

Estado da publicação: Não informado pelo autor submissor

Comparação de métodos para o controle do viés da aquiescência em escalas balanceadas e desbalanceadas

Gustavo Henrique Martins, Felipe Valentini, Bruno Bonfá-Araujo , Adriana Satiko Ferraz,
Fernanda Otoni, André Pereira Gonçalves

<https://doi.org/10.1590/SciELOPreprints.6284>

Submetido em: 2023-06-20

Postado em: 2023-06-26 (versão 1)

(AAAA-MM-DD)

A moderação deste preprint recebeu o endosso de:

Nelson Hauck Filho (ORCID: <https://orcid.org/0000-0003-0121-7079>)

Comparação de métodos para o controle do viés da aquiescência em escalas balanceadas e desbalanceadas

Comparison of methods for controlling acquiescence bias in balanced and unbalanced scales

Comparación de métodos para controlar el sesgo de aquiescencia en escalas balanceadas y no balanceadas

Título abreviado: Métodos de controle de aquiescência

Gustavo Henrique Martins (<https://orcid.org/0000-0002-5125-2553>)¹²³

Universidade São Francisco e Laboratório de Ciências para a Educação (eduLab21) do Instituto Ayrton Senna

Felipe Valentini (<https://orcid.org/0000-0002-0198-0958>)⁴

Universidade São Francisco

Bruno Bonfá-Araujo (<https://orcid.org/0000-0003-0702-9992>)

University of Western Ontario

Adriana Satiko Ferraz (<https://orcid.org/0000-0002-9856-0094>)

Centro Universitário de Brasília (CEUB)

Fernanda Otoni (<https://orcid.org/0000-0002-9347-7144>)

Universidade São Francisco

André Pereira Gonçalves (<https://orcid.org/0000-0002-2470-4040>)

Universidade Federal da Bahia

¹ O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES) – Código de Financiamento 001. Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) no âmbito do convênio FAPESP/CAPES (nº processos de bolsa no país 2020/04064-7).

² Endereço para correspondência: R. Waldemar César da Silveira, 105, Jardim Cura D'Ars (SWIFT), Campinas - São Paulo, CEP 13045-510. E-mail: gustavoh.martins95@gmail.com.

³ Agradecemos o Professor Doutor Nelson Hauck Filho pela cuidadosa revisão do nosso manuscrito.

⁴ Bolsista de produtividade no Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – PQ2.

Resumo

O controle do viés de aquiescência normalmente envolve a aplicação de itens positivos e negativos. Contudo, pouco se sabe sobre o efeito do balanceamento entre itens positivos e negativos sobre o controle do viés. O objetivo deste estudo foi comparar três modelos de Análise Fatorial Confirmatória (sem controle, MIMIC e Intercepto Randômico) para recuperar a estrutura fatorial de instrumentos desbalanceados e balanceados, a partir de dados simulados e reais (procedentes de um instrumento que avalia Personalidade). Mediante o controle da aquiescência, os resultados indicaram que a performance de escalas balanceadas foi melhor do que de escalas desbalanceadas, bem como na ausência de controle desse viés de resposta, ao considerar as escalas balanceadas e desbalanceadas. Dessa maneira, esta pesquisa aponta para a possibilidade de controle de aquiescência por meio de instrumentos balanceados associada ao uso dos métodos estatísticos na modelagem.

Palavras-chave: Análise Fatorial; Psicometria; Testes Psicológicos; Viés de resposta; Estilo de Resposta.

Abstract

The control of acquiescence bias typically involves the application of positive and negative keyed items. However, the effect of balancing positive and negative items on bias control remains an open question. In this sense, the objective of this study was to compare three models of Confirmatory Factor Analysis (without control, MIMIC, and Random Intercept) to recover the factor structure of unbalanced and balanced instruments based on simulated and real data (from an instrument that assesses Personality). Through the control of acquiescence, the results indicated that the performance of balanced scales was better than that of unbalanced scales, as well as in the absence of control of this response bias, when considering balanced and unbalanced scales. Thus, this research indicates to the possibility of controlling acquiescence through balanced instruments associated with the use of statistical methods in modeling.

Keywords: Factorial Analysis; Psychometrics; Psychological Testes; Response Bias; Response Styles.

Resumen

El control del sesgo de aquiescencia implica la aplicación de ítems positivos y negativos. Sin embargo, el efecto del equilibrio entre ítems positivos y negativos en el control del sesgo sigue siendo una pregunta abierta. En este sentido, el objetivo de este estudio fue comparar tres modelos de Análisis Factorial Confirmatorio (sin control, MIMIC y Intercepto Aleatorio) para recuperar la estructura factorial de instrumentos desbalanceados y balanceados, a partir de datos simulados y reales (a partir de un instrumento que evalúa personalidad). El control de este sesgo de respuesta indicó que el desempeño de escalas balanceadas fue mejor que el de escalas no balanceadas, así como en la ausencia del control de la aquiescencia, al considerar escalas balanceadas y no balanceadas. Por lo tanto, esta investigación sugiere la posibilidad de controlar este sesgo de respuesta por medio de instrumentos balanceados asociados con el uso de métodos estadísticos modelado.

Palabras-clave: Análisis Factorial; Psicometría; Testes Psicologicos; Sesgo de Respuesta; Estilo de Respuesta.

Instrumentos de medida que se utilizam de escalas Likert são suscetíveis a vieses de concordância (ou discordância). Nesse sentido, uma pessoa, por exemplo, pode concordar fortemente com o item “sou falante”, e, ao mesmo tempo, concordar fortemente com o item “sou quieto”. Esse viés de resposta (também compreendido como estilo de resposta) é chamado de aquiescência, caracterizada pela tendência do respondente de endossar os itens de forma sistemática, independentemente do seu conteúdo (Kam & Meyer, 2015; Paulhus, 1991; Wetzel et al., 2016).

O estilo de resposta aquiescente pode ser proveniente da dificuldade do indivíduo para entender o conteúdo dos itens e refletir sobre o quanto as características ali descritas são compatíveis com os seus comportamentos, pensamentos e sentimentos (He et al., 2014). Evidências sugerem que o padrão de resposta aquiescente tende a ser maior na infância (i.e., 10 anos) e vai diminuindo no decorrer da adolescência, até se estabilizar no início da idade adulta (i.e., 18 e 19 anos), sugerindo ser influenciado por elementos cognitivos (Soto et al., 2008). Somado a esses aspectos, a apresentação da aquiescência também pode ser decorrente da construção de instrumentos cujos itens apresentam somente itens de sentido positivo ou negativo, o que contribui para o padrão de resposta aquiescente (Henninger, 2019; Plieninger, 2018; Podsakoff et al., 2003).

Um dos problemas atrelados a aquiescência é a presença de variância de erro sistemática nas respostas e, conseqüentemente, nos escores do instrumento (Danner et al., 2015; Kam & Meyer, 2015; Kuru & Pasek, 2016; Lechner & Rammstedt, 2015), o que pode levar a um comprometimento das suas características psicométricas (Valentini, 2017; Zanon et al., 2018). No caso das estimativas de precisão, as inferências estatísticas sobre os parâmetros estimados podem reproduzir médias amostrais que não correspondem ao traço latente do respondente. Além disso, nas investigações das evidências de validade, a aquiescência tende a inflar correlações de itens redigidos na mesma direção e suprimir correlações de itens que foram

redigidos em direções opostas (Danner et al., 2015; Kam & Meyer, 2015; Kuru & Pasek, 2016; Lechner & Rammstedt, 2015). Isso pode acarretar malefícios na testagem psicológica em diversos contextos da Psicologia, pois a aquiescência tende a enviesar os escores dos indivíduos, pois se distanciam do seu escore verdadeiro.

Uma tentativa de evitar problemas com a aquiescência é construir instrumentos de medida que apresentem escalas balanceadas, compostos por itens de direção positiva e negativa (exemplos de itens: “sou falante” e “sou quieto” (Kam & Zhou, 2015; Lorenzo-Seva & Ferrando, 2009; Soto & John, 2017). Além disso, recomenda-se que: a) evite-se o termo “não” para a inversão dos itens, pois tende a dificultar a leitura e compreensão (Barnette, 2000; Gehlbach & Artino Junior, 2018); b) a chave de resposta seja compatível com o construto/habilidade avaliada pelo instrumento, bem como os itens positivos e negativos devem ser passíveis de serem respondidos por meio delas. Os termos âncoras da escala Likert sejam, de fato, simétricos (i.e., “1–Nada”, “2–Pouco”, “3–Moderadamente”, “4–Muito” e “5–Totalmente”; Sliter & Zickar, 2014; Wetzel et al., 2016).

Outra forma de controle da aquiescência versa sobre a aplicação da modelagem nos dados obtidos com a aplicação do instrumento (Henninger, 2019; Plieninger, 2018; Podsakoff et al., 2003). No presente estudo, buscamos demonstrar duas estratégias para modelagem da aquiescência – Múltiplos Indicadores e Múltiplas Causas (MIMIC – *Multiple Indicators, Multiple Causes*) e Intercepto Randômico; bem como avaliar a importância do balanceamento entre itens positivos e negativos, por meio de dados simulados e reais.

Principais modelos para controlar a aquiescência: MIMIC e Interceptos Randômicos

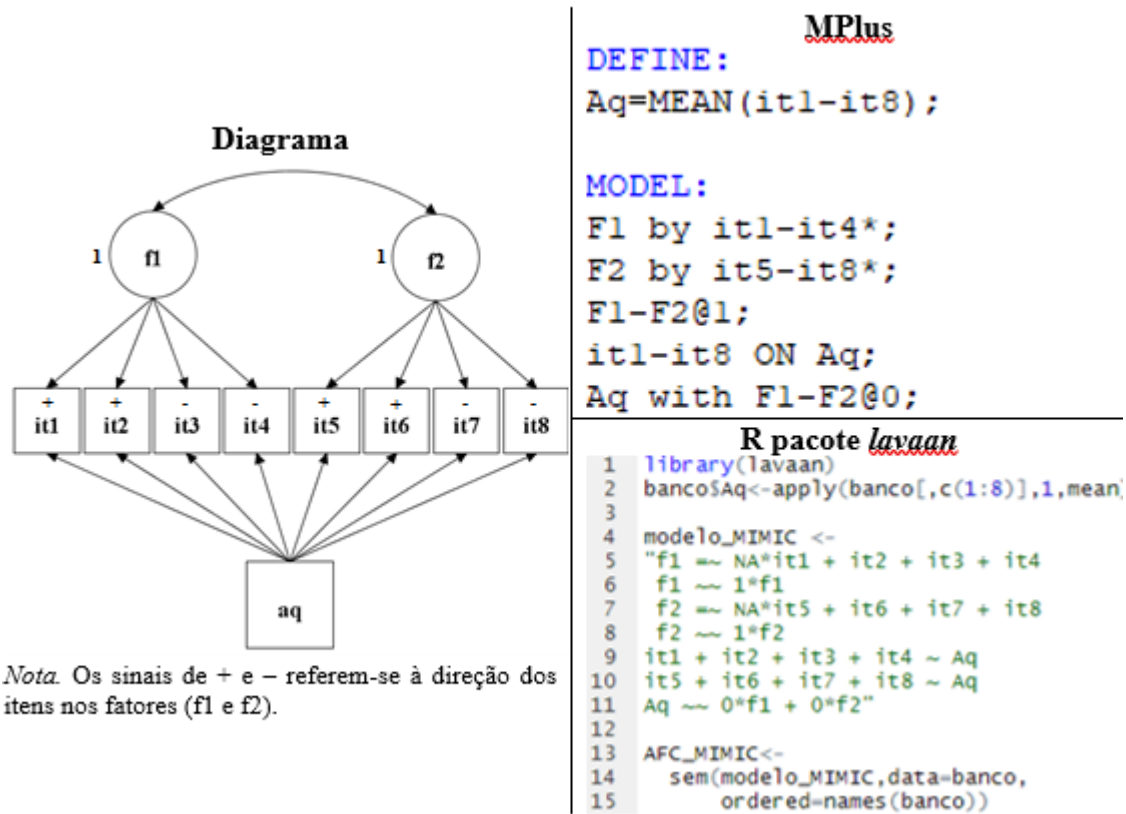
Métodos estatísticos como o MIMIC e Intercepto Randômico podem ser adotados para controlar o viés da aquiescência (Henninger, 2019; Plieninger, 2018; Podsakoff et al., 2003), por meio de análises fatoriais exploratórias (AFE) e confirmatórias (AFC). Para utilizar esses

métodos de controle (MIMIC e Intercepto Randômico), é necessária a presença de itens com direções opostas (positivos e negativos; Geiser, et al., 2008), pois existe a probabilidade de incorrer em *'overcontrol'* (controle além do necessário) quando aplicada em escalas desbalanceadas, isto é, constituídas apenas de itens de direção positiva ou negativa (Valentini & Hauck Filho, 2020).

No MIMIC, além do fator latente de conteúdo (i.e., o construto que se pretende avaliar com o instrumento), os itens também são explicados por (ou regredidos em) uma variável externa, observada, cujos escores representam o índice de aquiescência (normalmente a média dos itens positivos e negativos, sem inversão). Em uma escala balanceada com chave de resposta Likert, por exemplo, a aquiescência é identificada quando o indivíduo concorda com duas situações opostas. Sendo assim, no MIMIC, ao atribuir o índice de aquiescência como causa das respostas aos itens, as cargas fatoriais obtidas no fator de conteúdo estarão controladas e, portanto, livres do viés da aquiescência (Wetzel & Carstensen, 2015). A Figura 1 ilustra um modelo de AFC com MIMIC para controle da aquiescência.

Figura 1

Diagrama de uma Análise Fatorial Confirmatória com controle de aquiescência por MIMIC e sintaxes do MPlus e lavaan



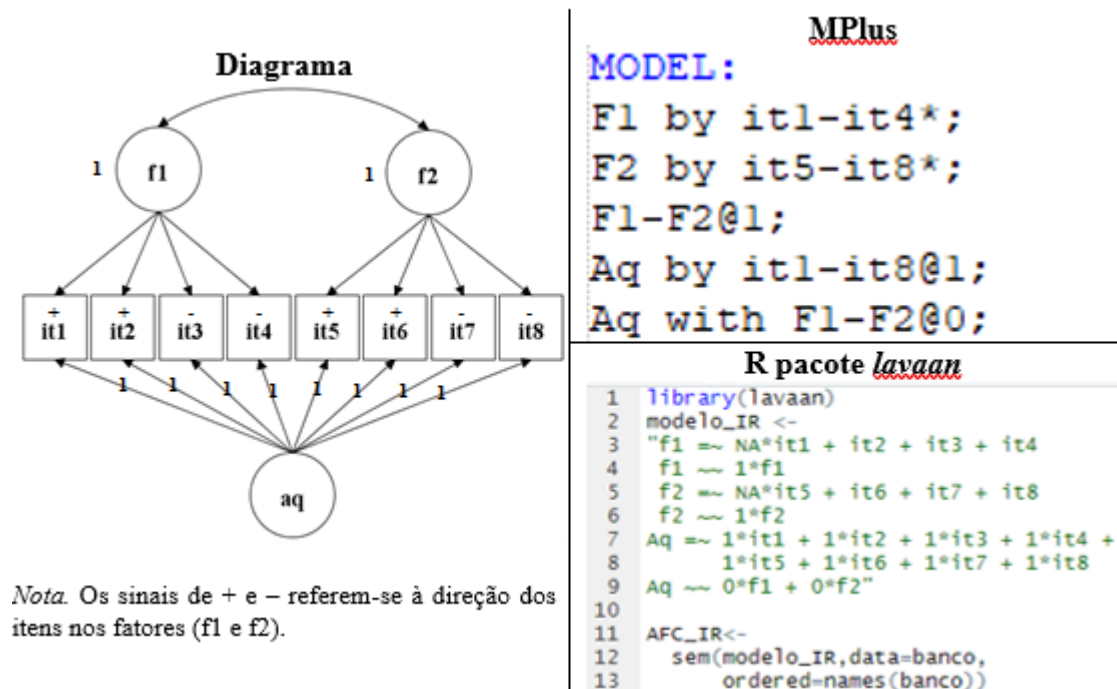
O diagrama apresenta graficamente o modelo MIMIC. Neste caso, são apresentados dois fatores (representados pelos círculos f1 e f2), ambos com dois itens positivos e negativos, além de um escore observado de aquiescência. No *software* MPlus, a partir do comando DEFINE, é calculada a média entre todos os itens, sendo este o escore de aquiescência. Em uma escala balanceada a média dos itens, sem invertê-los, é o valor central da escala Likert. Por exemplo, para o item ‘falante’ um participante assinalou a resposta 5; e para o item ‘quieto’, a resposta 1. A média dele é 3 (5 + 1 / 2), ponto central da escala. No entanto, qualquer desvio dessa média é um indicativo de respostas aquiescentes ou desaquiescentes. Em seguida, a partir do comando MODEL, para além da análise fatorial, os itens são regredidos no escore recém-criado de

aquiescência, bem como a correlação desta com os fatores é fixada em zero (Aq with F1-F2@0). Para o *software* R, fazendo uso do pacote *lavaan*, a primeira linha da sintaxe indica como estimar a média entre todos os itens, seguida da estimação do modelo fatorial.

Por sua vez, o Intercepto Randômico modela um fator latente de aquiescência não correlacionado aos fatores de conteúdo (Maydeu-Olivares & Coffman, 2006). As cargas dos itens no fator de aquiescência (Intercepto Randômico) são fixadas em +1 (tanto para os itens positivos quanto para os negativos). Assim, o modelo estimará o fator de aquiescência captando apenas a variância estritamente comum e unidirecional entre os itens positivos e negativos (por exemplo, tendência de concordância aos itens positivos e negativos). A Figura 2 ilustra um modelo de AFC com controle de aquiescência por Intercepto Randômico. Ressaltamos que a restrição nas cargas não é uma condição para a identificação do modelo de equações estruturais, mas é utilizada para captar a variância de aquiescência. Considerando que a aquiescência reflete uma tendência de endossar positivamente a todos os itens, sendo positivos ou negativos, a restrição imposta às cargas em +1 busca captar a variância comum que é independente da direção dos itens (i.e., a aquiescência). Assim, liberar as cargas para serem estimadas livremente pode incorrer em um modelo que capta outros tipos de vieses e, até mesmo, conteúdo psicológico legítimo do item, incorrendo em *overcontrol* (i.e., superestimar o viés e remover variância do construto psicológico).

Figura 2

Diagrama de uma Análise Fatorial Confirmatória com controle de aquiescência por Intercepto Randômico e sintaxes do MPlus e lavaan



Em ambos os softwares MPlus ou R (*lavaan*), estima-se uma AFC comum, adicionando uma linha para modelar a variável latente de aquiescência. A correlação entre aquiescência e fatores é fixada em zero, para identificação do modelo. Note que a variância de aquiescência é estimada (não fixada), permitindo a comparação com as variâncias genuínas dos fatores.

Na presente pesquisa utilizamos exemplos de escalas balanceadas, isto é, com o mesmo número de pares de itens opostos (positivos e negativos) em cada fator. No entanto, MIMIC e Intercepto Randômico permitem, hipoteticamente, o desbalanceamento entre itens positivos e negativos. Nos casos de desbalanceamento, a performance da modelagem permanece adequada? Assim, nosso objetivo é comparar a performance de três modelos de AFC (sem controle algum, método de controle MIMIC e método de controle Intercepto Randômico) para recuperar a estrutura fatorial de instrumentos desbalanceados e balanceados. Nossa pesquisa

envolve dois estudos: o primeiro com dados simulados e o segundo com dados reais coletados com os itens do *Big Five Inventory* (BFI-2; Soto & John, 2017), a partir de dados simulados e reais.

O BFI-2 (Soto & John, 2017) avalia a personalidade na perspectiva do modelo *Big Five* que, em linhas gerais, compreende esse construto em cinco domínios amplos, a saber, extroversão, amabilidade, conscienciosidade, neuroticismo e abertura à experiência. Esse instrumento de autorrelato é composto por itens que demarcam os adjetivos de cada traço de personalidade, sendo estruturados em sentenças curtas e simples e respondido por meio de uma escala Likert. A inteligibilidade do BFI-2 e a sua qualidade psicométrica (i.e., estimativas de precisão e evidências de validade) (John, 2021; Soto & John, 2017) determinaram a sua seleção para a condução do presente estudo.

Estudo de Simulação

Método

Condições de simulação

Foram testadas 18 condições variando os seguintes critérios:

a) Cargas fatoriais baixas/moderadas (0,30 a 0,50) e altas (0,60 a 0,80). Justificamos o critério em função de cargas comumente observadas nos instrumentos psicológicos, pela prática comum de descarte de itens com cargas menores que 0,30 e itens com cargas acima de 0,80 serem incomuns.

b) Variâncias de aquiescência (0,05, 0,10 e 0,20). Variâncias de aquiescência entre 0,05 e 0,10 são comuns na literatura (Soto et al., 2008), contudo, também inserimos a condição com variância de 0,20 para testar o desempenho dos métodos de controle de aquiescência neste cenário menos comum, porém possível;

c) Tamanhos amostrais (200, 500 e 1000). Inserimos essas condições para avaliar se os métodos de controle de aquiescência são eficazes em diferentes tamanhos de amostra.

Cada condição representa uma interação dos três critérios. No que se refere aos *thresholds*, eles foram sorteados a partir de uma distribuição uniforme, mantendo o mesmo valor para os pares opostos de itens. As correlações simuladas entre os fatores variaram de 0,26 a 0,59, representando situações bem comuns no contexto da Psicologia.

Procedimentos de Simulação e Análise

Os dados foram simulados para reproduzir um instrumento com 30 itens e cinco fatores. Essa quantidade de itens e fatores é comum em outros instrumentos psicológicos do contexto brasileiro (Ambiel et al., 2022; Zanon et al., 2014; Siqueira, 2008). Para tanto, foram seguidos quatro passos: simulação dos bancos; seleção de itens; modelagem; recuperação e resumo dos parâmetros modelados.

No primeiro passo foram simulados os bancos de dados. Embora nosso objetivo fosse reproduzir um instrumento com 30 itens, para testar diferentes condições de balanceamento foi necessário, nesta etapa, simular bancos de dados com 60 itens balanceados. Tais itens representam pares psicometricamente opostos, isto é, com a mesma magnitude de carga fatorial, *threshold* e variância de aquiescência entre os itens positivos e negativos. Consideramos que itens de instrumentos desbalanceados são uma amostra de itens balanceados. Nesse sentido, a simulação dos itens extras, permitiu a seleção de amostras de itens balanceados e desbalanceados, na segunda etapa. Caso tivéssemos simulado bancos de dados separados para as condições balanceadas e desbalanceadas, poderíamos ter um efeito de delineamento contaminando principalmente os itens desbalanceados. Simulamos bancos de representar respostas a uma escala Likert de cinco pontos. A estrutura fatorial simulada foi composta por cinco fatores, com 12 itens cada.

Na segunda etapa, selecionamos 30 itens para representar um instrumento balanceado (três positivos e três negativos por fator) e 30 itens para representar um instrumento desbalanceado (cinco positivos e um negativo por fator). Desta forma tivemos quatro itens comuns por fator em ambas as versões, sendo três positivos e um negativo. Além disso, optamos por não escolher os itens que eram os perfeitos pares opostos, pelo de fato de pares psicometricamente opostos serem mais raros em instrumentos reais.

Na terceira etapa, de modelagem, realizamos três AFC para cada uma das versões (i.e., balanceada e desbalanceada). A diferença dos três modelos de AFC ocorreu no controle da aquiescência. No primeiro a aquiescência não foi controlada, no segundo controlamos pelo método MIMIC e o terceiro controlamos pelo método Intercepto Randômico. Todos os modelos de AFC replicaram a estrutura simulada formada por cinco fatores. Além disso, o estimador utilizado foi o *Diagonally Weighted Least Squares* (DWLS), considerando a matriz de peso total para corrigir o cálculo do erro padrão, da média e da estatística de teste ajustada à variância, que é adequado para dados ordinais (Li, 2016). Para cada análise de cada condição testada realizamos 500 replicações.

Na quarta etapa, para comparar o desempenho das AFC testadas, foram realizados os cálculos dos vieses que se referem à distância da carga obtida na análise em comparação com a carga que foi simulada (i.e., carga obtida – carga simulada/carga simulada). Neste estudo, consideramos vieses de até 10% (i.e., 0,10) como aceitáveis. Um viés positivo indica que a carga fatorial obtida teve magnitude superior à simulada, já um viés negativo significa que a carga obtida foi de magnitude menor que a simulada. Para calcular o viés de um conjunto de itens positivos e negativos, calculamos a média dos vieses obtidos nos itens positivos e negativos, respectivamente, em uma determinada condição. Dessa forma, podemos obter uma medida geral do viés presente em um conjunto de itens, considerando tanto os itens positivos quanto os itens negativos. Os dados foram analisados no *software* R por meio do pacote

“lavaan” (Rosseel, 2012). O *script* da simulação realizada neste estudo está disponível no link: https://github.com/GustavHM/simulacao_aquiescencia.

Resultados

A Tabela 1 apresenta o resumo dos vieses das cargas fatoriais obtidas nas condições testadas.

Tabela 1*Resumo da média dos vieses das cargas obtidas nas condições testadas na simulação*

| Condições | | | Itens desbalanceados | | | | | | Itens balanceados | | | | | |
|-----------|---------|---------|----------------------|--------------|-----------|-----------|-----------|-----------|-------------------|--------------|-----------|-----------|-----------|-----------|
| | | | Sem controle | | MIMIC | | IR | | Sem controle | | MIMIC | | IR | |
| Cargas | Var Aqu | Amostra | Positivos | Negativos | Positivos | Negativos | Positivos | Negativos | Positivos | Negativos | Positivos | Negativos | Positivos | Negativos |
| Baixas | 0,20 | 200 | 0,50 | -1,60 | 0,12 | -0,20 | | -0,15 | 0,11 | -2,11 | | | | |
| Baixas | 0,20 | 500 | 0,50 | -1,61 | 0,13 | -0,22 | | | 0,12 | -2,10 | | | | |
| Baixas | 0,20 | 1000 | 0,50 | -1,61 | 0,13 | -0,22 | | | 0,12 | -2,10 | | | | |
| Baixas | 0,10 | 200 | 0,29 | -1,00 | 0,12 | -0,20 | | -0,10 | -0,48 | -0,56 | | | | |
| Baixas | 0,10 | 500 | 0,30 | -1,01 | 0,12 | -0,22 | | | -0,59 | -0,52 | | | | |
| Baixas | 0,10 | 1000 | 0,29 | -1,01 | 0,12 | -0,22 | | | -0,74 | -0,44 | | | | |
| Baixas | 0,05 | 200 | 0,17 | -0,55 | | -0,14 | | | | | | | | |
| Baixas | 0,05 | 500 | 0,17 | -0,55 | | -0,16 | | | | | | | | |
| Baixas | 0,05 | 1000 | 0,17 | -0,56 | | -0,16 | | | | | | | | |
| Altas | 0,05 | 200 | | -0,18 | | | | | | | | | | |
| Altas | 0,05 | 500 | | -0,18 | | | | | | | | | | |
| Altas | 0,05 | 1000 | | -0,18 | | | | | | | | | | |
| Altas | 0,10 | 200 | 0,11 | -0,35 | | | | | | | | | | |
| Altas | 0,10 | 500 | 0,11 | -0,35 | | | | | | | | | | |
| Altas | 0,10 | 1000 | 0,11 | -0,35 | | | | | | | | | | |
| Altas | 0,20 | 200 | 0,20 | -0,68 | | | | | | | | | | |
| Altas | 0,20 | 500 | 0,20 | -0,68 | | | | | | | | | | |
| Altas | 0,20 | 1000 | 0,20 | -0,68 | | | | | | | | | | |

Nota. Var Aqu = Variância da Aquiescência; Cargas baixas = 0,30 a 0,50; Cargas altas = 0,60 a 0,80. Os vieses iguais ou maiores que 0,50 estão em negrito e os menores que 0,10 foram omitidos.

De forma geral, na Tabela 1, é possível observar que a maior frequência de vieses superiores a 0,10 foram observados nos dados desbalanceados, sem controle de aquiescência e com as cargas fatoriais de conteúdo (i.e., cargas no fator que representa o construto que se pretende medir com o instrumento) baixas (i.e., 0,30 a 0,50). Nota-se também que na maioria das condições, os vieses indicaram uma superestimação das cargas dos itens positivos (i.e., vieses positivos) e subestimação das cargas dos itens negativos (i.e., vieses negativos). Para melhor visualizar os resultados, selecionamos algumas condições que apresentaram vieses moderados ($\geq 0,50$) e graves (≥ 1) e apresentamos, na Tabela 2, os dados desbalanceados, e na Tabela 3, os dados balanceados. Tendo em vista que os vieses permaneceram praticamente os mesmos independentemente do tamanho da amostra, foi então selecionada a amostra de 1000 casos para exemplificar os vieses obtidos nas diferentes variâncias de aquiescência e cargas simuladas.

Tabela 2*Cargas fatoriais obtidas e simuladas nos dados com os itens desbalanceados*

| Cargas simuladas | Simulações com cargas baixas | | | Simulação com cargas altas | | |
|------------------|------------------------------|---------------------|---------------------|----------------------------|---------------------|-------|
| | Var Aqu 0,05 | Var Aqu 0,10 | Var Aqu 0,20 | Var Aqu 0,20 | | |
| | Cargas sem controle | Cargas sem controle | Cargas sem controle | Cargas simuladas | Cargas sem controle | |
| 0,43 | 0,48 | 0,52 | 0,59 | 0,73 | 0,85 | |
| 0,38 | 0,44 | 0,49 | 0,58 | 0,68 | 0,82 | |
| 0,42 | 0,48 | 0,52 | 0,59 | 0,72 | 0,85 | |
| 0,40 | 0,45 | 0,50 | 0,58 | 0,70 | 0,83 | |
| 0,33 | 0,41 | 0,46 | 0,56 | 0,63 | 0,78 | |
| -0,31 | -0,09 | 0,09 | 0,31 | -0,61 | -0,12 | |
| 0,33 | 0,41 | 0,46 | 0,55 | 0,63 | 0,78 | |
| 0,38 | 0,45 | 0,49 | 0,57 | 0,68 | 0,82 | |
| 0,30 | 0,38 | 0,45 | 0,54 | 0,60 | 0,76 | |
| 0,42 | 0,48 | 0,53 | 0,59 | 0,72 | 0,85 | |
| 0,45 | 0,51 | 0,54 | 0,60 | 0,75 | 0,87 | |
| -0,45 | -0,22 | -0,03 | 0,22 | -0,75 | -0,27 | |
| 0,47 | 0,52 | 0,56 | 0,62 | 0,77 | 0,89 | |
| 0,39 | 0,45 | 0,51 | 0,59 | 0,69 | 0,83 | |
| 0,38 | 0,45 | 0,50 | 0,58 | 0,68 | 0,82 | |
| 0,41 | 0,47 | 0,52 | 0,59 | 0,71 | 0,84 | |
| 0,49 | 0,54 | 0,57 | 0,63 | 0,79 | 0,91 | |
| -0,43 | -0,21 | -0,02 | 0,23 | -0,73 | -0,26 | |
| 0,41 | 0,47 | 0,51 | 0,58 | 0,71 | 0,84 | |
| 0,35 | 0,42 | 0,48 | 0,57 | 0,65 | 0,80 | |
| 0,34 | 0,42 | 0,47 | 0,56 | 0,64 | 0,79 | |
| 0,32 | 0,39 | 0,45 | 0,55 | 0,62 | 0,77 | |
| 0,46 | 0,52 | 0,55 | 0,61 | 0,76 | 0,88 | |
| -0,49 | -0,25 | -0,06 | 0,20 | -0,79 | -0,31 | |
| 0,48 | 0,53 | 0,56 | 0,61 | 0,78 | 0,89 | |
| 0,33 | 0,41 | 0,47 | 0,56 | 0,63 | 0,78 | |
| 0,37 | 0,44 | 0,49 | 0,57 | 0,67 | 0,81 | |
| 0,45 | 0,51 | 0,54 | 0,60 | 0,75 | 0,87 | |
| 0,39 | 0,46 | 0,50 | 0,58 | 0,69 | 0,83 | |
| -0,41 | -0,18 | 0,01 | 0,26 | -0,71 | -0,21 | |
| | Viés | | Viés | | Viés | |
| | Pos | Neg | Pos | Neg | Pos | Neg |
| | 0,17 | -0,56 | 0,29 | -1,01 | 0,50 | -1,61 |
| | | | | | 0,20 | -0,68 |

Nota. Var Aqu = Variância da Aquiescência. Nas quatro condições foram selecionadas a amostra de 1000 casos. As cargas dos itens negativos estão em negrito.

Os resultados das condições apresentadas na Tabela 2 evidenciam que a aquiescência superestimou as cargas dos itens positivos. Contudo, os principais erros de estimação foram referentes as cargas dos itens negativos, os quais nas três condições apresentadas, tiveram suas cargas fatoriais fortemente subestimadas. Na primeira e quarta condições, os itens negativos tiveram cargas negativas, mas abaixo de 0,30. Na segunda, os itens negativos tiveram cargas

próximas a zero, e na terceira, os itens negativos tiveram suas cargas invertidas, apresentando cargas positivas nos fatores que eram para serem negativas. Resumidamente, nestes quatro casos, todos os cinco itens negativos apresentaram cargas fatoriais baixas e incongruentes com o esperado teoricamente. Na Tabela 3, são apresentadas as cargas fatoriais simuladas e obtidas em duas condições dos dados balanceados.

Tabela 3

Cargas fatoriais obtidas e simuladas nos dados com os itens balanceados

| Cargas baixas simuladas | Var Aqu 0,10 | | Var Aqu 0,20 | |
|-------------------------|---------------------|------------|---------------------|------------|
| | Cargas sem controle | | Cargas sem controle | |
| 0,43 | 0,10 | | 0,43 | |
| 0,38 | 0,08 | | 0,43 | |
| 0,42 | 0,10 | | 0,43 | |
| -0,40 | -0,20 | | 0,44 | |
| -0,33 | -0,18 | | 0,44 | |
| -0,31 | -0,18 | | 0,44 | |
| 0,33 | 0,09 | | 0,43 | |
| 0,38 | 0,10 | | 0,42 | |
| 0,30 | 0,07 | | 0,43 | |
| -0,42 | -0,24 | | 0,46 | |
| -0,45 | -0,25 | | 0,46 | |
| -0,45 | -0,25 | | 0,46 | |
| 0,47 | 0,14 | | 0,42 | |
| 0,39 | 0,11 | | 0,42 | |
| 0,38 | 0,11 | | 0,42 | |
| -0,41 | -0,24 | | 0,45 | |
| -0,49 | -0,27 | | 0,46 | |
| -0,43 | -0,24 | | 0,45 | |
| 0,41 | 0,12 | | 0,41 | |
| 0,35 | 0,09 | | 0,42 | |
| 0,34 | 0,09 | | 0,42 | |
| -0,32 | -0,19 | | 0,44 | |
| -0,46 | -0,26 | | 0,45 | |
| -0,49 | -0,26 | | 0,45 | |
| 0,48 | 0,13 | | 0,42 | |
| 0,33 | 0,08 | | 0,42 | |
| 0,37 | 0,09 | | 0,42 | |
| -0,45 | -0,24 | | 0,45 | |
| -0,39 | -0,22 | | 0,45 | |
| -0,41 | -0,23 | | 0,45 | |
| | Viés | | Viés | |
| | Pos | Neg | Pos | Neg |
| | -0,74 | -0,44 | 0,12 | -2,10 |

Nota. Var Aqu = Variância da Aquiescência. Em ambas as condições foram selecionadas a amostra de 1000 casos. As cargas dos itens negativos estão em negrito.

Nota-se na Tabela 3, que nas duas condições apresentadas, os problemas envolveram todos os itens, positivos e negativos. Na primeira, todas as cargas foram subestimadas a ponto de todos os itens terem cargas abaixo de 0,30, especialmente os itens positivos. Na terceira condição, todos os itens apresentaram cargas na mesma direção e com a mesma magnitude, variando de 0,41 a 0,46, sugerindo estar representando somente a variância do viés da aquiescência (i.e., 0,20) e não a variância de conteúdo. Portanto, nas duas condições todos os itens apresentaram problemas, pelo fato de as cargas fatoriais serem muito baixas e/ou incoerentes teoricamente.

Além dos vieses nas cargas, observamos alguns erros de identificação nas condições testadas, para amostra de 200 e 500 casos. Os casos com mais problemas de identificação, variando de 14,40% a 19,60% (das 500 replicações), ocorreram nas três análises (sem controle, MIMIC e Intercepto Randômico) com os itens balanceados, cargas fatoriais baixas, amostra de 200 casos e independentemente da variância da aquiescência. Além destas, com cargas altas, variância de aquiescência de 0,20 e 200 casos, a análise com Intercepto Randômico apresentou erros de identificação em 12,40% das replicações realizadas. Nas demais condições os problemas de identificação foram inferiores a 4% das replicações. Ressaltamos que os problemas de identificação ocorreram em função de parâmetros não plausíveis, como variância negativa.

Ilustração com dados reais

Buscamos ilustrar o uso dos Interceptos Randômicos e MIMIC para controle de aquiescência em um banco de dados real, no qual foi utilizado um instrumento balanceado de personalidade. Além disso, testamos uma versão desbalanceada de seleção de itens para avaliar o impacto dessa condição na aquiescência.

Método

Participantes

A amostra foi composta por 888 adultos de diversas regiões do Brasil. A maioria dos participantes eram do sexo feminino (70,9%) e as idades variaram de 18 a 73 anos ($M = 23,34$; $DP = 7,65$).

Instrumentos

Big Five Inventory (BFI-2; Soto & John, 2017).

É composto por 60 itens balanceados que se dividem igualmente nas cinco dimensões do modelo Big Five: extroversão (ex. Sou extrovertido, sociável – $\alpha = 0,87$), amabilidade (ex. Me importo com os outros e tento ajudá-los quando precisam de mim – $\alpha = 0,83$), conscienciosidade (ex. Sou responsável, faço minha parte do trabalho – $\alpha = 0,88$), neuroticismo (ex. Meu humor varia com muita rapidez, facilmente fico mal-humorado – $\alpha = 0,91$) e abertura (ex. Tenho curiosidade sobre muitos assuntos diferentes – $\alpha = 0,84$). A chave de resposta é dada em uma escala Likert de cinco pontos que varia de “1 - Nada. Não tem nada a ver comigo” a “5 - Totalmente. Tem tudo a ver comigo”.

Procedimentos

A amostra deste estudo é produto de duas dissertações de Mestrado que tiveram seus projetos submetidos e aprovados pelo Comitê de Ética em Pesquisa da Universidade São Francisco (CAAE: 01465718.3.0000.5514; CAAE: 08033419.9.0000.5514). A coleta ocorreu de forma presencial ($n = 285$) e on-line ($n = 603$). A coleta presencial foi realizada coletivamente com universitários no formato lápis e papel. Por sua vez, o link da coleta on-line foi disponibilizado por meio de redes sociais.

Análise de dados

Foram replicadas as AFC sem controle, com MIMIC e com Intercepto Randômico nos dados reais do BFI-2. Em todas as análises, foi utilizado o estimador DWLS considerando a matriz de peso total para corrigir o cálculo do erro padrão, da média e da estatística de teste ajustada à variância. Para replicar as análises, foram selecionados 30 itens balanceados e desbalanceados dentro dos 60 itens do BFI-2. Ressalta-se que cada fator do BFI-2 possui seis itens positivos e seis itens negativos. A escolha dos itens foi realizada seguindo uma sistemática parecida com a adotada na seleção dos itens simulados, isto é: a) na versão balanceada, foram selecionados os três primeiros itens positivos e os três últimos itens negativos, de acordo com a ordem que os itens são apresentados no instrumento completo, e b) na versão desbalanceada, foram selecionados os cinco primeiros itens positivos e o último item negativo de cada fator. Com essa forma sistemática de escolha dos itens, conseguimos selecionar sempre quatro itens iguais por fator nas duas versões (i.e., balanceada e desbalanceada), sendo três positivos e um negativo. Os índices de ajuste foram interpretados da seguinte forma: Confirmatory Fit Index (CFI > 0,95), Tucker-Lewis Index (TLI > 0,95), Root Mean Square Error of Approximation (RMSEA < 0,06; Hu & Bentler, 1999). Os dados foram analisados no *software* R por meio do pacote “lavaan” (Rosseel, 2012). O *script* e o banco das análises realizadas neste estudo estão disponíveis no link: https://github.com/GustavHM/simulacao_aquiescencia.

Resultados

A seguir, foram testadas as mesmas análises da simulação (sem controle, MIMIC e Intercepto Randômico) com itens desbalanceados e balanceados a partir de dados reais do BFI-2. Na Tabela 4, estão os índices de ajuste obtidos nas análises.

Tabela 4*Índices de ajuste das análises com o BFI-2*

| | χ^2 (gl) | CFI | TLI | RMSEA |
|----------------------------|---------------|------|------|-------|
| Desbalanceada sem controle | 3240,49 (395) | 0,86 | 0,84 | 0,09 |
| Desbalanceada MIMIC | 2922,16 (395) | 0,88 | 0,85 | 0,09 |
| Desbalanceada IR | 2917,26 (394) | 0,87 | 0,86 | 0,09 |
| Balanceada sem controle | 2222,19 (395) | 0,85 | 0,84 | 0,07 |
| Balanceada MIMIC | 1860,23 (395) | 0,89 | 0,87 | 0,07 |
| Balanceada IR | 1866,66 (394) | 0,88 | 0,87 | 0,07 |

De forma geral, os índices de ajuste foram todos aceitáveis e parecidos nas três análises (sem controle, MIMIC e Intercepto Randômico). Houve melhora dos índices de ajuste dos itens balanceados nas análises com MIMIC e Intercepto Randômico em comparação com a análise sem controle. Na Tabela 5, são descritos os resultados das cargas fatoriais obtidas nas três análises realizadas com os itens desbalanceados e balanceados do BFI-2.

Tabela 5*Cargas fatoriais nas análises com os itens desbalanceados e balanceados do BFI-2*

| Desbalanceado | | | | Balanceado | | | |
|---------------|--------------|-------|-------|------------|--------------|-------|-------|
| Itens | Sem controle | MIMIC | IR | Itens | Sem controle | MIMIC | IR |
| BFI1 | 0,77 | 0,77 | 0,73 | BFI1 | 0,74 | 0,75 | 0,76 |
| BFI6 | 0,40 | 0,41 | 0,35 | BFI6 | 0,39 | 0,47 | 0,40 |
| BFI21 | 0,52 | 0,53 | 0,48 | BFI21 | 0,57 | 0,65 | 0,58 |
| BFI41 | 0,79 | 0,77 | 0,78 | BFI31- | -0,74 | -0,71 | -0,75 |
| BFI46 | 0,58 | 0,58 | 0,54 | BFI36- | -0,44 | -0,40 | -0,43 |
| BFI51- | -0,52 | -0,49 | -0,58 | BFI51- | -0,64 | -0,57 | -0,63 |
| BFI2 | 0,82 | 0,82 | 0,78 | BFI2 | 0,47 | 0,58 | 0,47 |
| BFI7 | 0,74 | 0,73 | 0,71 | BFI7 | 0,74 | 0,81 | 0,75 |
| BFI27 | 0,41 | 0,41 | 0,38 | BFI27 | 0,31 | 0,39 | 0,32 |
| BFI32 | 0,87 | 0,85 | 0,84 | BFI37- | -0,72 | -0,61 | -0,72 |
| BFI52 | 0,76 | 0,76 | 0,72 | BFI42- | -0,40 | -0,26 | -0,39 |
| BFI47- | -0,36 | -0,30 | -0,45 | BFI47- | -0,45 | -0,30 | -0,44 |
| BFI13 | 0,61 | 0,64 | 0,57 | BFI13 | 0,53 | 0,56 | 0,54 |
| BFI18 | 0,75 | 0,72 | 0,72 | BFI18 | 0,81 | 0,83 | 0,80 |
| BFI33 | 0,83 | 0,80 | 0,80 | BFI33 | 0,87 | 0,90 | 0,87 |
| BFI38 | 0,61 | 0,61 | 0,57 | BFI28- | -0,69 | -0,65 | -0,69 |
| BFI43 | 0,68 | 0,71 | 0,62 | BFI48- | -0,83 | -0,81 | -0,85 |
| BFI58- | -0,54 | -0,47 | -0,65 | BFI58- | -0,60 | -0,52 | -0,60 |
| BFI14 | 0,73 | 0,61 | 0,71 | BFI14 | 0,81 | 0,73 | 0,81 |
| BFI19 | 0,55 | 0,46 | 0,53 | BFI19 | 0,60 | 0,48 | 0,59 |
| BFI39 | 0,91 | 0,87 | 0,89 | BFI39 | 0,68 | 0,55 | 0,67 |
| BFI54 | 0,93 | 0,88 | 0,90 | BFI29- | -0,64 | -0,70 | -0,66 |
| BFI59 | 0,71 | 0,59 | 0,69 | BFI44- | -0,63 | -0,70 | -0,64 |
| BFI49- | -0,37 | -0,41 | -0,39 | BFI49- | -0,42 | -0,48 | -0,44 |
| BFI10 | 0,57 | 0,57 | 0,53 | BFI10 | 0,57 | 0,63 | 0,58 |
| BFI15 | 0,46 | 0,46 | 0,40 | BFI15 | 0,48 | 0,59 | 0,50 |
| BFI20 | 0,80 | 0,80 | 0,77 | BFI20 | 0,53 | 0,55 | 0,55 |
| BFI35 | 0,83 | 0,83 | 0,80 | BFI45- | -0,75 | -0,69 | -0,75 |
| BFI40 | 0,48 | 0,47 | 0,41 | BFI50- | -0,54 | -0,50 | -0,55 |
| BFI55- | -0,38 | -0,40 | -0,51 | BFI55- | -0,45 | -0,41 | -0,44 |

Nota. Os conjuntos de seis itens representam, nesta ordem, as seguintes dimensões: Extroversão, Amabilidade, Conscienciosidade, Emocionalidade Negativa e Abertura à Experiências. O símbolo – na descrição de alguns itens indicam que estes são de direção negativa.

Nota-se que com os itens desbalanceados as cargas fatoriais dos itens negativos foram menores nas análises sem controle e com MIMIC, em comparação com a análise com Intercepto Randômico. Ao passo que não foram obtidas diferenças substanciais entre as cargas fatoriais dos itens balanceados, nas três análises. A variância da aquiescência nas análises com Intercepto Randômico, foram de 0,05 (desbalanceado) e 0,04 (balanceado) no BFI-2.

Discussão

O presente estudo teve o objetivo de comparar três modelagens (sem controle algum, método de controle MIMIC e Intercepto Randômico) para instrumentos que podem sofrer influência da aquiescência, utilizando dados simulados e um banco real. De maneira geral, os resultados apontaram para a importância de modelar a aquiescência nas análises fatoriais de instrumentos formados por itens positivos e negativos. Além disso, o modelo de Intercepto Randômico foi levemente superior ao MIMIC, e ambos tiveram performance variada nas diferentes condições.

Quanto ao estudo de simulação, a performance variou consideravelmente nas interações entre as condições de balanceamento, método de controle e tamanhos das cargas fatoriais. Além disso, o tamanho da aquiescência influenciou a performance dos modelos. No entanto, o tamanho da amostra não apresentou grande impacto.

Para escalas desbalanceadas houve superestimação das cargas fatoriais positivas (i.e., itens em maior quantidade) e sobretudo, subestimação das cargas fatoriais negativas (i.e., itens em menor quantidade). A subestimação, portanto, esteve associada às cargas dos itens negativos. Uma explicação plausível para este fato é que a aquiescência pode inflar as correlações entre itens redigidos na mesma direção e suprimir correlações entre itens de direções opostas (Danner et al., 2015; Kam & Meyer, 2015; Kuru & Pasek, 2016; Lechner & Rammstedt, 2015). Ressalta-se que apenas inverter o desbalanceamento, ou seja, ter mais itens negativos do que positivos, só irá inverter o problema. Em escalas desbalanceadas, sobretudo, o controle de aquiescência é fundamental para que os itens (positivos ou negativos) que estiverem em menor quantidade no instrumento não tenham suas cargas fatoriais suprimidas pela aquiescência.

As condições com cargas fatoriais de conteúdo (do traço psicológico) mais baixa, de modo geral, apresentaram maior viés na estimação, principalmente para escalas desbalanceadas

e sem nenhum método de controle de aquiescência. Pesquisas anteriores apontaram para a influência da aquiescência sobre a estrutura interna dos instrumentos (Kam & Meyer, 2015; Kuru & Pasek, 2016; Lechner & Rammstedt, 2015; Valentini, 2017). O presente estudo avança nesta literatura e apresenta evidências que algumas estruturas podem sofrer mais do que outras. Itens de menor qualidade (i.e., com menor relação com o traço latente) sofrem mais com o viés de aquiescência, ainda que a variância da aquiescência, em si, seja a mesma. Isso ocorre pelo fato de um item de menor qualidade ter menos variância do item sendo explicada pelo fator de conteúdo e conseqüentemente uma proporção considerável sendo explicada pelo viés de resposta aquiescente.

Em que pese à forte recomendação da modelagem da aquiescência para instrumentos com cargas fatoriais baixas, ressaltamos as diferenças de performance. Nestes casos, a ausência de controle é fortemente desaconselhada por enviesar substancialmente as cargas fatoriais, tanto para escalas desbalanceadas, quanto para escalas balanceadas. Em alguns casos, o viés ultrapassou 200%, na média. As conclusões do MIMIC e do Interceptos Randômicos variam: Intercepto Randômico parece resolver o problema do viés na estimação das cargas fatoriais baixas em ambas as escalas desbalanceadas e balanceadas, enquanto o controle por meio do MIMIC parece resolver apenas em escalas balanceadas. Isso pode ter acontecido em função do cálculo do indicador de aquiescência utilizando no MIMIC. Para escalas desbalanceadas, utilizamos um número reduzido de pares opostos para o cálculo do indicador de aquiescência. Isso pode ter diminuído consideravelmente a precisão do indicador de aquiescência.

Os resultados para as condições de cargas fatoriais altas no conteúdo são menos heterogêneos. Neste caso, observamos baixa performance apenas na ausência de controle em escalas desbalanceadas. Interceptos Randômicos e MIMIC não apresentaram vieses significativos em nenhuma das condições em que as cargas fatoriais eram altas. Isso era esperado, pois quando maior for a carga de conteúdo, maior a precisão do escore fatorial.

Portanto, cargas fatoriais altas podem servir como proteção à aquiescência, desde que o viés de resposta seja modelado ou que a escala seja balanceada. Isso também oferece sustentação adicional a nossa conclusão de que algumas estruturas podem sofrer mais ou menos com aquiescência.

O tamanho da aquiescência também foi relevante para a performance da modelagem. No geral, alta aquiescência (i.e., variância de 0,20), sem o controle, implicou maiores problemas nas cargas de conteúdo, mesmo para escalas balanceadas. Esse resultado é uma evidência adicional à literatura sobre a força da influência da aquiescência na piora da estrutura fatorial (Kam & Meyer, 2015; Kuru & Pasek, 2016; Lechner & Rammstedt, 2015; Valentini, 2017).

Por outro lado, o tamanho da amostra não parece exercer grande influência no viés gerado pela aquiescência. Entende-se, portanto, que caso o modelo tenha poder estatístico suficiente, ou seja um número amostral suficiente, a aquiescência parece não exigir um número amostral maior. Apesar disso, problemas com a identificação do modelo foram mais comuns em amostras pequenas, tornando-se um problema não apenas para o controle da aquiescência, mas para verificação da estrutura interna do instrumento. O tamanho da amostra irá depender da complexidade do modelo testado, portanto não existe um número mínimo de tamanho amostral padrão (Muthén & Muthén, 2002). Para verificar o tamanho amostral ideal, é indicado uma simulação simples tendo como base os modelos que serão testados.

Sobre a ilustração com os dados reais do BFI-2, percebemos que os controles de aquiescência testados aumentaram levemente o ajuste dos modelos. Ainda que o ajuste tenha sido aquém do esperado, esse resultado não é diferente do encontrado na literatura (Soto & John, 2017). Isso pode ser atribuído parcialmente à complexidade dos modelos fatoriais para a personalidade. Normalmente, itens nesse contexto são multidimensionais (i.e., contêm cargas cruzadas); e a estimação de modelos muito restritos, sem cargas cruzadas, acaba reduzindo bastante o ajuste (Aichholzer, 2014).

No que diz respeito às cargas fatoriais do BFI-2, os resultados com dados reais são semelhantes às simulações. Alguns itens negativos, aparentemente, tiveram suas cargas subestimadas, principalmente na ausência de controle na composição desbalanceada da escala. Contudo, para a composição balanceada da escala, não observamos grandes diferenças nas cargas fatoriais de conteúdo após o controle da aquiescência, seja por MIMIC ou Interceptos Randômicos. Destacamos que a amostra foi composta por adultos, e a variância da aquiescência foi baixa (i.e., 0,04 e 0,05). Soto et al. (2008) evidenciaram que a variância da aquiescência tende a ser superior na fase da infância e adolescência (i.e., entre 10 e 18 anos) e parece se estabilizar na fase adulta (i.e., 19 e 20 anos).

Considerações finais

Este estudo reúne evidências sobre a importância do balanceamento de itens e do controle da aquiescência para a estimação de modelos fatoriais. Destacamos que a performance de escalas balanceadas foi melhor, e que é necessário modelar a aquiescência (por métodos como MIMIC e Interceptos Randômicos).

Vale destacar aqui um ponto importante referente a algumas recomendações feitas a respeito do não uso de itens negativos em instrumentos de autorrelato, argumentando que estes são inferiores do ponto de vista psicométrico (Checa & Espejo, 2018; Gehlbach & Artino Junior, 2018; Lai, 1994; Ray et al., 2015; Salazar, 2015; Sliter & Zickar, 2014). Notamos a partir dos resultados obtidos neste estudo, que os itens negativos não demonstraram ser inferiores aos positivos após ter sido controlado o viés da aquiescência, assim como apontam outros autores (Marsh, 1996; Peterson & Peterson, 1976; Primi et al., 2019a; Tamir, 1993). Vale ainda lembrar que tal controle somente pode ser operacionalizado em instrumentos de autorrelato com chaves de resposta no formato Likert, mediante a presença de itens positivos e negativos. Portanto, ao eliminar os itens negativos, erroneamente, por parecerem piores

psicometricamente, o pesquisador também está inviabilizando a possibilidade do controle da aquiescência, podendo estar contribuindo assim para uma estrutura e escores enviesados.

Como limitação, podemos destacar que testamos uma única estrutura fatorial, formada por cinco fatores. Além disso, adotamos apenas uma condição de desbalanceamento, sendo ela igual em todos os fatores (i.e., um item negativo e cinco positivos). Não testamos também os vieses da aquiescência em modelos formados somente com itens positivos. Além disso, não testamos eventuais diferenças de erros padrões em cada condição; o que poderia indicar maior instabilidade na estimação dos parâmetros de algumas condições. Ademais, não consideramos a possibilidade de *misfit* dos modelos, o que pode limitar a aderência das simulações à realidade. Consideramos que todas essas questões não abordadas neste artigo podem gerar distinções nos vieses da aquiescência. Portanto, sugerimos que novos estudos sejam realizados testando o viés da aquiescência em instrumentos que abordem essas diferentes condições de estrutura e balanceamento, incluindo amostras pequenas (i.e., menor que 150 casos). Por fim, a amostra dos dados reais utilizada foi composta por adultos, sendo relatado na literatura como um grupo que tende a apresentar menor aquiescência do que adolescentes e crianças. Sendo assim, sugerimos que estudos futuros testem empiricamente os vieses da aquiescência em instrumentos balanceados e desbalanceados a partir de amostras de crianças ou adolescentes.

Em que pese às limitações, os dados apresentados no estudo são suficientes para sustentar a recomendação de construir instrumentos com itens positivos e negativos, preferencialmente balanceado. Destacamos que a literatura apresenta pouca evidência do efeito do balanceamento de itens na construção de escalas de autorrelato. Ademais, sugerimos a utilização de alguma modelagem da aquiescência (por MIMIC ou Interceptos Randômicos) no momento de analisar a estrutura fatorial do instrumento, para assim evitar cometer erros na interpretação da estrutura fatorial. A Tabela 6 pode funcionar como um guia para pesquisadores

na tomada de decisão para desenvolvimento de escalas e modelagem de vieses, verificando qual o melhor modelo para as características específicas da escala que está sendo desenvolvida.

Tabela 6

Recomendações de análises para as condições testadas neste artigo

| Condições | | | Desbalanceado | | | Balanceado | | |
|-----------|---------|---------|---------------|-----------|-----------|---------------|-------|----|
| Cargas | Var Aqu | Amostra | Sem controle | MIMIC | IR | Sem controle | MIMIC | IR |
| Baixas | 0,20 | 200 | Viés grave | Viés leve | Viés leve | Viés grave | | |
| Baixas | 0,20 | 500 | Viés grave | Viés leve | | Viés grave | | |
| Baixas | 0,20 | 1000 | Viés grave | Viés leve | | Viés grave | | |
| Baixas | 0,10 | 200 | Viés grave | Viés leve | Viés leve | Viés moderado | | |
| Baixas | 0,10 | 500 | Viés grave | Viés leve | | Viés moderado | | |
| Baixas | 0,10 | 1000 | Viés grave | Viés leve | | Viés moderado | | |
| Baixas | 0,05 | 200 | Viés moderado | Viés leve | | | | |
| Baixas | 0,05 | 500 | Viés moderado | Viés leve | | | | |
| Baixas | 0,05 | 1000 | Viés moderado | Viés leve | | | | |
| Altas | 0,05 | 200 | Viés leve | | | | | |
| Altas | 0,05 | 500 | Viés leve | | | | | |
| Altas | 0,05 | 1000 | Viés leve | | | | | |
| Altas | 0,10 | 200 | Viés leve | | | | | |
| Altas | 0,10 | 500 | Viés leve | | | | | |
| Altas | 0,10 | 1000 | Viés leve | | | | | |
| Altas | 0,20 | 200 | Viés moderado | | | | | |
| Altas | 0,20 | 500 | Viés moderado | | | | | |
| Altas | 0,20 | 1000 | Viés moderado | | | | | |

Nota. Cargas baixas variaram entre 0,30 e 0,50; Cargas altas variaram de 0,60 a 0,80. Desbalanceado refere-se a um instrumento com cinco itens positivos e um negativo por fator. Balanceado refere-se a um instrumento com o mesmo número de itens positivos e negativos por fator, no caso deste estudo foram três positivos e três negativos.

Referências

- Aichholzer, J. (2014). Random intercept EFA of personality scales. *Journal of Research in Personality*, 53, 1–4. <https://doi.org/10.1016/j.jrp.2014.07.001>
- Ambiel, R. A., Moreira, T. D. C., Barros, L. D. O., Martins, G. H., Salvador, A. P., & Wille, B. (2022). Measuring career adaptabilities in the Brazilian context: Development and validation of the CAAS+ C Brazilian Form. *International Journal for Educational and Vocational Guidance*, 1-17. <https://doi.org/10.1007/s10775-022-09523-5>
- Barnette, J. J. (2000). Effects of Stem and Likert Response Option Reversals on Survey Internal Consistency: If You Feel the Need, There is a Better Alternative to Using those Negatively Worded Stems. *Educational and Psychological Measurement*, 60(3), 361-370. <https://doi.org/10.1177/00131640021970592>
- Checa, I., & Espejo, B. (2018). Method Effects Associated with Reversed Items in the 29 Items Spanish Version of Ryff's Well-Being Scales. *Neuropsychiatry (London)*, 8(5), 1533-1540.
https://www.researchgate.net/profile/Begona_Espejo/publication/327751696_Method_Effects_Associated_with_Reversed_Items_in_the_29_Items_Spanish_Version_of_Ryff%27s_Well-Being_Scales/links/5bbb69c2a6fdcc9552d992e4/Method-Effects-Associated-with-Reversed-Items-in-the-29-Items-Spanish-Version-of-Ryffs-Well-Being-Scales.pdf
- Danner, D., Aichholzer, J., & Rammstedt, B. (2015). Acquiescence in personality questionnaires: Relevance, domain specificity, and stability. *Journal of Research in Personality*, 57, 119-130. <https://doi.org/10.1016/j.jrp.2015.05.004>
- Gehlbach, H., & Artino Junior, A. R. (2018). The survey checklist (manifesto). *Academic Medicine*, 93(3), 360-366. <https://doi.org/10.1097/ACM.0000000000002083>

- Geiser, C., Eid, M., & Nussbeck, F. W. (2008). On the meaning of the latent variables in the CT-C(M-1) model: A comment on Maydeu-Olivares and Coffman (2006). *Psychological Methods*, 13(1), 49–57. <https://doi.org/10.1037/1082-989X.13.1.49>
- He, J., Bartram, D., Inceoglu, I., van de Vijver, J. R. (2014). Response styles and personality traits: A multilevel analysis. *Journal of Cross-Cultural Psychology*, 45(7), 1028-1045. <https://doi.org/10.1177/0022022114534773>
- Henninger, M. (2019). *Psychometric modeling as a tool to investigate heterogeneous response scale use* [Tese de doutorado, Universität Mannheim]. <https://madoc.bib.uni-mannheim.de/52490/>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- John, O. P. (2021). History, Measurement, and Conceptual Elaboration of the Big-Five Trait Taxonomy: The Paradigm Matures. Em O. P. John, & R. W. Robins (Eds.), *Handbook of personality: Theory and Research*, (pp. 35-82). Guilford Press.
- Kam, C. C. S., & Meyer, J. P. (2015). How careless responding and acquiescence response bias can influence construct dimensionality: The case of job satisfaction. *Organizational Research Methods*, 18(3), 512-541. <https://doi.org/10.1177/1094428115571894>
- Kam, C. C. S., & Zhou, M. (2015). Does acquiescence affect individual items consistently?. *Educational and psychological measurement*, 75(5), 764-784. <https://doi.org/10.1177/0013164414560817>
- Kuru, O., & Pasek, J. (2016). Improving social media measurement in surveys: Avoiding acquiescence bias in Facebook research. *Computers in Human Behavior*, 57, 82-92. <https://doi.org/10.1016/j.chb.2015.12.008>

- Lai, J. C. (1994). Differential predictive power of the positively versus the negatively worded items of the Life Orientation Test. *Psychological Reports*, 75(3), 1507-1515.
<https://doi.org/10.2466/pr0.1994.75.3f.1507>
- Lechner, C. M., & Rammstedt, B. (2015). Cognitive ability, acquiescence, and the structure of personality in a sample of older adults. *Psychological assessment*, 27(4), 1301-1310.
<https://doi.org/10.1037/pas0000151>
- Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior research methods*, 48(3), 936-949. <https://doi.org/10.3758/s13428-015-0619-7>
- Lorenzo-Seva, U., & Ferrando, P. J. (2009). Acquiescent responding in partially balanced multidimensional scales. *British Journal of Mathematical and Statistical Psychology*, 62(2), 319–326. <https://doi.org/10.1348/000711007X265164>
- Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors?. *Journal of Personality and Social Psychology*, 70(4), 810-819. <https://doi.org/10.1037/0022-3514.70.4.810>
- Maydeu-Olivares, A., & Coffman, D. L. (2006). Random intercept item factor analysis. *Psychological methods*, 11(4), 344-362. <https://doi.org/10.1037/1082-989X.11.4.344>
- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 9(4), 599-620.
https://doi.org/10.1207/S15328007SEM0904_8
- Paulhus, D. L. (1991). Measurement and control of response bias. Em J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of social psychological attitudes: Measures of personality and social psychological attitudes* (pp. 17-59). Academic Press.

Peterson, C. C., & Peterson, J. L. (1976). Linguistic determinants of the difficulty of true-false test items. *Educational and Psychological Measurement*, 36(1), 161-164.

<https://doi.org/10.1177/001316447603600115>

Plieninger, H. (2018). *Towards a deeper understanding of response styles through psychometrics* [Tese de doutorado, Universität Mannheim]. <https://ub-madoc.bib.uni-mannheim.de/44325>

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903. <https://doi.org/10.1037/0021-9010.88.5.879>

Primi, R., De Fruyt, F., Santos, D., Antonoplis, S., & John, O. P. (2019a). True or False? Keying Direction and Acquiescence Influence the Validity of Socio-Emotional Skills Items in Predicting High School Achievement. *International Journal of Testing*, 1-25. <https://doi.org/10.1080/15305058.2019.1673398>

Ray, J. V., Frick, P. J., Thornton, L. C., Steinberg, L., & Cauffman, E. (2015). Positive and Negative Item Wording and Its Influence on the Assessment of Callous-Unemotional Traits. *Psychological Assessment*, 28(4), 394-404. <https://doi.org/10.1037/pas0000183>

Rosseel, Y. (2012). Quantitative aspects of blood flow and oxygen uptake in the human forearm during rhythmic exercise. *Journal of Statistical Software*, 48(2), 1-93. <https://pubmed.ncbi.nlm.nih.gov/5329335/>

Salazar, M. S. (2015). The dilemma of combining positive and negative items in scales.

Psicothema, 27(2), 192-199. <https://doi.org/10.7334/psicothema2014.266>

Siqueira, M. M. M. (2008). Satisfação no trabalho. Em M. M. M. Siqueira (Org), *Medidas do comportamento organizacional: ferramentas de diagnóstico e gestão* (pp. 265-274).

Artmed.

- Sliter, K. A., & Zickar, M. J. (2014). An IRT examination of the psychometric functioning of negatively worded personality items. *Educational and Psychological Measurement*, 74(2), 214-226. <https://doi.org/10.1177/0013164413504584>
- Sliter, K. A., & Zickar, M. J. (2014). An IRT Examination of the Psychometric Functioning of Negatively Worded Personality Items. *Educational and Psychological Measurement*, 74(2), 214-226. <https://doi.org/10.1177/0013164413504584>
- Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, 113, 117-143. <https://doi.org/10.1037/pspp0000096>
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of big five self-reports: acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, 94(4), 718-737. <https://doi.org/10.1037/0022-3514.94.4.718>
- Tamir, P. (1993). Positive and Negative Multiple Choice Items: How Different Are They?. *Studies in Educational Evaluation*, 19(3), 311-325. Tamir, P. (1993). Positive and negative multiple choice items: How different are they? *Studies in Educational Evaluation*, 19(3), 311-325. [https://doi.org/10.1016/s0191-491x\(05\)80013-6](https://doi.org/10.1016/s0191-491x(05)80013-6)
- Valentini, F. (2017). Editorial: Influência e controle da aquiescência na análise fatorial. *Avaliação Psicológica*, 16(2). <http://dx.doi.org/10.15689/ap.2017.1602.ed>
- Valentini, F., & Hauck Filho, N. (2020). O impacto da aquiescência na estimação de coeficientes de validade. *Avaliação Psicológica*, 19(1), 1-3. <http://dx.doi.org/10.15689/ap.2020.1901.ed>

- Wetzel, E., & Carstensen, C. H. (2015). Multidimensional modeling of traits and response styles. *European Journal of Psychological Assessment, 33*(5), 352–364. <https://doi.org/10.1027/1015-5759/a000291>
- Wetzel, E., Böhnke, J. R., & Brown, A. (2016). Response biases. Em F. T. L. Leong, D. Bartram, F. M. Cheung, K. F. Geisinger, & D. Iliescu (Eds.), *The ITC international handbook of testing and assessment* (pp. 349–363). Oxford University Press.
- Zanon, C., Bardagi, M. P., Layous, K., & Hutz, C. S. (2014). Validation of the Satisfaction with Life Scale to Brazilians: Evidences of measurement noninvariance across Brazil and US. *Social Indicators Research, 119*, 443-453. <https://doi.org/10.1007/s11205-013-0478-5>
- Zanon, C., Lessa, J. P. A., & Dellazzana-Zanon, L. L. (2018). Aquiescência em autorrelatos de personalidade: uma comparação de métodos. *Avaliação Psicológica, 17*(4), 428-438. <http://dx.doi.org/10.15689/ap.2018.1704.3.03>

Contribuição de autoria

Gustavo Henrique Martins: conceitualização, análise de dados, escrita do artigo

Felipe Valentini: conceitualização, análise de dados, revisão do artigo, supervisão

Bruno Bonfá-Araujo: análise de dados, escrita do artigo

Adriana Satico Ferraz: escrita do artigo

Fernanda Otoni: escrita do artigo

André Pereira Gonçalves: escrita do artigo

Conflito de interesses

Os autores declaram não possuir nenhum conflito de interesses relacionado ao presente artigo.

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença [Creative Commons CC-BY](#).
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.