Seguindo o mesmo procedimento, a estrutura interna e fidedignidade das escalas secundárias foram avaliadas. Cada uma das sete escalas, voltadas ao estudo da integridade, foram estudadas sob os mesmos modelos e pressupostos das escalas primárias. O resultado da FIFA para a escala de Proficiência indicou que oito itens obtiveram cargas fatoriais adequadas (acima de 0,30). Não foram observados padrões de respostas inesperados aos itens da escala de Proficiência, a média do *infit* foi 0,94 (DP = 0,10), e o *outfit* teve um média de 0,94 (DP = 0,11). O valor do  $\alpha = 0,68$  sugere que a escala tem consistência interna, de acordo com essa medida. O valor do *omega* foi idêntico ao apresentado pelo  $\alpha$  de Cronbach, 0,68, com intervalo de confiança de 95% entre [0,65, 0,72].

Sete itens apresentaram carga fatorial acima de 0,30 depois da FIFA para a escala de Orientação para o Trabalho. Entretanto, os itens não apresentaram distorções no padrão de respostas, como indicado pelo valor médio de *infit*, 0,95 (DP = 0,07), e do *outfit*, 0,96 (DP = 0,07). O valor do  $\alpha = 0,66$  foi abaixo do valor de referência recomendado, o *omega* também apresentou o valor de 0,66, com intervalo de confiança de 95% entre [0,62, 0,70]. Os resultados para a escala de Paciência demonstraram que oito itens obtiveram

cargas fatoriais suficientemente altas (acima de 0,30). Em relação ao estudo dos resíduos da escala, não existiram indicativos de desajuste nos itens, a média do *infit* foi 0,93 (DP = 0,12) e a do *outfit*, 0,95 (DP = 0,14). Os resultados sugeriram que a escala apresenta consistência interna com  $\alpha = 0,71$  e *omega* com o mesmo valor e intervalo de confiança de 95% entre [0,68, 0,74].

A escala Imparcialidade apresentou sete itens com cargas fatoriais acima de 0,30. Os resultados do *infit* e *outfit*, que indicam que não foram identificados itens com padrões idiossincráticos de resposta, as médias foram 0,95 (DP = 0,06) para ambos *infit* e *outfit*. Em relação a consistência interna da escala, o valor do  $\alpha$  foi 0,65, enquanto que o *omega* foi 0,64, com intervalo de confiança (95%) entre os valores [0,59, 0,68]. Ambos os coeficientes estão abaixo do valor de referência. Sete itens obtiveram uma carga fatorial adequada após a FIFA para a escala Lealdade. Não foram identificados desajuste entre os itens após as análises de resíduos, sob o modelo Rasch. A média do *infit* foi 0,94 (DP = 0,08), e a do *outfit* foi 0,94 (DP = 0,09). Tanto o  $\alpha$  de Cronbach quanto o *omega* para a escala foram

0,67, sendo que o intervalo de confiança de 95% para o *omega* variou entre [0,62, 0,71].

A escala Transparência, que avalia a desejabilidade social do teste, apresentou oito itens com carga fatorial acima de 0,30. Quanto a análise de resíduos sob o modelo Rasch, nenhum item foi identificado com padrões não esperados de respostas, sendo as médias do *infit* e *outfit* iguais a 0,95 (DP = 0,13). A escala apresentou evidências de fidedignidade com ambos os coeficientes iguais a 0,71, o intervalo de confiança de 95% para o *omega* foi [0,68, 0,75]. Seis itens apresentaram carga fatorial acima de 0,30 para a escala de Iniciativa. As análises dos resíduos da escala demonstram que os itens não apresentam *misfit*, ou seja, padrões inesperados não foram encontrados; a média do *infit* foi 0,95 (DP = 0,12), a do *outfit* foi 0,96 (DP = 0,12). Tanto o  $\alpha$  de Cronbach quanto o valor de *omega* foram 0,70, um indicativo de que a escala apresenta evidências de consistência interna. Para o *omega*, o intervalo de confiança de 95% foi entre [0,67, 0,73]. Um resumo dos resultados encontrados, incluindo os índices de ajuste de cada uma das escalas pode ser observado na Tabela 1.

Tabela 1

Resumo dos Resultados Referentes às Escalas Primárias e Secundárias

	Escala	Itens com carga fatorial $\geq 0,30$	Média do <i>misfit</i>	Índices de ajuste	$\alpha$ / <i>omega</i>
Escalas Primárias	Companheirismo	9	<i>Infit</i> = 0,90 (DP = 0,09) <i>Outfit</i> = 0,96 (DP = 0,11)	RMSEA = 0,08 RMR = 0,05 CFI = 0,71 TLI = 0,68	0,74/0,73
	Autoridade	6	<i>Infit</i> = 0,96 (DP = 0,06) <i>Outfit</i> = 0,97 (DP = 0,07)	RMSEA = 0,07 RMR = 0,05 CFI = 0,67 TLI = 0,62	0,69/0,67
	Conformidade	4	<i>Infit</i> = 0,96 (DP = 0,05) <i>Outfit</i> = 0,96 (DP = 0,05)	RMSEA = 0,07 RMR = 0,04 CFI = 0,38 TLI = 0,32	0,68/0,53
	Emoção	10	<i>Infit</i> = 0,95 (DP = 0,12) <i>Outfit</i> = 0,96 (DP = 0,12)	RMSEA = 0,05 RMR = 0,03 CFI = 0,92 TLI = 0,91	0,77/0,77
	Detalhe	6	<i>Infit</i> = 0,95 (DP = 0,13) <i>Outfit</i> = 0,96 (DP = 0,13)	RMSEA = 0,07 RMR = 0,04 CFI = 0,80 TLI = 0,75	0,70/0,70
Escalas Secundárias	Proficiência	8	<i>Infit</i> = 0,94 (DP = 0,10) <i>Outfit</i> = 0,94 (DP = 0,11)	RMSEA = 0,07 RMR = 0,04 CFI = 0,92 TLI = 0,90	0,68/0,68
	Orientação para o Trabalho	7	<i>Infit</i> = 0,95 (DP = 0,07) <i>Outfit</i> = 0,96 (DP = 0,07)	RMSEA = 0,10 RMR = 0,05 CFI = 0,82 TLI = 0,74	0,66/0,66
	Paciência	8	<i>Infit</i> = 0,93 (DP = 0,12) <i>Outfit</i> = 0,95 (DP = 0,14)	RMSEA = 0,06 RMR = 0,03 CFI = 0,95 TLI = 0,93	0,71/0,71
	Imparcialidade	7	<i>Infit</i> = 0,95 (DP = 0,06) <i>Outfit</i> = 0,95 (DP = 0,06)	RMSEA = 0,08 RMR = 0,04 CFI = 0,87 TLI = 0,81	0,65/0,64
	Lealdade	7	<i>Infit</i> = 0,94 (DP = 0,08) <i>Outfit</i> = 0,94 (DP = 0,09)	RMSEA = 0,10 RMR = 0,05 CFI = 0,85 TLI = 0,77	0,67/0,67
	Transparência	8	<i>Infit</i> = 0,95 (DP = 0,13) <i>Outfit</i> = 0,95 (DP = 0,13)	RMSEA = 0,05 RMR = 0,03 CFI = 0,97 TLI = 0,95	0,71/0,71
	Iniciativa	6	<i>Infit</i> = 0,95 (DP = 0,12) <i>Outfit</i> = 0,96 (DP = 0,12)	RMSEA = 0,05 RMR = 0,02 CFI = 0,98 TLI = 0,97	0,70/0,70

Nota: DP = desvio padrão

### Teste – reteste

Os resultados do teste-reteste, que foi realizado através da comparação dos escores dos mesmos indivíduos em momentos distintos, revelaram que o teste apresenta resultados consistentes e estáveis. O tempo decorrido entre a testagem a retestagem foi de quatro meses, tendo em vista que o OBPI é um teste de personalidade, este período foi suficiente para evitar efeitos de treino ou de memória dos itens. Os valores do coeficiente de correlação, utilizados como coeficientes de fidedignidade de teste-reteste, para as escalas primárias e secundárias são apresentados na Tabela 2. Como pode ser notado nenhum dos valores estão abaixo do valor de referência 0,70 (Urbina, 2009), com exceção da escala de Lealdade. Todos os coeficientes apresentaram valores de  $p < 0,05$ .

Tabela 2  
Correlação entre os escores de teste e reteste das escalas primárias e secundárias

	Escala	Correlação de Pearson
Escalas Primárias	Companheirismo	0,83
	Autoridade	0,76
	Conformidade	0,72
	Emoção	0,80
	Detalhe	0,73
Escalas Secundárias	Proficiência	0,77
	Orientação para o Trabalho	0,73
	Paciência	0,84
	Imparcialidade	0,77
	Lealdade	0,69
	Transparência	0,77
	Iniciativa	0,81

### Discussão

Grande parte das mudanças que ocorreram nos itens não alteraram o seu conteúdo como um todo e estavam voltadas para otimizar a sua inteligibilidade. Com exceção do item 180, todos os outros itens sofreram alterações mínimas, principalmente, voltadas para troca de palavras por sinônimos. As informações acerca das características quantitativas indicaram que a associação entre as respostas sugeria uma leve concordância, considerando os parâmetros desenvolvidos por Landis e Koch (1977), a concordância bruta, por sua vez, indicou que havia uma concordância substancial, aproximando-se do valor de 75%.

Apesar de serem largamente utilizados, os parâmetros de referência de Landis e Koch (1977) devem ser observados com cuidado, pois estes não levam em consideração alguns critérios que podem influenciar os valores estatísticos, como o viés e a prevalência de um valor sobre o outro. Se a prevalência for muito alta em um estudo (muitos valores idênticos, por exemplo), os valores de concordância por acaso podem ser altamente acrescidos, afetando negativamente a magnitude dos coeficientes. O viés, por outro lado, diminui os valores de concordância ao acaso, e conseqüentemente, os coeficientes podem ser inflados (Sim & Wright, 2005). Estes autores ainda descrevem que estes valores de referência, apesar de sua utilidade, são valores arbitrários.

Os itens respondidos pelos juízes, eram de caráter dicotômico (“sim” e “não”) e podem ter sofrido um decréscimo dos valores de concordância justamente pela prevalência do valor “sim” nas repostas dos juízes. Além disto, itens dicotômicos contribuem para um aumento na probabilidade de as concordâncias serem consideradas ao acaso. Como sugerido por Gisev et al (2013, p. 333)

“valores baixos de kappa podem não ser um indicativo de baixa concordância”. Na literatura, há a recomendação para reportar o valor do PABAK (*Prevalence-Adjusted-Bias-Adjusted-Kappa*), entretanto, esta estatística é limitada para apenas dois juízes. Neste trabalho, cinco juízes participaram das análises e devido a esta característica, houve uma limitação maior quanto aos coeficientes disponíveis. É importante notar que, em termos de significância estatística, todas as análises demonstraram quão uniforme foram as avaliações realizadas pelos juízes. Assim, tantos os resultados quantitativos quanto qualitativos contribuíram para evidências de que o conteúdo dos itens traduzidos reflete bem a teoria *Big Five* e possui a contextualização adequada à cultura brasileira.

Os resultados da FIFA, baseados na MIRT, indicaram a existência de um grupo de itens com carga fatorial abaixo de 0,30. Estes achados são reflexo do processo de construção destas escalas, que em sua formulação original não passaram pelos tratamentos estatísticos disponíveis atualmente. Estudos anteriores de validade do OBPI foram voltados principalmente para o exame da consistência interna do instrumento, sendo que a versão inglesa não havia passado por análises multivariadas mais sofisticadas, como a análise fatorial.

O modelo teórico considerou que cada uma das escalas era independente, de maneira que dentro do OBPI estas funcionariam como subtestes distintos. Entretanto, os índices de ajuste de algumas escalas demonstraram uma inadequação do modelo unidimensional aos dados. Na Tabela 1 é possível notar que a escala Conformidade, apresentou os menores índices incrementais (CFI = 0,38 e TLI = 0,32) indicando que o modelo teórico unidimensional não está levando em consideração outros aspectos que podem estar influenciando as respostas dos indivíduos. Estes achados podem ser explicados pelos estudos de Anusic, Schimmack, Pinkus & Lockwood (2009), que demonstrou que os fatores do Big Five apresentam interrelações, de maneira que um fator pode interferir nas respostas de outro. Como o modelo utilizado foi unidimensional, não foi possível identificar se tais interferências ocorreram nas escalas do OBPI.

Os estudos de fidedignidade trouxeram em grande parte resultados positivos, assim como o teste-reteste. Comparados aos resultados encontrados por Daouk-Oyry (2008), os valores do alfa de Cronbach encontrados para as escalas do OBPI foram superiores aqueles observados para todos os idiomas nas escalas de Conformidade e Companheirismo, maiores do que a adaptação para árabe, chinês e espanhol nas escalas de Autoridade e Emoção e, por fim, Detalhe obteve valores melhores do que os idiomas chinês e espanhol.

Este trabalho se configurou como o primeiro estudo da adaptação transcultural do OBPI no Brasil, e como tal, encontrou algumas limitações como o tamanho da amostra, que pode ter sido insuficiente para a estimação dos parâmetros e a dificuldade de encontrar coeficientes de concordância entre juízes para itens dicotômicos, por exemplo. Devido ao seu caráter exploratório, os resultados encontrados reuniram um conjunto de informações que subsidiarão estudos futuros, como a necessidade de investigar possíveis correlações entre as escalas do teste e o desenvolvimento de um modelo teórico mais robusto, que examine a estrutura fatorial dos itens diante de tais interrelações. Além disso, faz-se necessário que sejam realizados estudos de simulação para identificar tamanhos amostrais adequados para o uso das técnicas baseadas em Teoria de Resposta ao Item Multidimensional.

### Referências

- Anusic, I., Schimmack, U., Pinkus, R. T., & Lockwood, P. (2009). The nature and structure of correlations among Big Five ratings: The halo-alpha-beta model.

- Journal of Personality and Social Psychology*, 97(6), 1142-1156. doi: [10.1037/a0017159](https://doi.org/10.1037/a0017159)
- Bond, T. G. & Fox, C. M. (2007). *Applying the Rasch model: fundamental measurement in the human sciences*. Mahwah, NJ: Erlbaum.
- Catano, V. M., O'Keefe, D. F., Francis, R. E., & Owens, S. M. (2018). Construct-based approach to developing a short, personality-based measure of integrity. *International Journal of Selection and Assessment*, 26(1), 75-92. doi: [10.1111/ijsa.12197](https://doi.org/10.1111/ijsa.12197)
- Choi, D., Oh, I.-S., & Colbert, A. E. (2015). Understanding organizational commitment: A meta-analytic examination of the roles of the five-factor model of personality and culture. *Journal of Applied Psychology*, 100(5), 1542-1567. doi: [10.1037/apl0000014](https://doi.org/10.1037/apl0000014)
- Cooper, C. A., Knotts, H. G., McCord, D. M., & Johnson, A. (2013). Taking Personality Seriously: The Five-Factor Model and Public Management. *The American Review of Public Administration*, 43(4), 397-415. doi: [10.1177/0275074012446509](https://doi.org/10.1177/0275074012446509)
- Comensoli, A., & MacCann, C. (2013). Misconstruing methods and meaning in the General Factor of Personality. *International Journal of Psychology*, 48(4), 625-630. doi: [10.1080/00207594.2012.688134](https://doi.org/10.1080/00207594.2012.688134)
- Cortez, P. A. (2017). Evidências de validação empírica de escala de atitude empreendedora e testagem de um modelo preditivo a partir dos Cinco Grandes fatores de personalidade. Dissertação de Mestrado, Instituto de Psicologia, Universidade Federal de Uberlândia. Recuperado de <https://repositorio.ufu.br/handle/123456789/18387>
- Do, C. B., & Batzoglou, S. (2008). What is the expectation maximization algorithm? *Nature Biotechnology*, 26(8), 897-899. doi: [10.1038/nbt1406](https://doi.org/10.1038/nbt1406)
- Daouk-Oyry, L. (2008). Towards a culture-free model of the Big Five: a Cross-cultural investigation of the Orpheus in four different language families. Doctor of Philosophy thesis, Department of Psychology, City University, London, United Kingdom. Recuperado de <http://openaccess.city.ac.uk/8715/>
- Fernandes, K. R., & Zanelli, J. C. (2006). O processo de construção e reconstrução das identidades dos indivíduos nas organizações. *Revista de Administração Contemporânea*, 10(1), 55-72. doi: <http://dx.doi.org/10.1590/S1415-65552006000100004>
- Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: key concepts, approaches, and applications. *Research in Social and Administrative Pharmacy*, 9(3), 330-338. doi: [10.1590/S1415-65552006000100004](https://doi.org/10.1590/S1415-65552006000100004)
- Grijalva, E., & Newman, D. A. (2015). Narcissism and counterproductive work behavior (CWB): Meta-analysis and consideration of collectivist culture, Big Five personality, and narcissism's facet structure. *Journal of Applied Psychology*, 64(1), 93-126. doi: [10.1111/apps.12025](https://doi.org/10.1111/apps.12025)
- Hambleton, R. K., Zenisky, A. L. (2011) *Translating and Adapting Tests for Cross-Cultural Assessments*. In Matsumoto, D. & Van de Vijver, F. J. R. (Org.), *Cross-cultural research methods in psychology* (pp. 46-74), New York: Cambridge University Press.
- He, J., & van de Vijver, F. (2012). Bias and Equivalence in Cross-Cultural Research. *Online Readings in Psychology and Culture*, 2(2). doi: [10.9707/2307-0919.1111](https://doi.org/10.9707/2307-0919.1111)
- Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: updated guidelines. *Industrial Management & Data Systems*, 116(1), 2-20. doi: [10.1108/IMDS-09-2015-0382](https://doi.org/10.1108/IMDS-09-2015-0382)
- Hogan, R., Hogan, J. (2007). *Hogan Personality Inventory Manual*. Tulsa, OK: Hogan Assessment System.
- Johnson, T. P. (2014). Snowball Sampling: Introduction. *Wiley StatsRef: Statistics Reference Online*, 1. doi: [10.1002/9781118445112.stat05720](https://doi.org/10.1002/9781118445112.stat05720)
- Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174. doi: [10.2307/2529310](https://doi.org/10.2307/2529310)
- McCrae, R. R., & John, O. P. (1998). An introduction to the five-factor model and its applications. *Personality: Critical Concepts in Psychology*, 60, 175-215. doi: [10.1111/j.1467-6494.1992.tb00970.x](https://doi.org/10.1111/j.1467-6494.1992.tb00970.x)
- Nei, K. S., Foster, J. L., Ness, A. M., & Nei, D. S. (2018). Rule breakers and attention seekers: Personality predictors of integrity and accountability in leaders. *International Journal of Selection and Assessment*, 26(1), 17-26. doi: [10.1111/ijsa.12201](https://doi.org/10.1111/ijsa.12201)
- Reckase, M. (2009). *Multidimensional item response theory* (Vol. 150). New York: Springer.
- Rust, J. (1999). The validity of the Giotto integrity test. *Personality and Individual Differences*, 27(4), 755-768. doi: [10.1016/S0191-8869\(98\)00277-3](https://doi.org/10.1016/S0191-8869(98)00277-3)
- Rust, J. (2001). *Orpheus Manual*. London, UK: The Psychological Corporation.
- Scalise, K. & Allen, D. D. (2015). Use of open-source software for adaptive measurement: Concerto as an R-based computer adaptive development and delivery platform. *Br. J. Math. Stat. Psychol.* 68, 478-496. doi: [10.1111/bmsp.12057](https://doi.org/10.1111/bmsp.12057)
- Syed, N., Saeed, A., & Farrukh, M. (2015). Organization commitment and five factor model of personality: Theory recapitulation. *Journal of Asian Business Strategy*, 5(8), 183. doi: [10.18488/journal.1006/2015.5.8/1006.8.183.190](https://doi.org/10.18488/journal.1006/2015.5.8/1006.8.183.190)
- Sim, J. and Wright, C.C. (2005) The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Physical Therapy*, 85, 257-268. doi: [10.1093/ptj/85.3.257](https://doi.org/10.1093/ptj/85.3.257)
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education*, 2, 53. doi: [10.5116/ijme.4dfb.8dfd](https://doi.org/10.5116/ijme.4dfb.8dfd)
- Templer, K. J. (2012). Five-Factor Model of Personality and Job Satisfaction: The Importance of Agreeableness in a Tight and Collectivistic Asian Society. *Applied Psychology*, 61(1), 114-129. doi: [10.1111/j.1464-0597.2011.00459.x](https://doi.org/10.1111/j.1464-0597.2011.00459.x)
- Urbina, S (2009). *Fundamentos da Testagem Psicológica*. Porto Alegre, RS: Artmed Editora.
- Van de Vijver, F. J., & Leung, K. (2013). Equivalence and bias: A review of concepts, models, and data analytic procedures. In: Matsumoto, D. & Van de Vijver, F. J. R. (Org.), *Cross-cultural research methods in psychology* (pp. 17-45), New York: Cambridge University Press.
- Wu, C. H. (2016). Personality change via work: A job demand-control model of Big-five personality changes. *Journal of Vocational Behavior*, 92, 157-166. doi: [10.1016/j.jvb.2015.12.001](https://doi.org/10.1016/j.jvb.2015.12.001)

Informações sobre o artigo

Recebido em: 28/5/2018

Primeira decisão editorial em: 20/11/2018

Versão final em: 10/12/2018

Aceito em: 14/12/2018