

Invariância Longitudinal e Entre Sexos no Questionário de Gratidão: Exemplificando o Processo

Júlio Frota Lisboa Pereira de Souza

Cristian Zanon

RESUMO

Análises de invariância permitem a comparação das propriedades psicométricas de um instrumento (e.g., cargas fatoriais, interceptos e erros de medida) entre grupos ou momentos. O presente estudo analisou a invariância entre grupos (e.g., homens e mulheres) e ao longo do tempo (intervalo de 5 meses entre as coletas) do Questionário de Gratidão. A amostra utilizada entre os grupos foi composta por 1.850 participantes com idade média de 25,13 anos ($DP = 5,36$), sendo 50% do sexo feminino e 50% do sexo masculino. A amostra utilizada ao longo do tempo contou com 617 participantes, todas do sexo feminino, com média de idade de 25,4 anos ($DP = 5,6$). Comparações de modelos aninhados e hierarquicamente restritivos indicaram alta invariância (e.g., configural, métrico, escalar e residual) para o questionário, indicando robustez na medida. Ademais, este estudo apresentou e exemplificou a importância das análises de invariância para a pesquisa em Psicologia.

Palavras-chave: gratidão; adaptação; psicometria, invariância; personalidade.

ABSTRACT

Longitudinal and Gender Invariance in the Gratitude Questionnaire: Exemplifying the Process

Invariance analyses make it possible to compare the psychometric properties (e.g., factor loadings, intercepts, and measurement errors) between groups or moments. The present study analyzed the invariance between groups (e.g., men and women) and over time (with a 5-month interval between data collections) of the Gratitude Questionnaire. The sample used between the groups consisted of 1850 participants with a mean age of 25.13 years ($SD = 5.36$), 50% female and 50% male. The sample used in longitudinal part of this study consisted of 617 female participants with a mean age of 25.4 years ($SD = 5.6$). Comparisons of nested and hierarchically restrictive models indicated high invariance (e.g., configural, metric, scalar, and residual) for the questionnaire, indicating robustness in the measure. Furthermore, this study presented and exemplified the importance of invariance analyses for psychology research.

Keywords: gratitude; adaptation; psychometrics; invariance; personality.

Sobre os Autores

J. F. L. P. S.
orcid.org/0000-0003-3822-5275
Universidade Federal do Rio
Grande do Sul (UFRGS) – Porto
Alegre, RS
julioflps@gmail.com

C. Z.
orcid.org/0000-0003-3822-5275
Universidade Federal do Rio
Grande do Sul (UFRGS) – Porto
Alegre, RS
autoroliveira@email.com

Direitos Autorais

Este é um artigo aberto e pode ser reproduzido livremente, distribuído, transmitido ou modificado, por qualquer pessoa desde que usado sem fins comerciais. O trabalho é disponibilizado sob a licença Creative Commons CC-BY-NC.



A gratidão tem sido, frequentemente, definida como uma disposição que indica “uma tendência geral de reconhecer e responder com gratidão aos papéis de benevolência de outras pessoas ou eventos em experiências pessoais e resultados positivos obtidos” (McCullough et al., 2002, p. 112; tradução nossa). Muitos estudos mostram que a gratidão está positivamente relacionada à autoestima, ao apoio social e ao bem-estar subjetivo, e negativamente relacionada a sintomas depressivos (Emmons & McCullough, 2003; Wood et al., 2010; Hill et al., 2013; Kong et al., 2015). A literatura aponta para a gratidão como um fenômeno importante para o aumento do bem-estar, o que justifica a relevância de estudá-la e medi-la.

Dada a importância da gratidão para o convívio social e para o bem-estar, intervenções foram criadas para o seu desenvolvimento através de diversas técnicas (Dickens, 2017). Contudo, para que os resultados de pesquisas sejam válidos e fidedignos, é necessário que os instrumentos usados apresentem propriedades psicométricas adequadas. Uma vez que comparações de homens e mulheres são comumente realizadas, assim como testes que comparam pré e pós-intervenções, faz-se relevante investigar se o teste usado apresenta equivalência entre os homens e as mulheres e se preserva suas propriedades psicométricas (e.g., estrutura, cargas fatoriais, interceptos e erros de medida) ao longo do tempo. Este manuscrito avalia a adequação de comparações de gratidão entre e intragrupos com o Questionário de Gratidão (QG-6) no contexto brasileiro. Ademais, busca-se apresentar a relevância sobre investigações de invariância de forma didática para encorajar mais estudos com outros instrumentos de medida na área da Avaliação Psicológica.

GRATIDÃO

Além de relacionada a vários indicadores de bem-estar como autoestima, esperança e otimismo (Layous et al., 2017), a gratidão caracteriza-se como uma tendência ou orientação no sentido de reconhecer e responder com a emoção positiva da gratidão aos outros e ao mundo em geral (Wood et al., 2010). Esse entendimento do construto é apoiado por evidências de diversos estudos que avaliaram a relação desse traço com as dimensões da teoria dos cinco grandes fatores de personalidade, demonstrando sua validade convergente para predição de bem-estar (Wood et al., 2010). Mais especificamente, verificou-se que indivíduos com alta gratidão disposicional¹ tendem a ser mais extrovertidos, sociáveis, abertos a experiências e conscienciosos, além de apresentarem menor neuroticismo (McCullough et al., 2002; McCullough et al., 2004). Além disso, a gratidão é diferente de outras características positivas. Por exemplo, a gratidão como traço é orientada para perceber resultados positivos na vida, enquanto o otimismo é orientado para esperar resultados futuros positivos (Wood et al., 2010).

Com o objetivo de avaliar a gratidão de forma válida e rápida, desenvolveu-se o Questionário de Gratidão (GQ-6;

[McCullough et al., 2002]). O GQ-6 é um instrumento de seis itens que mede o construto da gratidão disposicional (McCullough et al., 2002) e tornou-se o mais utilizado para avaliar a gratidão em populações saudáveis em todo o mundo. No estudo de McCullough et al. (2002), a análise fatorial confirmatória (CFA) mostrou que existia uma estrutura robusta de fator único do GQ-6, replicada em outros estudos (Froh et al., 2009; Yüksel & Oğuz Duran, 2012; Caputo, 2016; Langer et al., 2016). Ademais, McCullough et al. (2002) encontraram evidências de que o GQ-6 está positivamente relacionado com satisfação de vida, bem-estar e características pró-sociais, e negativamente relacionado com sintomas psicológicos e materialismo. Em suma, os resultados sugerem que o GQ-6 possui boas propriedades psicométricas internacionalmente e no Brasil ($\chi^2(9) = 59, p < 0,001$; CFI = 0,956, RMSEA = 0,078; [Froh et al., 2009; Yüksel & Oğuz Duran, 2012; Caputo, 2016; Langer et al., 2016]). Este instrumento foi adaptado para o português pelos autores deste estudo como parte da tese de doutorado do primeiro autor, sendo que o estudo de adaptação se encontra em desenvolvimento.

INVARIÂNCIA DA MEDIDA

A maior parte dos estudos que utilizam instrumentos psicométricos assume que uma determinada pontuação em um teste representa o mesmo nível do construto subjacente para diversos grupos e que prediz o mesmo escore, independentemente de outros fatores (Chen, 2008). Os estudos de invariância investigam justamente essa premissa. Entretanto, é possível que um teste seja enviesado em função da sua própria estrutura. Diferenças no funcionamento do teste para grupos distintos ou ao longo do tempo tornam os resultados pouco confiáveis e as comparações em intervenções e entre grupos enviesadas.

Um teste tendencioso “é aquele que superestima ou subestima sistematicamente o valor da variável que se pretende avaliar” em função de um indivíduo ser parte de um grupo, como etnia ou gênero (Reynolds & Suzuki, 2013, p. 94). Esse fenômeno refere-se ao que às vezes é chamado de viés de medição, no qual a relação entre os escores dos testes e o atributo latente medido pelas pontuações varia para diferentes grupos (Borsboom et al., 2008; Millsap, 1997).

VIÉS DE TESTE

Se as pontuações dos testes puderem ser usadas para prever algum resultado futuro (e.g., um critério), é possível afirmar que indicam validade preditiva (Clearly, 1968). Em geral, um teste é considerado não tendencioso (i.e., demonstrando “invariância preditiva”) se suas pontuações previrem resultados futuros igualmente bem para indivíduos de diferentes grupos. Essa hipótese consiste em estimar se os escores do Questionário de Gratidão (GQ-6) predizem mudanças em escores de satisfação de vida em homens de forma similar em mulheres, ao se calcular regressões.

No entanto, se os resultados dos testes são melhores em prever resultados para alguns grupos do que para outros, diz-se que os resultados dos testes refletem validade preditiva diferencial (Camilli, 2006). Se as pontuações dos testes superestimam ou subestimam sistematicamente os critérios de um grupo em relação a outro, diz-se que os resultados refletem viés de intercepto (i.e., *interception bias*).

Nos últimos anos, a Invariância de Medida (IM) foi investigada com mais frequência e rigor do que a invariância preditiva devido a avanços nos procedimentos estatísticos, incluindo o desenvolvimento de softwares que permitem aos pesquisadores realizar testes de IM com mais facilidade. Isso também se deve ao fato de que a facilidade de avaliar o viés preditivo varia muito, dependendo da integridade psicométrica da variável critério em questão. É possível que a validade preditiva diferencial reflita um artefato estatístico devido ao erro de medição nas pontuações no critério e no preditor (Kane & Mroch, 2010; Warne et al., 2014). Além disso, outras variáveis de confusão (e.g., intervalo de tempo entre a obtenção de dados de teste e critério) precisam ser controladas entre os grupos ao examinar a invariância preditiva.

Essas complicações são problemáticas, sobretudo em pesquisas interculturais que envolvem medidas traduzidas. Portanto, argumenta-se que os psicólogos deveriam favorecer o teste de IM em vez do teste preditivo de invariância na definição e mensuração do viés (Borsboom et al., 2008). Por isso, o restante desse artigo focará na IM.

NÍVEIS DE INVARIÂNCIA

A IM costuma ser avaliada usando análise fatorial confirmatória com múltiplos grupos. Identifica-se, primeiro, um modelo fatorial bem ajustado para servir como “linha de base” e, em seguida, aplica-se uma série de modelos estatísticos mais restritivos nas amostras de interesse para se investigar o nível de invariância a ser alcançado. Estes modelos refletem uma ordem hierárquica de equivalência de parâmetros psicométricos como: (a) invariância configural (i.e., estrutura ou forma das ligações causais dos fatores aos itens); (b) invariância métrica (i.e., magnitude e direção das relações causais entre os fatores e os indicadores); (c) invariância escalar (i.e., interceptos/limiares em relação ao ponto de origem); (d) invariância única ou residual, (i.e., precisão de cada escala). Esses níveis de invariância envolvem diferentes parâmetros de medição do modelo e são considerados tipos de IM ou equivalência da medida.

A invariância configural (Meredith, 1993; Vandenberg & Lance, 2000) constitui o primeiro e o mais baixo nível na sequência de testes de IM. Ela requer uma demonstração de que os mesmos itens ou indicadores carreguem os mesmos fatores entre os grupos. O ajuste de um teste de invariância configural é, aproximadamente, a média dos ajustes de CFA para cada grupo considerado separadamente. Por exemplo, se

um modelo de quatro fatores específicos mostrar um bom ajuste separadamente para duas amostras de grupos étnicos, é provável que a invariância configural seja estabelecida, sugerindo que ambos os grupos demonstram o construto de maneira semelhante. Uma vez atingido um modelo invariante configural, este pode ser usado como modelo de linha de base para testes de invariância adicionais (Vandenberg & Lance, 2000).

A invariância métrica, o segundo nível na sequência de teste de IM, exige invariância configural e que os valores das cargas fatoriais para cada item sejam idênticos entre os grupos. Ela é também chamada de invariância fraca, porque caracteriza o nível mínimo de equivalência para que seja considerado invariante (Han et al., 2019). Cargas fatoriais equivalem a coeficientes de regressão que têm como variável independente o fator e como variável dependente o indicador em um modelo de fator comum.

Após constatada a invariância métrica, a invariância escalar (i.e., “invariância forte”) pode ser examinada observando se os interceptos dos itens são idênticos entre os grupos (i.e., a pontuação prevista de um item quando sua pontuação no fator latente é zero; isto é, o ponto de origem de indicadores contínuos) ou limites de classificações a partir dos itens (i.e., o ponto em que as categorias de resposta mudam, por exemplo, de “discordo totalmente” para “discordo em parte” ou de “falso” para “verdadeiro” (Han et al., 2019).

Quando estabelecida a invariância escalar, pode-se considerar que a escala possui as mesmas unidades de medida (e.g., cargas fatoriais) e origem (e.g., interceptos/limiares) para os diversos grupos. Consequentemente, a invariância escalar – igualdade de cargas fatoriais e interceptos de itens dos grupos – é um pré-requisito para a comparação de médias de fatores latentes, pois permite concluir que as diferenças entre grupos nas médias de fatores latentes refletem verdadeiras diferenças médias entre grupos em construtos latentes e não erros de medida (Van de Velde et al., 2009).

A variação residual refere-se à variação de um indicador que não é atribuível à variação da variável latente associada (Cheung & Rensvold, 2002). A invariância residual implica que, além das cargas e interceptos invariantes, a exclusividade da variância comum do item, a precisão ou o erro de medição também são invariáveis entre os grupos. A invariância residual, também conhecida como “invariância estrita”, constitui um modelo altamente rígido que não é comumente encontrado em análises na prática.

É normal que não seja encontrada invariância residual em função de problemas nos erros do teste, nos resíduos. Diferenças na familiaridade com o formato das respostas, a estrutura da sentença do item, o uso de expressões idiomáticas ou experiências culturais podem contribuir para a variação residual (Cheung & Rensvold, 2002). A importância da

invariância residual deve-se ao fato desta permitir afirmar com confiança que diferenças na pontuação média entre os grupos são efetivamente diferenças reais entre os grupos no construto e não ocasionadas por outros fatores (Meredith, 1993). Entretanto, a invariância escalar é considerada, muitas vezes, suficiente para permitir a comparação relevante de diferenças médias latentes, pois os resíduos não fazem parte dos fatores latentes (Chen, 2008).

O estabelecimento de critérios para determinar a invariância é uma tarefa complexa, porque inúmeras características de um modelo precisam ser consideradas. As características dos indicadores (e.g., contínuos, ordenados categóricos ou binários) e das distribuições (e.g., grau de normalidade multivariada) podem exigir diferentes tipos de métodos de comparação, que, por sua vez, também exigem diferentes abordagens para avaliar o ajuste do modelo. Atualmente, a abordagem mais usada para modelos com indicadores categóricos, contínuos ou ordenados que não apresentam desvios sérios da normalidade multivariada é o método de *maximum likelihood estimation* (Han et al., 2019).

Uma vez que comparações de grupo são desencorajadas até que a IM de um instrumento seja estabelecida, os pesquisadores às vezes tratam a IM como um obstáculo que deve ser superado (Meade & Bauer, 2007). Consequentemente, a avaliação da IM nem sempre é realizada com o cuidado e a consideração que os pesquisadores trazem aos seus interesses primários de pesquisa, e os resultados nem sempre são tão informativos quanto poderiam ser.

SOBRE INVARIÂNCIA DE SEXOS NA GRATIDÃO

Outra questão importante ao usar o instrumento GQ-6 é a invariância da medida em relação ao sexo. Por exemplo, estudos anteriores compararam as diferenças de escores no GQ-6 entre homens e mulheres (Kashdan et al., 2009; Sun & Kong, 2013; Kong et al., 2015). Kashdan et al. (2009) descobriram que as mulheres adultas pontuaram mais no GQ-6 do que homens adultos (Kashdan et al., 2009; Sun & Kong, 2013). Em outro estudo, Sun e Kong (2013) descobriram que as mulheres tiveram maior gratidão que os homens no final da adolescência (Sun & Kong, 2013). No entanto, ao testar a diferença de sexo em escores GQ-6, temos que garantir que o GQ-6 mede a mesma estrutura latente no sexo. Ou seja, qualquer diferença de sexo nas pontuações GQ-6 é devida a variações quantitativas reais e não resultado do viés de medição nos grupos de homens e mulheres. Dessa forma, outro objetivo deste estudo foi examinar se a estrutura de medição subjacente à gratidão é invariante para sexo.

SOBRE INVARIÂNCIA TEMPORAL/LONGITUDINAL

Um dos objetivos das pesquisas sobre desenvolvimento é examinar mudanças individuais em construtos ao longo do tempo. A precisão dos modelos que endereçam essas

questões de pesquisa depende da suposição de invariância longitudinal da medição: as variáveis medidas repetidamente precisam representar o mesmo construto na mesma métrica ao longo do tempo. A invariância da medida pode ser estudada através de modelos de fatores que examinam as relações entre os indicadores observados e os construtos latentes. Na pesquisa longitudinal, são geralmente usados indicadores categóricos ordenados, como escalas de tipo Likert de autorrelato ou relatadas por terceiros. Infelizmente, tais medidas frequentemente não se aproximam das distribuições normais contínuas presumidas nas comparações (Han et al., 2019).

O PRESENTE ESTUDO

Este estudo investiga a possível existência de diferenças de gratidão entre homens e mulheres no contexto brasileiro. Essas diferenças são relevantes porque podem indicar a necessidade de intervenções de gratidão distintas para cada sexo. Considerando que a gratidão tem sido uma variável bastante usada para promoção de bem-estar (Dickens, 2017, 2019), a compreensão de sua manifestação no contexto brasileiro é essencial. Para isso, buscou-se testar a invariância entre sexos e a estabilidade temporal (e.g., invariância longitudinal) do Questionário de Gratidão (McCullough et al., 2002), instrumento mundialmente usado para possibilitar o estudo da gratidão de forma sistematizada no Brasil. O segundo objetivo deste artigo é oferecer uma visão geral introdutória da análise fatorial confirmatória com múltiplos grupos, dos procedimentos metodológicos relacionados à análise da invariância entre grupos e longitudinalmente e da sua importância para a avaliação psicológica.

Duas hipóteses norteiam este estudo:

H1 - O Questionário de Gratidão apresenta evidências de invariância configural, métrica, escalar e residual entre sexos em sua versão adaptada para o português.

H2 - O Questionário de Gratidão apresenta evidências de invariância configural, métrica, escalar e residual longitudinalmente.

MÉTODO

DELINEAMENTO

Tratou-se de um delineamento longitudinal correlacional com coleta de dados on-line. Os dados foram coletados em dois momentos, com uma diferença de 5 meses entre as datas finais das coletas. O primeiro momento foi realizado em dezembro de 2019, e o segundo momento foi feito em maio de 2020. O período de 5 meses entre as coletas ocorreu em função da mudança de objetivos iniciais deste projeto de pesquisa devido ao início da pandemia. Ademais, este período é adequado por reduzir a lembrança dos itens e das respostas dos participantes na primeira coleta.

PARTICIPANTES

Visando uma melhor organização, a descrição dos participantes será dividida entre o primeiro e o segundo momento de coleta:

Primeiro momento. A amostra inicial contou com 9.845 participantes, tendo idades variando de 18 a 86 anos ($M = 26,3$; $DP = 7,5$) e sendo 89% mulheres. Entretanto, para produzir uma amostra mais homogênea, foi realizada uma limpeza do grupo, eliminando os 0,7% da amostra que indicam outro gênero, os 5% que possuíam idade acima de 40 anos e as participantes que não foram sorteadas após ser realizada uma equiparação aleatória do número de mulheres com o número de homens usando a idade como critério. Estes critérios para eliminar participantes foram usados para se obter grupos similares (exceto em relação a sexo) para minimizar os efeitos de variáveis aleatórias nos resultados. Após esses procedimentos, obteve-se uma amostra de 1.850 participantes, sendo 50% do sexo feminino e 50% do sexo masculino, com média de idade de 25,13 anos ($DP = 5,36$).

Segundo momento. Para o segundo momento, convidamos todos os participantes da primeira aplicação para uma nova coleta. Com isso, obtivemos 767 novas respostas. Em seguida, também para produzir uma amostra mais homogênea e semelhante, foi realizada uma limpeza do grupo, sendo eliminados todos os participantes que tinham acima de 40 anos de idade. Desse modo, a amostra final contou com 617 participantes, sendo todas do sexo feminino e tendo média de idade de 25,4 anos ($DP = 5,6$). Em relação à raça autodeclarada, 61% era branca, 26% parda, 9% negra, 2% amarela.

INSTRUMENTO

Questionário de Gratidão (GQ-6): O GQ-6 inclui seis itens classificados em uma escala do tipo Likert de 7 pontos, variando de 1 (discordo totalmente) a 7 (concordo totalmente). Alguns exemplos de itens são: "Sinto-me agradecido pelo que recebi na vida" e "Às vezes me sinto grato para as menores coisas".

PROCEDIMENTOS

Tanto para o primeiro quanto para o segundo momento de coleta de dados, os participantes responderam ao questionário em formato on-line e nas mesmas condições. Foi utilizada a plataforma *SurveyAnyplace*, na qual foi desenvolvido o questionário e coletado os dados. O convite para a participação na pesquisa foi feito através das redes sociais. Quando os participantes acessavam a página da pesquisa, logo era apresentado o Termo de Consentimento Livre e Esclarecido, sendo permitido o prosseguimento da coleta somente após o aceite do indivíduo. A pesquisa obteve aprovação do Comitê de Ética em Pesquisa sob o CAEE: 14351319.0.0000.5334.

ANÁLISES DE DADOS

Para avaliar a invariância do modelo unidimensional entre homens e mulheres, procederam-se as análises fatoriais confirmatórias multi-grupo. Para avaliar a invariância longitudinal do modelo em duas ocasiões (i.e., antes e durante a pandemia), foram realizadas análises fatoriais confirmatórias. Em ambos os procedimentos, testaram-se os níveis configural (i.e., mesma estrutura fatorial), métrico (i.e., mesmas cargas fatoriais), escalar (i.e., mesmas cargas fatoriais e interceptos) e residual (i.e., mesmas cargas fatoriais, interceptos e variância dos erros), nessa sequência. Estes níveis de invariância representam modelos aninhados e hierárquicos que se tornam, gradualmente, mais restritivos, na medida em que novos parâmetros vão sendo fixados (e.g., cargas fatoriais, interceptos e variâncias de erro).

Para identificar o grau de invariância, ou equivalência, entre os grupos e nas diferentes ocasiões, utilizou-se o critério da variação do $CFI < 0,01$ e do $RMSEA < 0,015$ (Chen, 2008; Cheung & Rensvold, 2002) entre os modelos testados. Ou seja, quando há redução de CFI superior a 0,01 ou aumento de $RMSEA$ superior a 0,015 de um modelo para outro (i.e., piora no ajuste) o modelo mais restritivo é rejeitado e nos demais casos o modelo mais restritivo é favorecido. O método de estimação utilizado em ambos os procedimentos foi o *Maximum Likelihood with Robust Standard Errors* (MLR) por lidar com variáveis contínuas cujo resíduo pode não ser normalmente distribuído. O programa utilizado para as análises fatoriais confirmatórias foi o MPlus7.11.

RESULTADOS

ANÁLISE DE INVARIÂNCIA ENTRE SEXOS

Inicialmente, o modelo unidimensional foi testado, independentemente, no grupo masculino ($\chi^2 = 59.60$, $df = 9$, $p < .001$, $CFI = .96$, $RMSEA = .08$, $SRMR = .04$) e feminino ($\chi^2 = 54.91$, $df = 9$, $p < .001$, $CFI = .95$, $RMSEA = .08$, $SRMR = .04$). Análises de invariância só devem ser conduzidas quando o modelo testado apresenta ajuste adequado nos grupos de interesse (ver Zanon et al., 2021, para um exemplo em que grupos foram removidos previamente por não apresentar ajuste adequado no modelo testado). Como observado na Tabela 1, os índices de ajuste foram satisfatórios (Cheung & Rensvold, 2002) e todas as cargas fatoriais ($0,43 < \lambda < 0,77$) foram significativas, indicando adequação do modelo unidimensional nos grupos.

Posteriormente, procedeu-se a análise de invariância testando, inicialmente, o modelo configural. Estes resultados indicam que o modelo unidimensional com todos os itens carregando na variável latente estimada é adequado para homens e mulheres. Uma vez que os índices foram adequados (ver Tabela 1), procedeu-se a avaliação do modelo métrico.

Verificou-se melhora nos índices de ajuste, e a variação de índices favorece o nível métrico ($\Delta CFI = 0,001$ e do $\Delta RMSEA = 0,009$). Estes resultados indicam que as cargas fatoriais dos itens de gratidão eram invariantes (i.e., equivalentes) entre homens e mulheres. Obtendo evidências de invariância métrica, testou-se o modelo escalar que apresentou índices satisfatórios e variações em relação ao modelo métrico que suportam o nível escalar ($\Delta CFI = 0,003$ e do $\Delta RMSEA = 0,005$). Estes resultados indicam que os interceptos dos itens, ou médias estimadas, eram invariantes entre os grupos e que

comparações de médias latentes são suportadas. Por fim, testou-se a invariância residual que apresentou índices de ajuste adequados e variações em relação ao modelo escalar satisfatório ($\Delta CFI = 0,002$ e do $\Delta RMSEA = 0,007$).

Estes resultados indicam que as variâncias residuais dos itens também eram invariantes entre os grupos e que comparações de médias são suportadas. Não foram verificadas diferenças nas médias latentes de gratidão entre homens e mulheres.

Tabela 1. Análise Fatorial Confirmatória Multi-Grupo da Escala de Gratidão entre Homens e Mulheres

Modelos aninhados testados	χ^2	gl	CFI	ΔCFI	RMSEA	$\Delta RMSEA$	NParamL
Configural	121	18	0,957	-	0,079	-	36
Métrico	126	23	0,958	0,001	0,070	0,009	31
Escalar	136	28	0,955	0,003	0,065	0,005	26
Residual	140	34	0,957	0,002	0,058	0,007	20

Nota. χ^2 = chi-quadrado, gl = graus de liberdade, CFI = Comparative Fit Index, ΔCFI = diferença de CFI entre os modelos comparados, RMSEA = Root Mean Square Error of Approximation, $\Delta RMSEA$ = diferença de RMSEA entre os modelos comparados, NParamL = Número de parâmetros livremente estimados.

MENSURAÇÃO DE INVARIÂNCIA AO LONGO DO TEMPO

Com o objetivo de analisar as invariâncias da escala entre o grupo do primeiro momento da aplicação (t1) e o grupo do segundo momento (t2), foi realizado primeiramente um teste do modelo unidimensional em ambos os momentos. Os índices de ajustes encontrados foram adequados para ambos (Cheung & Rensvold, 2002), e todas as cargas fatoriais foram significativas, indicando a adequação de um modelo unidimensional nos dois momentos.

A seguir, analisou-se a invariância em ordem hierárquica crescente: configural, métrica, escalar e residual (Tabela 2). Para o modelo configural, os resultados mostraram que o modelo unidimensional com todos os itens carregando na variável latente gratidão é adequado para t1 e t2. Com o modelo métrico, notou-se que as cargas fatoriais dos itens da escala de gratidão foram invariantes entre t1 e t2, pois verificou-se uma melhora nos índices de ajuste e que a variação de índices favorece o nível métrico ($\Delta CFI = 0,000$ e do $\Delta RMSEA = 0,002$).

Tabela 2. Análise de Invariância Longitudinal (Solução Unidimensional) da Escala de Gratidão em Dois Momentos

(t1 e t2)

Modelos aninhados testados	χ^2	gl	CFI	ΔCFI	RMSEA	$\Delta RMSEA$	NParamL
Configural	151	47	0,956	-	0,059	-	43
Métrico	156	52	0,956	0,000	0,057	0,002	38
Escalar	174	58	0,951	0,005	0,056	0,002	32
Residual	190	64	0,946	0,005	0,056	0	26

Já no modelo escalar, os resultados indicaram que os interceptos estimados eram invariantes entre os momentos t1 e t2 e que são suportadas comparações de médias latentes, já que apresentaram índices adequados e variações (em relação ao modelo métrico) que suportam o nível escalar ($\Delta CFI = 0,005$ e do $\Delta RMSEA = 0,002$). Por fim, no modelo residual, os resultados indicaram que as variâncias residuais dos itens também eram invariantes entre os momentos t1 e t2 e que as comparações de médias são suportadas, tendo em vista que apresentaram índices de ajuste adequados e variações (em relação ao modelo escalar) satisfatórias ($\Delta CFI = 0,005$ e do $\Delta RMSEA = 0,000$). Não foram verificadas diferenças significativas nas médias latentes de gratidão entre t1 e t2.

DISCUSSÃO

Este estudo objetivou avaliar a invariância de sexo e longitudinal da versão em português do Questionário de Gratidão de 6 Itens (GQ-6) e fornecer uma descrição minuciosa do procedimento utilizado, com o propósito de promover a realização de mais estudos de invariância no contexto brasileiro. Este trabalho é o primeiro a investigar a invariância da escala e a demonstrar sua equivalência de sexo e estabilidade temporal por meio da invariância longitudinal.

Os resultados deste estudo contribuíram para o estabelecimento de evidências psicométricas sólidas para o GQ-6 no contexto brasileiro, uma vez que as análises revelaram que as invariâncias configural, métrica, escalar e residual foram alcançadas. Além disso, este estudo não identificou diferenças significativas em níveis de gratidão entre indivíduos de diferentes sexos e ao longo do tempo. É importante ressaltar que a pesquisa em instrumentos psicométricos que realize e reporte análises de invariância é escassa, e ainda mais rara é a demonstração de invariância nos níveis configural, métrico, escalar e residual, destacando a relevância deste estudo.

Alguns estudos internacionais sobre gratidão utilizando a GQ-6 têm analisado a invariância de gênero com resultados distintos quanto à existência de diferenças entre homens e mulheres (Rey et al., 2018; Liu et al., 2017; Kong et al., 2017; Valdez et al., 2017; Langer et al., 2016). De forma geral, a literatura aponta para uma maior percepção de gratidão média para mulheres que poderia ser explicada por uma maior disposição a demonstrar emoções (Kong et al., 2017). Contudo, esse resultado não foi replicado no contexto brasileiro, mas poderá ser futuramente investigado com grupos mais homogêneos e controlando variáveis de confusão como bem-estar e psicopatologias.

Não foram encontrados estudos que tenham realizado análises de invariância longitudinal da GQ-6, cujas evidências são especialmente importantes quando o instrumento é utilizado como variável desfecho ou moderadora em intervenções – precisamente o caso da GQ-6 (Davis et al.,

2016). A verificação de invariância estrita longitudinalmente da GQ-6 no Brasil constitui um incentivo à condução de pesquisas sobre gratidão no Brasil, especialmente em intervenções de promoção do bem-estar (Dickens, 2017).

O aumento do interesse sobre a IM pode refletir a globalização das ciências sociais, maior ênfase na avaliação justa para diversas populações e avanços estatísticos que permitem testes rigorosos de IM (Davidov et al., 2014; Sass, 2011; Vandenberg, 2002). As análises de invariância e os níveis aceitáveis (e.g., métrico, escalar ou até mesmo residual) dependem do contexto de uso dos instrumentos. Níveis mais altos de invariância são especialmente importantes para a garantia de diagnósticos psicológicos precisos, avaliações psicológicas, admissões em concursos públicos e seleção no mercado de trabalho.

Já no contexto da pesquisa, sobretudo a transcultural com instrumentos adaptados, a invariância pode ser flexibilizada até certo ponto. Oliveira et al. (2022), por exemplo, apresentam uma descrição detalhada de um processo de adaptação de um instrumento para medir otimismo em crianças e adolescentes brasileiros e conduzem uma análise de invariância que permite a comparação de médias latentes entre os grupos. Este estudo alcançou invariância escalar e detectou o aumento do pessimismo e a redução do otimismo com o passar dos anos.

A principal limitação do presente estudo foi a amostra consideravelmente menor e predominantemente composta por mulheres que respondeu nos dois momentos em relação aos participantes do primeiro levantamento. Isso não permitiu que a invariância longitudinal fosse testada para ambos os sexos. Ademais, é possível que esta amostra apresente um viés de seleção referente a ter participantes mais cooperativos e altruístas que responderam ao questionário duas vezes com o objetivo de obter autoconhecimento (referente ao seu nível de gratidão) e contribuir para o avanço da ciência psicológica.

O estudo da gratidão no contexto brasileiro ganha uma relevância ainda maior diante dos desafios impostos pela pandemia de COVID-19. A crise sanitária ampliou significativamente os índices de depressão, de estresse e de ansiedade em todo o país, destacando a necessidade premente de compreender e promover o bem-estar da população. Nesse contexto, a investigação e a promoção da gratidão assumem um papel crucial, uma vez que está associada a uma série de benefícios para a saúde mental e emocional, incluindo a redução desses sintomas negativos. Além disso, compreender como a gratidão pode ser cultivada mediante o enfrentamento de adversidades pode fornecer insights valiosos para o desenvolvimento de intervenções eficazes que visem mitigar os impactos psicológicos da pandemia no Brasil e promover uma sociedade mais resiliente e saudável.

CONCLUSÃO

Os resultados deste estudo fornecem evidências robustas de equivalência dos parâmetros do QG-6 entre homens e mulheres, mantendo-a constante ao longo do tempo no contexto brasileiro. Essa constatação é de grande importância para a pesquisa, permitindo a utilização do QG-6 de maneira consistente em estudos transversais e longitudinais, bem como no desenvolvimento de intervenções focadas no aprimoramento da gratidão na população brasileira.

CONTRIBUIÇÃO DE CADA AUTOR

Certificamos que todos os autores participaram suficientemente do trabalho para tornar pública sua responsabilidade pelo conteúdo. A contribuição de cada autor pode ser atribuída como se segue:

J.F.L.P.S. foi orientando de doutorado de C.Z. J.F.L.P.S. e C.Z. contribuíram para a conceitualização, investigação, visualização do artigo, redação inicial e final do artigo. C.Z. conduziu as análises estatísticas.

AGRADECIMENTOS

Os autores agradecem à banca de qualificação que indicou melhorias importantes na versão anterior deste manuscrito.

DECLARAÇÃO DE CONFLITOS DE INTERESSES

Os autores declaram que não há conflitos de interesse no manuscrito submetido.

REFERÊNCIAS

- Borsboom, D., Romeijn, J. W., & Wicherts, J. M. (2008). Measurement invariance versus selection invariance: Is fair selection possible? *Psychological Methods*, 13(2), 75–98. <http://dx.doi.org/10.1037/1082-989X.13.2.75>
- Camilli, G. (2006). Test fairness. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 221–256). Praeger.
- Caputo, A. (2016). Italian translation and validation of the GQ-6. *International Journal of Wellbeing*, 6(2), 80–92. <https://doi.org/10.5502/ijw.v6i2.492>
- Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology*, 95(5), 1005–1018. <http://dx.doi.org/10.1037/a0013193>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5
- Clearly, T. A. (1968). Test bias: Prediction of grades of Negro and White students in integrated colleges. *Journal of Educational Measurement*, 5(2), 115–124. <http://dx.doi.org/10.1111/j.1745-3984.1968.tb00613.x>
- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, 40, 55–75. <https://doi.org/10.1146/annurev-soc-071913-043137>
- Davis, D. E., Choe, E., Meyers, J., Wade, N., Varjas, K., Gifford, A., Quinn, A., Hook, J. N., Van Tongeren, D. R., Griffin, B. J., & Worthington, E. L., Jr. (2016). Thankful for the little things: A meta-analysis of gratitude interventions. *Journal of Counseling Psychology*, 63(1), 20–31. <https://doi.org/10.1037/cou000107>
- Dickens, L. R. (2017). Using gratitude to promote positive change: A series of meta-analyses investigating the effectiveness of gratitude interventions. *Basic and Applied Social Psychology*, 39(4), 193–208. <https://doi.org/10.1080/01973533.2017.1323638>
- Dickens, L. R. (2019). Gratitude Interventions: Meta-analytic Support for Numerous Personal Benefits, with Caveats. In L. E. Van Zyl & S. Rothmann (Eds.), *Positive Psychological Intervention Design and Protocols for Multi-Cultural Contexts* (pp. 127–147). Springer, Cham.
- Emmons, R. A., & McCullough, M. E. (2003). Counting blessings versus burdens: An experimental investigation of gratitude and subjective well-being in daily life. *Journal of Personality and Social Psychology*, 84(2), 377–389. <https://doi.org/10.1037/0022-3514.84.2.377>
- Froh, J. J., Yurkewicz, C., & Kashdan, T. B. (2009). Gratitude and subjective well-being in early adolescence: Examining gender differences. *Journal of Adolescence*, 32(3), 633–650. <https://doi.org/10.1016/j.adolescence.2008.06.006>
- Han, K., Colarelli, S. M., & Weed, N. C. (2019). Methodological and statistical advances in the consideration of cultural diversity in assessment: A critical review of group classification and measurement invariance testing. *Psychological Assessment*, 31(12), 1481–1496. <https://doi.org/10.1037/pas0000731>
- Hill, P. L., Allemand, M., and Roberts, B. W. (2013). Examining the pathways between gratitude and self-rated physical health across adulthood. *Personality and Individual Differences*, 54(1), 92–96. <https://doi.org/10.1016/j.paid.2012.08.011>
- Kane, M. T., & Mroch, A. A. (2010). Modeling group differences in OLS and orthogonal regression: Implications for differential validity studies. *Applied Measurement in Education*, 23(3), 215–241. <http://dx.doi.org/10.1080/08957347.2010.485990>
- Kashdan, T. B., Mishra, A., Breen, W. E., & Froh, J. J. (2009). Gender differences in gratitude: Examining appraisals, narratives, the willingness to express emotions, and changes in psychological needs. *Journal of Personality*, 77(3), 691–730. <https://doi.org/10.1111/j.1467-6494.2009.00562.x>
- Kong, F., Ding, K., & Zhao, J. (2015). The relationships among gratitude, self-esteem, social support and life satisfaction among undergraduate students. *Journal of Happiness Studies*, 16, 477–489. <https://doi.org/10.1007/s10902-014-9519-2>
- Kong, F., You, X., & Zhao, J. (2017). Evaluation of the gratitude questionnaire in a Chinese sample of adults: Factorial validity, criterion-related validity, and measurement invariance across sex. *Frontiers in Psychology*, 8, 1498. <https://doi.org/10.3389/fpsyg.2017.01498>

- Langer, Á. I., Ulloa, V. G., Aguilar-Parra, J. M., Araya-Véliz, C., & Brito, G. (2016). Validation of a Spanish translation of the Gratitude Questionnaire (GQ-6) with a Chilean sample of adults and high schoolers. *Health and Quality of Life Outcomes*, 14(1), 1-9. <https://doi.org/10.1186/s12955-016-0450-6>
- Layous K, Sweeny K, Armenta C, Na S, Choi I, Lyubomirsky S (2017) The proximal experience of gratitude. *PLoS ONE* 12(7): e0179123. <https://doi.org/10.1371/journal.pone.0179123>
- Liu, Y., Millsap, R. E., West, S. G., Tein, J. Y., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological Methods*, 22(3), 486-506. <https://doi.org/10.1037/met0000075>
- McCullough, M. E., Emmons, R. A., & Tsang, J.-A. (2002). The grateful disposition: A conceptual and empirical topography. *Journal of Personality and Social Psychology*, 82(1), 112-127. <https://doi.org/10.1037/0022-3514.82.1.112>
- McCullough, M. E., Tsang, J. A., & Emmons, R. A. (2004). Gratitude in intermediate affective terrain: Links of grateful moods to individual differences and daily emotional experience. *Journal of Personality and Social Psychology*, 86(2), 295-309. <https://doi.org/10.1037/0022-3514.86.2.295>
- Meade, A. W., & Bauer, D. J. (2007). Power and precision in confirmatory factor analytic tests of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(4), 611-635. <http://dx.doi.org/10.1080/10705510701575461>
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525-543. <http://dx.doi.org/10.1007/BF02294825>
- Millsap, R. E. (1997). Invariance in measurement and prediction: Their relationship in the single-factor case. *Psychological Methods*, 2(3), 248-260. <https://doi.org/10.1037/1082-989X.2.3.248>
- Oliveira, M. C., Zanon, C., Bandeira, M. C., Heath, P. J., & Giacomoni, H. C. (2022). Evaluating optimism in children and adolescents: Adaptation, factor structure, convergent validity, and invariance of the Brazilian version of the Youth Life Orientation Test (YLOT). *Psychological Assessment*, 34(1), e1-e14. <https://doi.org/10.1037/pas0001090>
- Remor, E. A. (2018). *Gratitude Questionnaire-Six Item Form (GQ-6): Adaptação brasileira e estudo psicométrico preliminar*. [Apresentação de poster]. 48ª Reunião Anual da Sociedade Brasileira de Psicologia, Unisinos, São Leopoldo, RS, Brasil.
- Rey, L., Sánchez-Álvarez, N., & Extremera, N. (2018). Spanish Gratitude Questionnaire: Psychometric properties in adolescents and relationships with negative and positive psychological outcomes. *Personality and Individual Differences*, 135, 173-175. <https://doi.org/10.1016/j.paid.2018.07.005>
- Reynolds, C. R., & Suzuki, L. A. (2013). Bias in psychological assessment: An empirical review and recommendations. In J. R. Graham, J. A. Naglieri, & I. B. Weiner (Eds.), *Handbook of Psychology: Assessment psychology* (pp. 82-113). Wiley.
- Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment*, 29(4), 347-363. <https://doi.org/10.1177/0734282911406661>
- Sun, P., & Kong, F. (2013). Affective mediators of the influence of gratitude on life satisfaction in late adolescence. *Social Indicators Research*, 114, 1361-1369. <https://doi.org/10.1007/s11205-013-0333-8>
- Valdez, J. P. M., Yang, W., & Datu, J. A. D. (2017). Validation of the Gratitude Questionnaire in Filipino secondary school students. *The Spanish Journal of Psychology*, 20. <https://doi.org/10.1017/sjp.2017.51>
- Van de Velde, S., Bracke, P., Levecque, K., & Meuleman, B. (2010). Gender differences in depression in 25 European countries after eliminating measurement bias in the CES-D 8. *Social Science Research*, 39(3), 396-404. <https://doi.org/10.1016/j.ssresearch.2010.01.002>
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4-70. <https://doi.org/10.1177/109442810031002>
- Vandenberg, R. J. (2002). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods*, 5(2), 139-158. <https://doi.org/10.1177/1094428102005002001>
- Warne, R. T., Yoon, M., & Price, C. J. (2014). Exploring the various interpretations of "test bias". *Cultural Diversity & Ethnic Minority Psychology*, 20(4), 570-582. <http://dx.doi.org/10.1037/a0036503>
- Wood, A. M., Froh, J. J., & Geraghty, A. W. (2010). Gratitude and well-being: A review and theoretical integration. *Clinical Psychology Review*, 30(7), 890-905. <https://doi.org/10.1016/j.cpr.2010.03.005>
- Yüksel, A., & Oguz Duran, N. (2012). Turkish adaptation of the gratitude questionnaire. *Eurasian Journal of Educational Research*, 46, 199-215.
- Zanon, C., Brenner, R. E., Baptista, M. N., Vogel, D. L., Rubin, M., Al-Darmaki, F. R., Gonçalves, M., Heath, P. J., Liao, H.-Y., Mackenzie, C. S., Topkaya, N., Wade, N. G., & Zlati, A. (2021). Examining the Dimensionality, Reliability, and Invariance of the Depression, Anxiety, and Stress Scale-21 (DASS-21) Across Eight Countries. *Assessment*, 28(6), 1531-1544. <https://doi.org/10.1177/1073191119887449>

NOTAS

1. Gratidão disposicional refere-se a uma tendência estável de uma pessoa de ser grata em diversas situações e ao longo do tempo, independentemente de circunstâncias específicas. É como um traço de personalidade, no qual a pessoa demonstra uma inclinação natural para reconhecer e valorizar coisas boas na vida, sentindo-se grata com mais frequência. Essa forma de gratidão vai além de momentos pontuais e se manifesta como uma característica constante no modo como a pessoa enxerga o mundo.

Recebido em: 29/11/2022

Primeira decisão editorial em: 18/09/2023

Aceito em: 10/11/2023