

# Versão brasileira da *Self-efficacy for home exercise programs scale* para a população idosa

*Brazilian version of Self-efficacy for home exercise programs scale for the elderly population*

*Versión brasileña de “self-efficacy for home exercise programs scale” para la población de ancianos*

Desiree Rodrigues da Veiga<sup>1</sup> ; Kelsey Jean Picha<sup>11</sup> ; Flávio Rebutini<sup>1</sup> 

<sup>1</sup>Universidade de São Paulo, São Paulo, Brasil; <sup>11</sup>University of Kentucky, Kentucky, Estados Unidos

## RESUMO

**Objetivo:** buscar evidências de validade da versão brasileira *Self-efficacy for home exercise programs scale* para a população idosa. **Método:** foram analisadas as evidências de conteúdo, processo de resposta e estrutura interna. Foram utilizados dois painéis de especialistas (equivalências e conteúdo) por ser um *assembly*. Três formatos da escala foram analisados: original, 5 pontos e 7 pontos. Para o processo de resposta e estrutura interna participaram 174 idosos. A Estrutura interna foi testada pela análise fatorial exploratória. **Resultados:** Sete dos 12 itens passaram por ajuste de redação, e um item foi adicionado por orientação dos especialistas. O processo de resposta indicou que a versão com 5 pontos foi a apontada pelos idosos como a melhor para ser respondida. Na estrutura interna, o modelo unidimensional com a escala de 5 pontos obteve com os melhores indicadores. **Conclusão:** recomenda-se a adoção do instrumento de forma unidimensional e o uso da escala de 5 pontos. **Descritores:** Psicometria; Estudo de Validação; Idoso; Exercício Físico; Autoeficácia.

## ABSTRACT

**Objective:** to seek evidence of validity of the Brazilian version *Self-efficacy for home exercise programs scale* for the elderly population. **Method:** evidence of content, response process and internal structure were analyzed. Two expert panels were used (equivalences and content) as it is an *assembly*. Three scale formats were analyzed: original, 5 points and 7 points. For the response process and internal structure, 174 older people participated. The Internal Structure was tested by exploratory factor analysis. **Results:** Seven of the 12 items underwent wording adjustments, and one item was added following expert guidance. The response process indicated that the version with 5 points was pointed out by the elderly as the best to be answered. In the internal structure, the one-dimensional model with the 5-point scale obtained the best indicators. **Conclusion:** it is recommended to adopt the instrument in a unidimensional way and use the 5-point scale. **Descriptors:** Psychometrics; Validation Study; Aged; Exercise; Self-Efficacy.

## RESUMEN

**Objetivo:** buscar evidencias de validez de la versión brasileña de “*Self-efficacy for home exercise programs scale*” (escala de autoeficacia de programas de ejercicios en el hogar) para la población anciana. **Método:** se analizaron evidencias de contenido, proceso de respuesta y estructura interna. Se utilizaron dos paneles de expertos (equivalencias y contenido) por tratarse de una asamblea. Se analizaron tres formatos de escala: original, 5 puntos y 7 puntos. 174 ancianos participaron en el proceso de respuesta y la estructura interna. Se probó la Estructura Interna mediante análisis factorial exploratorio. **Resultados:** Siete de los 12 ítems se sometieron a ajustes de redacción y se agregó un ítem siguiendo la opinión de expertos. El proceso de respuesta indicó que los ancianos indicaron la versión con 5 puntos como siendo la mejor para ser respondida. En la estructura interna, el modelo unidimensional con escala de 5 puntos obtuvo los mejores indicadores. **Conclusión:** se recomienda adoptar el instrumento de forma unidimensional y utilizar la escala de 5 puntos. **Descriptores:** Psicometría; Estudio de Validación; Anciano; Ejercicio Físico; Autoeficacia.

## INTRODUÇÃO

A qualidade de vida, de maneira geral, pode ser considerada um conceito muito amplo e complexo, que engloba diversas questões como: a saúde física e psicológica, nível de independência, relações sociais, eficácia cognitiva, produtividade, a continuidade de papéis familiares e ocupacionais, além de condições ambientais adequadas<sup>1</sup>. Existe um consenso que prática regular de atividade física é fator determinante para a promoção de saúde, além de auxiliar no tratamento e prevenção de diversas doenças crônicas. A alta incidência e prevalência em diversas populações já é visto como uma questão de saúde pública<sup>2</sup>.

Para a população idosa surgiram iniciativas mundiais para a promoção da saúde, mediante a prática de atividades físicas, que foram desenvolvidas pela Organização Mundial da Saúde (OMS), em 1996, quando a instituição elaborou o protocolo denominado “*The Heiderberg Guidelines for Promotion Physical Among Older Persons*”, que traz evidências científicas dos benefícios da prática de atividade física nos aspectos biológico, psicológico e social, além da importância de políticas públicas no incentivo à atividade física. O documento também fala sobre os fatores motivadores e as barreiras para a atividade física nessa população<sup>3</sup>.

No Brasil, no ano de 1999, a Sociedade Brasileira de Medicina do Esporte, em parceria com a Sociedade Brasileira de Geriatria e Gerontologia, elaborou o documento “Atividade Física e Saúde dos Idosos – posição oficial”, cujo propósito seria ampliar a recomendação das práticas de atividade física pelos profissionais da área da saúde que lidam com indivíduos idosos<sup>4</sup>. Após 20 anos, em 2021, o Ministério da Saúde no ano de lançou o “Guia de Atividade Física para a População Brasileira” onde se pode encontrar recomendações específicas sobre as atividades físicas para os idosos<sup>5</sup>.

Destaca-se que nem sempre a inatividade física em pessoas idosas se deve à falta de prática intencional por parte do indivíduo mas, em muitos casos, podem ser resultado de imposições socioculturais, o que nada tem a ver com uma incapacidade funcional. Culturalmente são atribuídos aos idosos a dependência e a imobilidade fazendo com que eles reduzam significativamente os seus níveis de atividade física.

Com a recente pandemia da COVID-19, que teve início no final de 2019 e chegou a Brasil no início de 2020, tivemos uma mudança na execução dos exercícios físicos em meados de março, as academias foram fechadas, devido ao reconhecimento do estado de calamidade pública. Este movimento forçou muitas pessoas a abandonarem ou realizarem seus exercícios físicos em casa, com a orientação dos profissionais de educação física de forma *online*.

Em um estudo recente de revisão de escopo, artigos que continham recomendações sobre a prática de atividade física e exercício físico para os brasileiros durante a pandemia da COVID-19 foram analisados<sup>6</sup>. Foram encontrados 22 trabalhos onde as recomendações para a prática de atividade e exercício físico contam com a utilização de recursos digitais para sua orientação, e incluem a recomendação para a execução em ambiente domiciliar e, apenas quando possível, em ambiente ao ar livre, respeitando medidas de segurança e o distanciamento social.

A *Self-efficacy for home exercise programs scale* (SEHEPS) é a primeira escala voltada para a avaliação da autoeficácia de um programa de exercícios em casa. A ideia dos autores<sup>7</sup> seria a utilização para a prescrição médica de exercícios físicos com a finalidade de reabilitação. A revisão da escala foi feita por um grupo de especialistas, do qual faziam parte treinadores esportivos, fisioterapeutas e um especialista em autoeficácia a fim de fornecer evidências da validade face e de conteúdo. Estes optaram por modificar a escala de resposta de 11 pontos para 7 pontos (0-6), para reduzir os níveis de discriminação na escala, as âncoras da escala variaram de 0 (não confiante) a 6 (muito confiante). Sendo que os rótulos foram combinados e as âncoras verbais utilizadas apenas em três pontos da escala: o rótulo de “não confiante” foi colocado entre 0 e 1, o rótulo “um pouco confiante” está entre o 2, 3 e 4 e o “muito confiante” entre o 5 e 6. O fato de não termos os rótulos verbais em todos os pontos da escala pode ser um problema quanto a compreensão dos respondentes sobre a resposta que mais se adequa a sua opinião no item<sup>8,9</sup> e podem transmitir informações discrepantes sobre o significado dos intervalos. Inúmeros artigos<sup>10-14</sup>, apesar da pouca atenção recebida sobre esse aspecto, têm apontado problemas nas análises de estrutura interna em decorrência de problemas de redação, formato da escala e de indução de resposta para os participantes gerando viés de resposta.

Como o formato da escala da SEHEPS não apresenta o rótulo em todos os pontos da escala, nós optamos por buscar evidências de validade para a população idosa brasileira com a testagem de duas versões adicionais da escala. Esse processo ganha relevo em razão da resistência em se realizar mudanças substanciais em instrumento, mesmo havendo respaldo técnico para tal. Outro aspecto, se refere à proposta com foco em uma população idosa diferente daquela originária do instrumento e por não termos informações se o formato da escala pode afetar as respostas dos idosos, pois não detectamos estudos que tenham sido desenvolvidos para analisar os efeitos da escala sobre o seu padrão de resposta.

## MÉTODO

Este é um estudo de busca de evidências de validade da versão brasileira do SEHEPS para Idosos e seguiu as recomendações do *Standards*<sup>15,16</sup>, com coletas de dados realizadas entre 2020 e 2022. Três fontes de evidências foram testadas: Conteúdo, a qual inclui a adaptação transcultural, Processo de Resposta e Estrutura Interna, descritas a seguir.

### Evidências de validade de conteúdo/estudo cross-cultural (adaptação transcultural – ATC)

A maior parte dos trabalhos atuais de instrumentos *cross*-culturais seguem os estágios do processo propostos por Beaton *et al.*<sup>17</sup>, envolvendo seis etapas: tradução, síntese da tradução, tradução reversa, comitê de especialistas, pré-teste e submissão a um comitê científico. Porém para as etapas do SEHEPS foram utilizadas as recomendações de Gjersing *et al.*<sup>18</sup>.

O processo de adaptação de instrumentos pode envolver três níveis<sup>19</sup>. A *adoption* (adoção) está pautada na suposição de que quanto menos intervenções no formato original, mais semelhante será a nova versão, modificando o instrumento o mínimo possível. A *adaptation* (adaptação) visa que as similaridades das duas versões só são alcançadas por modificações no instrumento, como mudança de redação do item, inclusão ou retirada de item, e por último. O *assembly* (remontagem), que se preocupa em aperfeiçoar o instrumento para o contexto, tempo e população ao qual

se destinará e não tem a intenção de ser plenamente equivalente ou mesmo imitar o instrumento desenvolvido na cultura original, complementando as questões que são importantes no original com as novas sugestões que podem tornar o instrumento adequado a nova população alvo<sup>20</sup>. Esse último inclui alteração do formato de instruções, formato da escala, acréscimo de itens, entre outros e pode ser, também, denominado como “Severo!” por alguns autores<sup>21,22</sup>.

Ainda é importante destacar que, por vezes, a tradução é usada como sinônimo de adaptação de teste, porém são dois processos diferentes. Traduzir está ligado a transformação do texto de um idioma para outro. Já a adaptação trata buscar comprovar que a adequação não se refere apenas a linguagem, mas também a transformar as características para que se adequem a população alvo de destino<sup>20</sup>.

O instrumento passou pelas etapas de tradução e retrotradução com dois tradutores independentes e um terceiro tradutor para a síntese em cada etapa. Respeitando a fluência entre os idiomas e a natividade. Optou-se pela aplicação de dois painéis de especialistas: o primeiro voltado as equivalências, portanto, na relação entre a versão original e a versão para o Brasil, e o segundo painel teve como objetivo avaliar as modificações realizadas no instrumento e no formato da escala. Os especialistas receberam a versão finalizada no primeiro painel com três formatos de escala a serem avaliadas: original (O) – portanto com os rótulos integrando mais de uma numeração, como apontado anteriormente, pode afetar as análises de estrutura interna, em função de trazer viés de resposta dos participantes, uma versão de 7 pontos com todos os rótulos (P7) e uma versão de 5 pontos com todos os rótulos.

Os especialistas foram selecionados por meio do currículo Lattes e tinham como experiência algum contato a assistência a pessoas idosas, desenvolvimento de ações de promoção da saúde e/ou Educação em Saúde voltada para a orientação de atividade e exercícios físicos, experiência com instrumentos de medida e com especialistas multidisciplinares<sup>23</sup> e que fossem de diversas regiões do país<sup>24,25</sup>. Àqueles que aceitaram o convite receberam após o aceite, um formulário eletrônico desenvolvido por meio do *software QuestionPro*<sup>®</sup> foi remetido por *e-mail*.

Para avaliar as respostas dos especialistas foi utilizado o CVR (*Content Validity Ratio*), desenvolvido por Lawshe<sup>26</sup> em que o escore mínimo da aderência do item à variável latente será baseado nos valores críticos do CVR a partir do número final de especialistas<sup>27</sup>.

### Processo de resposta

Estudos usando essa evidência não são comuns<sup>28</sup>, sendo reportadas evidências de Processo de resposta em apenas 1,8% de 283 testes analisados em estudo prévio<sup>29</sup>. No Brasil, são ainda mais raros.

Em razão dessa raridade dos estudos de processo de resposta, ainda mais com idosos, e da realização do *assembly*, optou-se por realizar o processo de resposta com todos os participantes da etapa de evidências de estrutura interna, portanto, todos os respondentes dos instrumentos, também responderam as questões sobre o processo de resposta.

### Evidências validade da estrutura interna

As evidências de estrutura interna referem-se ao grau em que as relações entre os itens de teste e os componentes de teste estão de acordo com o construto<sup>16</sup>.

O recrutamento dos participantes foi realizado via redes sociais, grupos de *Whatsapp*<sup>®</sup> e em eventos esportivos. Eles receberam um *link* do *QuestionPro*<sup>®</sup>, para que respondessem o instrumento nas suas três versões e avaliassem as três versões como apontado durante a descrição do processo de resposta.

### Análise Fatorial Exploratória (AFE)

A primeira etapa da análise visa a verificação se os dados são passíveis de fatorial, ou seja, a fatorabilidade por meio das Medidas de Adequação da Amostra (*Measure of Sampling Adequacy* - MSA). Para essa etapa, foram avaliados pela esfericidade de Bartlett, determinante da matriz e Kaiser-Meyer-Olkin (KMO) para a amostra total e para os itens, conforme recomendação<sup>30</sup>.

A testagem da dimensionalidade foi realizada com a Análise Paralela por meio da *Optimal implementation of Parallel Analysis* (PA) com *Minimun rank factor analysis* que minimiza a variância comum dos resíduos<sup>31</sup>, a partir da matriz polidórica.

As análises foram realizadas com um *bootstrap* de 5000. O uso da replicação interna é uma orientação contemporânea que tem o intuito de testar a estabilidade e replicação do modelo<sup>32</sup>. Além dessa, foi adotado o índice G-H para avaliar a replicabilidade do modelo.

A extração dos fatores foi feita pela técnica RULS (*Robust Unweighted Least Squares*/mínimos quadrados não ponderados) que reduz os resíduos das matrizes, mais robusta em dados não normais. Ainda, foram adotados como indicadores de avaliação da unidimensionalidade<sup>33</sup>, *Unidimensional Congruence* (UNICO>0,95), *Explained Common Variance* (ECV>0,80) e *Mean of Item Residual Absolute Loadings* (MIREAL<0,30). Caso o instrumento demonstrasse ser

multidimensional, seria utilizada a rotação oblíqua Promin<sup>34</sup> e *Unique Directional Correlation* (Eta) pela Pratt's Measure<sup>35</sup>, como forma complementar para avaliação do modelo.

### Confiabilidade

A confiabilidade foi aferida por três indicadores: alpha de Cronbach<sup>36</sup>, Ômega<sup>37</sup> e *Overall Reliability of fully-Informative prior oblique N-EAP scores* (ORION)<sup>38</sup>.

Para manter ou retirar os itens do modelo, foram avaliadas a magnitude das cargas fatoriais, das comunalidades, a inexistência de *cross-loading*, *heywood cases* e a interpretabilidade dos fatores. Os bancos foram analisados de forma individual para cada uma das versões, sendo denominados da seguinte maneira: versão 1: escala original (O); versão 2: escala de 7 pontos (7P); versão 3: escala de 5 pontos (5P).

Vale ressaltar que este estudo foi realizado atendendo aos princípios éticos, segundo as Resoluções vigentes do Conselho Nacional de Saúde, sendo aprovado pelo Comitê de Ética em Pesquisa da instituição envolvida.

### RESULTADOS

Os resultados são apresentados de acordo com as etapas do procedimento metodológico.

#### Evidências de validade de Conteúdo - tradução e retrotradução

A Figura 1 apresenta o resultado da Tradução 1 e 2 e a síntese realizada pelo terceiro tradutor.

	Original	Tradutor 1	Tradutor 2	Síntese das traduções
	How confident are you that you could perform the prescribed exercises correctly...	Você tem segurança de que poderia praticar os exercícios prescritos corretamente...	O quão confiante você está de que realizará os exercícios prescritos...	Quão confiante você está de que realizará corretamente os exercícios prescritos...
1	...as often as prescribed by your clinician?	... na mesma frequência prescrita por seu clínico?	...na frequência prescrita pelo seu médico?	... com a frequência prescrita pelo seu médico?
2	...when you are bored by the program?	... quando se sente aborrecido(a) pelo programa?	...quando estiver entediado com o programa?	... quando estiver entediado com o programa?
3	...when you feel pain when exercising?	...quando sente dores durante o exercício?	...quando sentir dor durante os exercícios?	... quando sentir dor durante os exercícios?
4	...when you have to exercise alone?	...quando tem de se exercitar sozinho(a)?	...quando tiver de se exercitar sozinho?	... quando tiver que se exercitar sozinho(a)?
5	...when you do not enjoy it?	...quando não tem mais satisfação?	...quando não estiver gostando?	... quando não estiver gostando?
6	...when you are given written exercise instruction?	...quando as instruções são dadas por escrito?	...quando as instruções forem dadas por escrito?	... quando as instruções dos exercícios forem dadas por escrito?
7	...when you are too busy with other activities?	...quando está muito ocupado(a) com outras atividades?	...quando estiver ocupado com outras atividades?	... quando estiver muito ocupado(a) com outras atividades?
8	...when you are given video exercise instruction?	...quando as instruções são dadas por vídeo?	... quando as instruções forem dadas por vídeo?	... quando as instruções dos exercícios forem dadas por vídeo?
9	... when you feel tired?	... quando se sente cansado(a)?	... quando estiver cansado?	... quando se sentir cansado(a)?
10	... when you feel stressed?	...quando se sente estressado(a)?	...quando estiver estressado?	... quando se sentir estressado(a)?
11	... when you feel depressed?	...quando se sente deprimido(a)?	...quando estiver deprimido?	... quando se sentir deprimido(a)?
12	...when you do not have supervision or clinician feedback?	... quando não tem supervisão ou acompanhamento clínico?	...quando não for supervisionado ou não tiver um retorno do médico?	... quando não tiver supervisão ou retorno do médico?

FIGURA 1: Resultados da tradução. São Paulo, SP, Brasil, 2022.

Entre os T1 e o T2 houve algumas discrepâncias em termos do cabeçalho, como: "Você tem segurança de que poderia praticar os exercícios prescrito corretamente..." e "O quão confiante você está de que realizará os exercícios

prescritos...”, que foi sintetizada da seguinte forma “Quão confiante você está de que realizará corretamente os exercícios prescritos...”.

No item 1, o T1 sugeriu a tradução da seguinte forma “... na mesma frequência prescrita por seu clínico?”, e o T2 “...na frequência prescrita pelo seu médico?”, na síntese T3 optou pela opção “...na frequência prescrita pelo seu médico?”, no item 4 também houve discrepâncias entre T1 e T2 “... quando não tem mais satisfação?” e “...quando não estiver gostando?”, na síntese o T3 optou pela mesma tradução do T2. No item 2 as traduções de T1 e T2 ficaram da seguinte maneira: “...quando se sente aborrecido(a) pelo programa?” e “.....quando estiver entediado com o programa?” e na síntese T3 optou pela opção “.....quando estiver entediado com o programa?” Nos demais itens não ocorreram discrepâncias substanciais e não alteram o entendimento ou compreensão sobre o tema.

### Primeiro painel de especialistas de equivalência

Para o primeiro painel de especialistas de (equivalência), enviamos a carta convite e o formulário via *link* do *QuestionPro*® para 50 profissionais das cinco regiões do Brasil por meio da plataforma Lattes. Obtivemos o retorno de 11 formulários com preenchimento completo, sendo que um especialista foi da região Centro-Oeste, dois da região Norte, dois especialistas da região Sul e seis da região sudeste. Não tivemos resposta de especialistas da região nordeste.

As versões foram encaminhadas para análise das equivalências semântica, idiomática, experimental e conceitual. A Tabela 1 apresenta os valores de CVR das equivalências.

**TABELA 1:** Cálculo CVR 1º painel de especialistas de equivalência. São Paulo, SP, Brasil, 2022.

	Itens avaliados	S	I	E	C
Instruções	Por favor, circule seu nível de confiança para concluir os exercícios prescritos para você fazer em casa.	1,00	0,82	0,45	1,00
Rótulos	Rótulos	1,00	1,00	1,00	1,00
Enunciado	Quão confiante você está de que realizará corretamente os exercícios prescritos...	0,82	1,00	0,64	1,00
1	... com a frequência prescrita pelo seu médico?	1,00	0,82	0,82	0,82
2	... quando estiver entediado com o programa?	1,00	1,00	1,00	1,00
3	... quando sentir dor durante os exercícios?	1,00	1,00	1,00	1,00
4	... quando tiver que se exercitar sozinho(a)?	1,00	1,00	1,00	1,00
5	... quando não estiver gostando?	0,64	1,00	1,00	1,00
6	... quando as instruções dos exercícios forem dadas por escrito?	1,00	1,00	1,00	1,00
7	... quando estiver muito ocupado(a) com outras atividades?	0,82	1,00	1,00	1,00
8	... quando as instruções dos exercícios forem dadas por vídeo?	1,00	1,00	1,00	0,82
9	... quando se sentir cansado(a)?	0,82	1,00	1,00	1,00
10	... quando se sentir estressado(a)?	0,82	1,00	1,00	1,00
11	... quando se sentir deprimido(a)?	0,82	1,00	1,00	1,00
12	... quando não tiver supervisão ou retorno do médico?	1,00	1,00	1,00	1,00

**Legenda:** CVR – Content Validity Ratio; S - Semântica; I – Idiomática; E – Experimental; C – Conceitual; CVR – Content Validity Ratio.

Todos os itens para as quatro equivalências tiveram valores acima de 0.59 (valor crítico para 11 especialistas).

Foram sugeridas algumas alterações. O termo "*completing*" seria mais bem compreendido em português por "realizar". Trocou-se "prescritos por "orientados". Para o item 1 da escala houve a indagação por um dos especialistas sobre quem deve prescrever exercícios físicos: "Exercícios físicos não são prescritos por educadores físicos?". No item 5, 7, 9, 10 e 11 foi sugerido que se inserisse o pronome "você", assim, como exemplo para o item 5 passasse de "Quando não estiver gostando dos exercícios" para "Quando você não estiver gostando?". Para o item 8 foi sugerido "Quando você estiver recebendo instruções dos exercícios por vídeo".

Para as equivalências, ainda foram adotadas questões complementares (Tabela 2).

**TABELA 2:** Avaliação das equivalências de maneira geral. São Paulo, SP, Brasil, 2022.

Perguntas	N	S	ND	NR
1 Os itens têm o mesmo significado ou significado muito semelhante nos dois idiomas?	0	11	0	0
2 Existem pistas gramaticais que podem tornar os itens mais fáceis ou mais difíceis na versão do idioma de destino?	6	3	1	1
3 Há alguma estrutura gramatical na versão dos itens no idioma de origem que não tem paralelos no idioma de destino?	9	1	0	1
4 Há alguma palavra nos itens que, ao ser traduzida, deixa de ter um significado para ter mais de um significado comum?	9	1	1	0
5 Existem diferenças entre as versões dos itens no idioma de destino e de origem em relação ao uso de metáforas, expressões idiomáticas ou coloquialismos?	11	0	0	0
6 O formato do item e a tarefa exigidos ao respondente serão igualmente familiares nas duas versões linguísticas?	2	9	0	0
7 Quando a passagem é traduzida do idioma de origem para o idioma de destino, as palavras e frases da versão traduzida transmitem conteúdo e ideias semelhantes à versão de origem?	0	11	0	0
8 A passagem inclui conteúdo ou requer habilidades que podem não ser familiares a alguns respondentes em qualquer um dos dois idiomas ou grupos culturais?	8	3	0	0
9 Existem diferenças culturais que afetariam a probabilidade de uma resposta ser escolhida quando o item é apresentado na versão do idioma de origem ou de destino?	8	2	1	0
10 Existem diferenças culturais que afetariam a probabilidade de uma resposta ser escolhida quando o item é apresentado na versão do idioma de origem ou de destino?	11	0	0	0
11 O conceito ou construção do item tem aproximadamente a mesma familiaridade e significado nas versões do idioma de origem e de destino?	0	11	0	0

**Legenda:** N – não; S – Sim; ND – não sei dizer; NR – não relevante

Os resultados, de forma geral, indicaram equivalência adequada entre o original e a versão brasileira. Mesmo assim, alguns apontamentos foram realizados: nas instruções, foi apontado que a expressão "nível de confiança" não é um termo comumente utilizado para a população idosa e poderia gerar dúvidas.

Sobre o item 1, um especialista destacou a redação "*as often as prescribed by your clinician?*" que pode ter sentido de "frequência" ou pode representar uma "submissão a uma diretriz médica", sem necessariamente ter relação com a "frequência" daquele evento. Sugerindo trocar "frequência" por "sempre". Exemplo: "Sempre que prescrito pelo seu...". Outro especialista sugeriu rever a redação sobre a prescrição de exercícios físicos por médicos, pois o exercício físico pode ser feito por outros tipos de profissionais, predominantemente por profissionais de educação física. Sobre a sentença "frequência prescrita" sugeriu colocar "frequência orientada" por ser uma expressão mais utilizada quando se fala em exercícios físicos.

### Segundo painel de especialistas de conteúdo

Para o segundo painel de especialistas, enviamos a carta convite e o formulário via *link* do *QuestionPro*® para 114 profissionais das cinco regiões do Brasil, por meio da plataforma Lattes e de forma direta no *e-mail* do profissional. Obtivemos o retorno de 87 profissionais, sendo que 12 desses responderam ao formulário de maneira completa. Os especialistas estavam distribuídos nas seguintes regiões: um do sul, três do nordeste (Bahia, Ceará e Rio Grande do Norte), uma do norte (Amazonas), sete do sudeste (Minas Gerais, Rio de Janeiro e cinco do estado de São Paulo). Não tivemos nenhum especialista da região centro-oeste nesta etapa.

Os especialistas na primeira parte do formulário avaliaram cada um dos itens com relação a: pertinência do item para a avaliação da autoeficácia; clareza na redação do item e se o item está redigido corretamente. Na tabela 3 podemos verificar os valores do CVR dos itens.

**TABELA 3:** Cálculo CVR para pertinência, clareza e redação do item. São Paulo, SP, Brasil, 2022.

		Pertinência	Clareza	Redação
1	... com a frequência prescrita pelo seu médico?	0,67	0,67	0,17
2	... quando estiver entediado com o programa?	0,67	0,33	0,50
3	... quando sentir dor durante os exercícios?	1,00	1,00	0,83
4	... quando tiver que se exercitar sozinho(a)?	1,00	1,00	1,00
5	... quando não estiver gostando?	1,00	1,00	0,83
6	... quando as instruções dos exercícios forem dadas por escrito?	0,83	1,00	0,83
7	... quando estiver muito ocupado(a) com outras atividades?	1,00	1,00	1,00
8	... quando as instruções dos exercícios forem dadas por vídeo?	0,83	1,00	0,67
9	... quando se sentir cansado(a)?	1,00	1,00	0,83
10	... quando se sentir estressado(a)?	1,00	1,00	1,00
11	... quando se sentir deprimido(a)?	1,00	0,83	0,83
12	... quando não tiver supervisão ou retorno do médico?	1,00	1,00	0,67

**Legenda:** CVR – Content Validity Ratio.

O item 1 para redação (0,17) e o item 2 para clareza (0,33) e redação (0,50) ficaram abaixo do valor crítico (0,56).

Com relação a pertinência do item foram feitas sugestões pelos especialistas: “Semelhança entre os termos “entediados”, “cansado” e a expressão “quando não estiver gostando”. Outro termo apontado foi “programa” para designar as atividades físicas orientadas, uma sugestão seria utilizar a expressão “exercícios prescritos” ao invés de “programa”. Foi indicado pelos especialistas a substituição do termo “médico”, pois este exclui outros profissionais da saúde, que no Brasil tem autorização legal para a prescrição. Foi sugerido rever os itens de conotação negativa, pois estes ficam conflitantes com o enunciado do instrumento “quão confiante...”. E rever também o item que trata da prática de atividades físicas quando estiver deprimido, pois a depressão independe da autoeficácia, é algo que foge do controle e gerenciamento do indivíduo, assim como a dor.

Na avaliação da clareza dos itens no enunciado foi solicitado rever a expressão “Quão confiante”, tivemos novamente um questionamento sobre os termos “programas”, já que este não é comum no contexto brasileiro da atividade física. Assim como a expressão: “Com a frequência prescrita pelo seu médico”, pois dependendo do grau de instrução do idoso, é possível que a palavra “frequência” não seja compreendida de maneira adequada, foi sugerida como opção “o número de vezes que o seu médico recomendou”.

Outra sugestão apontada seria substituir a palavra “prescrita” pela palavra “receitada”, pois a essa palavra pode ser uma expressão mais usual, facilitando a compreensão os indivíduos avaliados. No item 2 foi sugerido completar a redação, especificando que tipo de programa, assim como no item 5 completar, gostando de quê? Nos itens com palavras negativas também foram solicitadas alterações como “... quando não estiver gostando?”, substituiria por: “...quando estiver insatisfeito?” E para o item 12 foi sugerido a divisão do item, criar um item para supervisão do profissional de saúde e um para retorno ao profissional de saúde

Houve sugestões para alterações, na palavra prescrição, pois esta é repetida na mesma sentença, pode ser que pacientes com menor nível instrucional não entendam o significado, sendo assim o especialista sugeriu a revisão deste termo. Assim como sugere-se substituir médico por algum profissional. No item 5 completar a sentença, pois não fica claro sobre o que está gostando pode ser da atividade, do dia, do corpo, instruções ou até mesmo do professor. Nos itens 3, 7 e 8 foi sugerido evitar a repetição da palavra exercício por “atividade”, por exemplo. Para melhorar a redação do item 8 foi sugerido alterar de “...quando as instruções dos exercícios forem dadas por vídeo?” por “oferecidas por vídeo”. No item 9 “Quando se sentir cansado (a)?”, deixar claro se o cansaço é dado pelo estado geral do idoso, do dia a dia ou se o cansaço dado pelo próprio exercício. Recomendam-se também revisões nos itens 2 e 11, sugere-se substituir os termos “entediado” e “deprimido”, por termos mais acessíveis a população de destino.

Os resultados do CVR para as questões complementares de ordem geral do instrumento são apresentados na Tabela 4.

**TABELA 4:** Avaliação do instrumento de maneira geral. São Paulo, SP, Brasil, 2022.

Perguntas	Não	Sim	CVR
1 O instrumento mede a variável latente Autoeficácia?	0	12	1,00
2 O instrumento mede apenas o que se propôs?	0	12	1,00
3 A linguagem adotada é adequada à população alvo no caso a população idosa brasileira?	5	7	0,17
4 O número de itens é suficiente para a medida?	1	11	0,83
5 A escala é coerente com a proposta?	1	11	0,83
6 O dimensionamento da escala é coerente com a variável latente?	2	10	0,67

**Legenda:** CVR – Content Validity Ratio

Apontam, novamente, que o critério que avaliava a adequação da linguagem para a população alvo resultou em um CVR (0,17) abaixo do valor crítico. Obviamente, resultado das diversas sugestões para ajuste de redação. As sugestões dos especialistas foram atendidas para solucionar o item 3 da Tabela 4, quanto foi pedido que avaliassem a linguagem adotada.

Em função da testagem de três possíveis configurações da escala, um conjunto de questões foram submetida aos especialistas (Tabela 5).

**TABELA 5:** Considerações dos especialistas com relação as opções de escala. São Paulo, SP, Brasil, 2022.

Perguntas	Original	7 pontos	5 pontos
1 Rótulos descrevem melhor como posso me sentir	0	3	9
2 Melhor visualização dos rótulos	0	3	9
3 Melhor compreensão sobre as opções que os rótulos nos trazem	0	2	10
4 Disposição dos rótulos e números torna melhor a visualização das respostas	0	2	10
5 Considerando a estrutura da escala de uma maneira geral depois de ter respondido as três opções, qual você acredita ser a mais adequada para as respostas aos itens pensando na população alvo os idosos?	0	2	10

Para as cinco questões formuladas ao painel, a escala com formato de 5 pontos teve a melhor avaliação. No polo oposto, a escala original não recebeu indicação para nenhuma das perguntas feitas.

#### Alterações realizadas no instrumento

Em razão das extensivas sugestões dos especialistas, o instrumento passou por diversos ajustes que podem ser visualizados na Figura 2: síntese das traduções, alterações sugeridas e redação final.

	SINTESE DAS TRADUÇÕES	ALTERAÇÕES SUGERIDAS	REDAÇÃO FINAL
	Quão confiante você está de que realizará corretamente os exercícios prescritos...	Qual o seu nível confiança para realizar corretamente os exercícios indicados...	Qual o seu nível confiança para realizar corretamente os exercícios indicados...
1	... com a frequência prescrita pelo seu médico?	... na quantidade que foi informada pelo profissional?	... na quantidade que foi informada pelo profissional?
2	... quando estiver entediado com o programa?	... quando estiver desmotivado com exercícios?	... quando me sinto motivado com os exercícios?
3	... quando sentir dor durante os exercícios?	... quando sentir dor durante os exercícios?	... quando sinto dor durante os exercícios?
4	... quando tiver que se exercitar sozinho(a)?	... quando tiver que realizar os exercícios sozinho(a)?	... quando tenho que realizar os exercícios sozinho(a)?
5	... quando não estiver gostando?	... quando não estiver gostando?	... quando estou gostando dos exercícios?
6	... quando as instruções dos exercícios forem dadas por escrito?	... quando as instruções dos exercícios forem passadas por escrito?	... quando as instruções são passadas por escrito?
7	... quando estiver muito ocupado(a) com outras atividades?	... quando estiver muito ocupado(a) com outras atividades?	... quando estou muito ocupado(a) com outras atividades?
8	... quando as instruções dos exercícios forem dadas por vídeo?	... quando as instruções dos exercícios forem passadas por vídeo?	... quando as instruções são passadas por vídeo?
9	... quando se sentir cansado(a)?	... quando se sentir cansado(a) por qualquer motivo?	... quando me sinto cansado(a) por qualquer motivo?
10	... quando se sentir estressado(a)?	... quando se sentir estressado(a) por qualquer motivo?	... quando me sinto estressado(a) por qualquer motivo?
11	... quando se sentir deprimido(a)?	... quando se sentir deprimido(a) por qualquer motivo?	... quando me sinto deprimido(a) por qualquer motivo?
12	... quando não tiver supervisão ou retorno do médico?	... quando não tiver supervisão ou retorno do profissional?	... quando tenho supervisão?
13			... quando tenho retorno do profissional?

**FIGURA 2:** Alterações realizadas na escala. São Paulo, SP, Brasil, 2022.

Foi sugerido, pelos especialistas, que os itens 2 e 5 fossem agrupados em um único, mas após a revisão optou-se por mantê-los separados. Sendo a escala no sentido ascendente, modificou-se o item 2 que é um item reverso para “quando me sinto motivado” e o item 5 para “quanto gosto do exercício”. Revisou-se a concordância entre a âncora e o complemento (itens) para manter a coerência da redação. Os itens 2, 3, 4, 7, 9, 10 e 11 foram alterados. Nos itens 6 e 8 os itens a redação foi modificada para o tempo presente. O item 12 foi dividido em dois itens, devido a ter a preposição “ou”, o que indica dois fatores sendo avaliados pelo item<sup>39,40</sup>, passando assim o instrumento a ter 13 itens. As sentenças dos itens 5, 12 e 13 foram passadas para o positivo.

## Evidências de Processo de Resposta e de Estrutura Interna

### Caracterização da amostra

A coleta foi realizada de 31 de agosto a 30 de novembro de 2022. No total, o questionário teve 2.211 visualizações, sendo que 412 pessoas iniciaram as respostas, com 240 desistências após o início das respostas e 174 completaram o questionário, com uma taxa de conclusão de 41,75%. O tempo médio para o preenchimento foi de 11 minutos, incluindo o questionário sociodemográfico e as três versões do instrumento.

Participaram do estudo idosos entre 60 e 95 anos, sendo a média de 65 anos (desvio padrão=6,72), sendo 68,53% do sexo feminino. A predominância da raça e/ou cor foi a branca (72,56%), casado (41,57%), católicos (51,81%), nível superior (35%), com uma renda de mais de 5 salários-mínimos (36,13%) e 53,89% a origem da renda vem de aposentadoria. Com relação a origem dos participantes, tivemos representantes das cinco regiões do país, sendo a sua maioria (85,71%) do Sudeste, sendo 73,55% residentes no estado de São Paulo; 3,23% eram da região Norte, 3,24% do Nordeste, 2,59% do Centro-Oeste e 5,17% do Sul.

Quando questionados sobre as doenças pré-existentes, a hipertensão foi a recorrente declarada por 20% dos participantes, seguida por ansiedade com 9,72% e hipotireoidismo com 9,03%. A maioria da amostra declarou tomar alguma medicação (75,30%) e, quando questionados sobre o tipo de acesso a tratamentos de saúde, 27,44% utilizam convênio individual e 23,78% o Sistema Único de Saúde (SUS). Sobre ingestão de bebidas alcoólicas, 60,48% declaram não fazer uso e 97,62% reportam não fumar.

Tivemos 88,10% dos participantes realizando atividades físicas e 11,09%, apesar de já terem praticado, não estão realizando no momento. A finalidade da prática da maioria dos participantes foi a saúde (38,42%), seguida pela estética e indicação médica (10,35%), sendo de intensidade moderada para 66,87% deles. O local apontado para a prática foi em casa para 24,64%; os tipos de atividade mais frequentes foram a caminhada (34,59%), a ginástica (18,90%), musculação (12,79%) e outros tipos (13,66%), sendo que 42,57% dos participantes as realizam sozinhos, 35,64% em grupo e 21,78% acompanhados de um profissional. Realizam na sua maioria no período matutino (61,20%), vespertino (27,32%) e noturno (11,48%), sendo que 45,58% dos participantes declararam que recebem a orientação para a realização das práticas por um professor.

### Processo de resposta

Após os participantes da amostra responderem às três versões do instrumento, sendo a versão 1: escala original (O); versão 2: escala de 7 pontos (7P); versão 3: escala de 5 pontos (5P), eles foram convidados a responder alguns questionamentos sobre essas 3 versões. Optamos em realizar essa etapa, junto com a etapa da coleta da estrutura interna, devido às dificuldades impostas pela pandemia em relação ao contato com os participantes e pelo desenho de pesquisa permitir a realização das etapas concomitantemente.

Além disso, as questões do processo de resposta diziam respeito à comparação entre a facilidade de uso das três versões. A tabela 6 mostra os resultados dessa etapa.

**TABELA 6:** Processo de resposta avaliação geral do instrumento. São Paulo, SP, Brasil, 2022.

	O		P7		P5	
	n	%	n	%	n	%
Mais facilidade para responder	50	32,47%	32	20,78%	<b>72</b>	<b>46,75%</b>
Mais fácil a compreensão	45	29,22%	31	20,13%	<b>78</b>	<b>50,65%</b>
Mais rápido de responder	35	22,44%	31	19,87%	<b>90</b>	<b>57,69%</b>
Mais segurança nas suas respostas	44	28,21%	27	17,31%	<b>85</b>	<b>54,49%</b>

**Legenda:** O – Original; P7 – 7 pontos; P5 – 5 pontos.

Para as quatro questões aplicadas a versão com 5 pontos obtive os melhores resultados. Interessante apontar que a versão de 7 pontos teve índices piores do que a versão original.

## Evidências de validade da estrutura interna

### Fatorabilidade

Os índices globais de adequação da amostra para a versão “O” foram: KMO=0,75, Esfericidade de Bartlett=1094,0 (df=78;  $p<0.001$ ) e o determinante da matriz=0,000337. Para a “P7” obteve-se KMO=0,75, Esfericidade de Bartlett=1183,9 (df=78;  $p<0.00001$ ) e o determinante da matriz=0,0000158. Para a “P5”, KMO=0,79, Esfericidade de Bartlett=1213,5 (df=78;  $p<0,00001$ ) e o determinante da matriz=0,0000159. Na versão “O”, o resultado pode indicar um problema de redundância de item ( $p>0,000001$ )<sup>41</sup>.

#### i. Dimensionalidade

Apesar da proposta original de modelo unidimensional, a análise paralela com os 13 itens indicou a possibilidade de um modelo de duas dimensões pela média dos autovalores e de uma dimensão pelo percentil 95. Em função disso, optamos por fazer análises *post hoc* da dimensionalidade<sup>39,42</sup>.

Assim, foram aplicadas mais duas técnicas de dimensionalidade. A *Bayesian Information Criterion* (BIC)<sup>43</sup> e a HULL<sup>44</sup>. Novamente tivemos as duas opções, enquanto a BIC indicou um modelo de duas dimensões a HULL indicou um modelo unidimensional para as três versões da escala.

Para os dados de *closeness of dimensionality*, há reforço de um modelo multidimensional para os três bancos. Sendo na versão “O”: UNICO=0,92; ECV=0,74 e MIREAL = 0,35, que podem indicar a existência de multidimensionalidade. Na versão “P7”: UNICO=0,88; ECV=0,68 e MIREAL=0,48. A versão 3 “P5”: UNICO=0,89; ECV=0,71 e MIREAL=0,42. Como as técnicas divergem, possibilitam mais de uma configuração nas dimensões, optamos por explorar a solução com uma e duas dimensões.

#### ii. Análise da versão unidimensional

Os resultados obtidos com a análise são apresentados na Tabela 7.

**TABELA 7:** Cargas fatoriais e  $h^2$  modelo unidimensional. São Paulo, SP, Brasil, 2022.

Item		O		P7		P5	
		$\lambda$	$h^2$	$\lambda$	$h^2$	$\lambda$	$h^2$
1	na quantidade que foi informada pelo profissional?	<b>0,65</b>	0,43	<b>0,68</b>	0,47	<b>0,67</b>	0,45
2	quando me sinto motivado com os exercícios?	<b>0,63</b>	0,40	<b>0,69</b>	0,48	<b>0,72</b>	0,52
3	quando sinto dor durante os exercícios?	0,42	0,18	<b>0,66</b>	0,44	<b>0,57</b>	0,33
4	quando tenho que realizar os exercícios sozinho(a)?	<b>0,67</b>	0,45	<b>0,81</b>	0,67	<b>0,81</b>	0,66
5	quando gosto dos exercícios?	<b>0,74</b>	0,55	<b>0,68</b>	0,46	<b>0,62</b>	0,39
6	quando as instruções são passadas por escrito?	<b>0,61</b>	0,37	<b>0,75</b>	0,57	<b>0,80</b>	0,64
7	quando estou muito ocupado(a) com outras atividades?	<b>0,58</b>	0,33	<b>0,55</b>	0,31	<b>0,56</b>	0,31
8	quando as instruções são passadas por vídeo?	<b>0,63</b>	0,39	<b>0,65</b>	0,42	<b>0,67</b>	0,46
9	quando me sinto cansado(a) por qualquer motivo?	<b>0,70</b>	0,49	<b>0,64</b>	0,41	<b>0,71</b>	0,50
10	quando me sinto estressado(a) por qualquer motivo?	<b>0,66</b>	0,43	<b>0,58</b>	0,34	<b>0,68</b>	0,47
11	quando me sinto deprimido(a) por qualquer motivo?	<b>0,58</b>	0,33	<b>0,63</b>	0,40	<b>0,59</b>	0,35
12	quando tenho supervisão?	<b>0,56</b>	0,31	<b>0,62</b>	0,39	<b>0,63</b>	0,39
13	quando tenho retorno do profissional?	<b>0,57</b>	0,32	<b>0,52</b>	0,27	<b>0,44</b>	0,19

**Legenda:** O – Original; P7 – 7 pontos; P5 – 5 pontos;  $\lambda$  – cargas fatoriais;  $h^2$  – comunalidades.

Os dados iniciais de “O” apresentaram cargas fatoriais variando de 0,43 e 0,74, sendo a recomendação de pelo menos 0,30, para uma amostra de mais de 300 indivíduos. Devido à nossa amostra ter sido menor do que 300 participantes, o valor de corte das cargas fatoriais ( $\lambda$ ) subiu para 0,45, conforme recomendação da literatura<sup>45</sup>. As comunalidades ( $h^2$ ) variaram entre 0,18 e 0,55. O modelo tem variância explicada de 47,20%.

Para a versão “P7”, as cargas fatoriais variaram de 0,52 a 0,82, comunalidade de 0,27 a 0,67. O modelo tem variância explicada de 49,14%.

Para a versão “P5”, as cargas fatoriais variaram de 0,56 a 0,82, comunalidade de 0,20 a 0,66. O modelo tem variância explicada de 54,02%.

O “P5” é o único com variância explicada acima de 50% e tem cargas fatoriais e comunalidades melhores do que os modelos “O” e “P7”.

Os índices de confiabilidade variaram entre os bancos de análise de 0,91 a 0,89 para o Cronbach e para o Ômega nas três versões. Os índices de qualidade da solução fatorial também mostraram patamares adequados. Os índices entre os três formatos, em sua maioria, são similares.

Contudo, a variância explicada para o modelo “P5” é determinante, porque indica que ele mensura melhor a variável latente, como demonstrado na Tabela 8.

**TABELA 8:** Síntese do modelo unidimensional para as versões. São Paulo, SP, Brasil, 2022.

Síntese	Índice	Técnica	O	P7	P5
<b>Exploratória</b>	Adequação da matriz de correlação	Determinante da matriz	0,0003371	0,0000015	0,0000159
		Bartlett (df = 78)	1094,0	1183,9	1213,5
		KMO (Kaiser-Meyer-Olkin)	0,75	0,75	0,79
	Variância explicada (AP)		47,20%	49,14%	54,22%
	Polychoric Correlation ( $r_p =$ )		0,031 - 0,820	0,012 - 0,899	0,035 - 0,838
<b>Confiabilidade</b>	Cronbach's Alpha		0,89	0,91	0,90
	McDonald's Ômega		0,89	0,91	0,90
<b>Replicabilidade</b>	Índice G-H (latente/Observado)		0,89/0,87	0,91/0,94	0,92/0,89
<b>Avaliação Unidimensional</b>	Unidimensional Congruence (UNICO)		0,92	0,88	0,89
	Explained Common Variance (ECV)		0,74	0,68	0,71
	Mean of item residual absolute loading (MIREAL)		0,35	0,45	0,42
<b>Qualidade e Eficácia</b>	Factor Determinacy Index (FDI) <sup>a</sup>		0,947	0,987	0,959
	Sensitivity Ratio (SR) <sup>a</sup>		2,956	3,314	3,393
	Expected percentage of true differences (EPTD) <sup>a</sup>		92,40%	93,30%	93,50%

**Legenda:** O – Original; P7 – 7 pontos; P5 - 5 pontos.

Soma-se que os três modelos indicaram boa replicabilidade (índice G-H) tanto na latente quanto na observada. Indicando que os resultados obtidos são consistentes e tendem a serem encontrados para outras sub-populações.

### iii. Análise da versão bidimensional

A Tabela 9 apresenta os valores primários do modelo bidimensional.

**TABELA 9:** Cargas fatoriais, comunalidades e ETA modelo bidimensional. São Paulo, SP, Brasil, 2022.

	Original				7P				5P						
	$\lambda$		$h^2$	Eta		$\Lambda$		$h^2$	Eta		$\lambda$		$h^2$	Eta	
	1D	2D	1D	2D	1D	2D	1D	2D	1D	2D	1D	2D	1D	2D	
Item 1	<b>0,86</b>	-0,09	0,67	<b>0,82</b>	0,07	<b>0,84</b>	-0,08	0,66	<b>0,81</b>	0,00	<b>0,87</b>	-0,14	0,67	0,82	0,00
Item 2	<b>1,03</b>	-0,25	0,85	<b>0,92</b>	0,00	<b>0,94</b>	-0,17	0,79	<b>0,89</b>	0,00	<b>0,81</b>	-0,02	0,64	<b>0,80</b>	0,00
Item 3	0,38	0,10	0,20	<b>0,40</b>	0,20	<b>0,14</b>	<b>0,76</b>	0,68	0,24	<b>0,79</b>	<b>0,00</b>	<b>0,75</b>	0,57	<b>0,03</b>	0,75
Item 4	<b>0,79</b>	-0,01	0,61	<b>0,76</b>	0,18	<b>0,65</b>	0,30	0,65	<b>0,70</b>	0,40	<b>0,62</b>	0,30	0,65	0,69	<b>0,42</b>
Item 5	<b>0,71</b>	0,15	0,63	<b>0,72</b>	0,34	<b>0,74</b>	0,02	0,55	<b>0,74</b>	0,08	<b>0,72</b>	0,04	0,49	<b>0,70</b>	0,00
Item 6	0,39	0,30	0,37	0,45	0,41	<b>0,80</b>	0,05	0,67	<b>0,81</b>	0,13	<b>0,81</b>	0,08	0,72	<b>0,83</b>	0,19
Item 7	0,15	<b>0,51</b>	0,36	0,25	<b>0,55</b>	0,00	<b>0,80</b>	0,64	0,00	<b>0,80</b>	-0,01	<b>0,75</b>	0,56	<b>0,00</b>	0,75
Item 8	0,25	<b>0,46</b>	0,40	0,33	<b>0,53</b>	<b>0,73</b>	-0,01	0,53	<b>0,73</b>	0,00	<b>0,62</b>	0,14	0,48	0,65	<b>0,24</b>
Item 9	0,21	<b>0,58</b>	0,52	0,33	<b>0,64</b>	-0,04	<b>0,89</b>	0,81	0,12	<b>0,89</b>	0,02	<b>0,93</b>	0,88	0,10	<b>0,90</b>
Item 10	0,07	<b>0,69</b>	0,53	0,19	<b>0,70</b>	-0,09	<b>0,98</b>	0,90	0,00	<b>0,95</b>	0,08	<b>0,80</b>	0,71	0,19	<b>0,82</b>
Item 11	-0,13	<b>0,82</b>	0,57	0,00	<b>0,76</b>	-0,07	<b>0,83</b>	0,73	0,15	<b>0,84</b>	-0,09	<b>0,90</b>	0,74	0,00	<b>0,86</b>
Item 12	-0,11	<b>0,77</b>	0,52	0,00	<b>0,72</b>	<b>0,79</b>	-0,11	0,57	<b>0,76</b>	0,00	<b>0,87</b>	-0,19	0,65	<b>0,80</b>	0,00
Item 13	-0,12	<b>0,79</b>	0,54	0,00	<b>0,74</b>	<b>0,58</b>	0,00	0,34	<b>0,58</b>	0,00	<b>0,66</b>	-0,20	0,36	0,60	0,00

**Legenda:** O – Original; P7 – 7 pontos; P5 - 5 pontos.

Os dados iniciais da versão (O), apresentaram cargas fatoriais variando de 0,46 e 1,02 (*Heywood case*) as comunalidades variaram entre 0,19 e 0,85 e Eta de 0,53 a 0,92. O item 2 apresentou uma violação da carga fatorial com um *Heywood Case*, isto é, a carga fatorial violou os limites teóricos das cargas fatoriais de -1 a 1. O modelo tem variância explicada de 50,67% (Tabela 9).

No modelo bidimensional com a escala original, os itens de 1 a 6 se alinharam em um domínio e os de 7 a 13 em outro, sendo que os itens 3 e 6 apresentaram um valor abaixo do desejável em ambas as dimensões. Esse dado é reforçado pelos valores do Eta (*Pratt's Measure*) que auxilia na confirmação dos valores das cargas fatoriais.

Para a versão "P7", as cargas fatoriais variaram de 0,58 a 0,98, comunalidade de 0,33 a 0,90 e Eta de 0,58 a 0,89. O modelo tem variância explicada de 51,88%. Esse modelo também apresentou duas dimensões sendo uma composta pelos itens de 1, 2, 4, 5, 6, 8 e 13, e a outra pelos itens 3, 7, 9, 10 e 11.

No último banco, no modelo "P5" as cargas fatoriais variaram de 0,61 a 0,92, comunalidade de 0,33 a 0,90 e Eta de 0,49 a 0,90. O modelo tem variância explicada de 53,03%. Esse modelo também apresentou duas dimensões sendo uma composta 1, 2, 4, 5, 6, 8, 12 e 13 e a outra 3, 7, 9, 10 e 11. Os modelos "P5" e "P7" apresentaram a mesma configuração de itens nos domínios com índices adequados, o que não ocorreu com a versão (O).

Em todas as três versões analisadas, o conteúdo dos domínios não permite uma interpretação. Portanto, apesar do alinhamento quantitativo não há o alinhamento qualitativo do conteúdo. Sendo esses dois elementos indissociáveis na análise da estrutura interna. Consequentemente, pela inviabilidade dos dois elementos, os índices de confiabilidade, replicabilidade, qualidade e eficácia dos escore perdem o poder interpretativo.

Desta forma, a solução unidimensional encontra respaldo quantitativo e qualitativo. O modelo "P5" tem índices melhores do que o "O" e "P7", sendo o único a ter variância explicada acima de 50%.

Os índices de confiabilidade variaram entre os bancos de análise de 0,89 a 0,91 para o alpha de Cronbach e para o Ômega nas três versões. A sínteses dos modelos podem ser vistas na Tabela 10.

**TABELA 10:** Cargas fatoriais e  $h^2$  modelo bidimensional. São Paulo, SP, Brasil, 2022.

Síntese	Índice	Técnica	O	P7	P5
Exploratória	Adequação da matriz de correlação	Determinante da matriz	0,0003371	0,0000015	0,0000159
		Bartlett (df = 78)	1094,0	1183,9	1213,5
		KMO (Kaiser-Meyer-Olkin)	0,75	0,75	80
		Variância explicada (AP)	50,67%	51,88%	53,03%
	Polychoric Correlation ( $r_p$ ) =		0,031 - 0,820	0,012 - 0,899	0,035 - 0,838
Confiabilidade		Cronbach's Alpha	0,89	0,91	0,91
		McDonald's Omega	0,89	0,91	0,91
		ORION <sup>a</sup>	0,928; 0,880	0,933; 0,954	0,926; 0,939
Avaliação Unidimensional		Unidimensional Congruence (UNICO)	0,92	0,878	0,887
		Explained Common Variance (ECV)	0,74	0,679	0,708
		Mean of item residual absolute loading (MIREAL)	0,35	0,447	0,417
Qualidade e Eficácia		Factor Determinacy Index (FDI) <sup>a</sup>	0,963; 0,938	0,966; 0,977	0,962; 0,969
		Sensitivity Ratio (SR) <sup>a</sup>	3,580; 2,708	3,742; 4,558	3,525; 3,940
		Expected percentage of true differences (EPTD) <sup>a</sup>	93,9%; 91,65	94,2%; 95,6%	93,8%; 94,6%

**Legenda:** O – Original; P7 – 7 pontos; P5 - 5 pontos; a – Valores por domínio.

Desta forma, o conjunto de técnicas e índices empregados apontam para um conjunto de evidências de validade de estrutura interna adequados, consistente, robustas e interpretáveis quantitativamente e qualitativamente, com recomendação para o instrumento com uma escala numérica e rotulada de 5 pontos.

## DISCUSSÃO

Este estudo teve como seu principal objetivo a adaptação e a busca de evidências de validade da versão brasileira do SEHEPS voltada para a população idosa, sendo a versão original voltada para adultos e com intuito de avaliar as práticas de atividades físicas voltadas para a reabilitação, as adaptações não só culturais, mas também contextuais foram necessárias.

Em algumas escalas, realizar uma adaptação ou uma adequação é quase como desenvolver um instrumento novamente, pois é muito mais difícil ou quase impossível de se obter a equivalência desejada, independentemente de quais sejam as áreas de adequação: linguística; cultural; psicométrica, no caso do nosso estudo, optar pelo “*assembly*” foi a melhor opção para fornecer uma medida viável, útil e válida do construto para a cultura e o idioma de destino<sup>46</sup>.

Após a avaliação dos especialistas tivemos sugestões de alteração para melhorar a equivalência em sete dos 12 itens do instrumento, reforçando assim as diferenças existentes entre as culturas e da população alvo do instrumento, no que tange à prática desenvolvida e aos termos empregados, ademais da linguagem.

Todos os especialistas concordaram que os itens têm significados similares nos dois idiomas, mas também todos concordaram que existiam diferenças entre as versões dos itens no idioma de destino e de origem em relação ao uso de metáforas, expressões idiomáticas ou coloquialismos, portanto os ajustes necessários foram realizados respeitando a equivalência do seu significado.

Apesar da adequação dos itens, não foi necessária a exclusão de nenhum. Porém, para uma melhor compreensão, foi necessário incluir um item, respeitando a necessidade de inclusão dos aspectos presentes na cultura-alvo e carentes no instrumento a ser adaptado, sendo assim o item 12 foi dividido em dois itens, deixando assim o instrumento com 13 itens. Soma-se a realização de dois painéis, o primeiro voltado às equivalências e o segundo efetivamente ao conteúdo. Ambos os painéis contaram com uma representação ampla de regiões do Brasil, no intuito de garantir uma compreensão do conteúdo nos mais diversos contextos<sup>23-25</sup>.

Há orientações<sup>39,40</sup> para o desenvolvimento e formato das escalas de resposta, principalmente, a necessidade de que todos os pontos da escala sejam rotulados e numerados, o que não acontecia na escala original.

A proposta e testagem de três configurações da escala nos permite um olhar não só sobre os aspectos do ajustamento estatístico, mas também compreender como os participantes respondem e geram a associação entre os itens do instrumento.

Outro aspecto para a sustentação da realização de um *assembly* reside no fato de que a população alvo era completamente diferente da população do instrumento original. Visto que o intuito do instrumento, base para o nosso estudo, era avaliar as atividades e exercícios físicos realizados de forma remota em população idosa, algo relativamente recente na nossa cultura, que se popularizou e se tornou viável com o início da pandemia e perdura até os dias atuais. Não encontramos na literatura, até o momento, instrumento que avalie o comportamento dos praticantes neste novo contexto.

Diante desse primeiro elemento da fase de conteúdo deve-se ter claro que a aplicação de dois painéis centrados em elementos de avaliação distintos contribuiu de forma extensiva para a melhoria do instrumento. Inúmeras contribuições surgiram do segundo painel que não foram apontados no processo de equivalência. A despeito da possibilidade de evocar painéis de especialistas em diversos momentos da busca de evidências<sup>15,16</sup>, eles raramente são adotados. Muitas das alterações propostas pelo segundo painel não foram sequer aludidas pelo painel de equivalências. Inclusive porque diversas perguntas são singulares aos objetivos de cada etapa, o que pode ser entendido também como uma contribuição para estudos futuros da nossa pesquisa. Esses apontamentos nos levam a refletir sobre a extensão, qualidade e precisão das adequações realizadas nos estudos que utilizam apenas o painel de equivalências como fonte para avaliação do conteúdo do instrumento.

Cabe ainda ressaltar que não é, usualmente, papel do painel de equivalências apontar alterações mais profundas no instrumento, novamente, porque não é seu objetivo primário e porque os protocolos não induzem ou recomendam aos pesquisadores que sejam direcionadas perguntas a pontos especificamente técnicos de redação, formato, orientação, integração e interpretação, além do sentido da equivalência. Tanto que os dois protocolos de adaptação citados nesse estudo<sup>17,18</sup>, sequer tratam, por exemplo, das questões técnicas de redação e formato de escala, apesar do olhar as essas questões serem apontadas em diversos livros<sup>20,39</sup>. Tal fato nos leva a pontar que o painel de equivalências, na prática, não seja suficiente para um processo de adaptação adequado.

A realização do processo de resposta amplo, com todos os participantes participando conjuntamente enquanto preenchiam as três versões, permitiu elencar elementos e evidências de que o modelo com 5 pontos da escala era superior aos outros dois modelos tanto do ponto de vista quantitativo (estrutura interna) quanto qualitativo (conteúdo e processo de resposta).

Devemos lembrar que o processo de resposta, apesar de estar estabelecido há décadas no *Standard*<sup>15,16</sup>, anteriormente integrado na fase de validade de conteúdo, é pouco utilizado e muito ainda há de ser desenvolvido.

Superadas as questões relativas aos significados, é possível verificar nos resultados da análise fatorial que eles diferem entre os três formatos da escala, apontando para um modelo de escala de 5 pontos com todos os rótulos numerados com todos os índices adequados e consistentes.

O fato central reside no fato da extensiva testagem necessária para se atingir esse ponto. Não se afasta o fato de que os dados oferecidos pelo modelo unidimensional, em uma primeira visão, indicarem a possibilidade da utilização dos três formatos de escala - não permitindo ignorar o problema original, já tratado no formato da escala. Por si, o problema da escala impede tratar os dados de estrutura interna do modelo original sem desconfiança.

Um aspecto que talvez passe despercebido para um leitor que não tenha familiaridade com estudo de evidências, é a existência de testagem de dimensionalidade *post hoc*, o que inclui técnicas adicionais a partir dos resultados encontrados e a integração de diversos indicadores de modelo.

Diante da possibilidade de mais de uma configuração dimensional dos instrumentos, fez-se necessária a extensão da testagem, bem como, a avaliação se os resultados da matriz fatorial seriam interpretáveis.

Apesar dos elementos encontrados da análise reforçarem e direcionarem a melhor solução para um instrumento unidimensional, é preciso ter em vista que existe a possibilidade de que estudos futuros possam ampliar o instrumento e aperfeiçoá-lo no intuito de encontrarmos um modelo que possa acomodar de forma acurada e precisa duas dimensões, que preliminarmente são apenas indícios.

## CONCLUSÃO

Diante da extensiva testagem e da comparação entre os três modelos, a recomendação é de uso da versão de 5 pontos da escala do SEHEPS para a população idosa que apresentou níveis adequados, satisfatórios, coerente e robustos de evidências de validade.

Destaca-se, até onde foi possível investigar, que este é o primeiro instrumento no Brasil que durante seu processo de busca de evidências de validade *cross-cultural* a ter passado por um processo de *assembly*.

Como apontado anteriormente, a realização de dois painéis de especialistas independentes para avaliar aspectos distintos do conteúdo e a adoção das evidências de processo de resposta com a amostra integral de respondentes e a aplicação de uma extensiva testagem na estrutura interna, assim como a realização de técnicas de *post hoc* para a dimensionalidade, podem contribuir para trabalhos de outros pesquisadores, principalmente, pelo fato que essa extensividade de procedimento e técnicas não serem usualmente utilizados.

## REFERÊNCIAS

1. Veras RP, Oliveira M. Aging in Brazil: the building of a healthcare model. *Ciênc saúde colet*. 2018 [cited 2020 Oct 13]; 23:1929-36. DOI: <https://doi.org/10.1590/1413-81232018236.04722018>.
2. Bottcher LB. Atividade física como ação para promoção da saúde. *Rev. G. & S*. 2019 [cited 2020 Oct 25]; 14:98-111. DOI: <https://doi.org/10.26512/gsv0i0.23324>.
3. World Health Organization. The Heidelberg guidelines for promoting physical activity among older persons. WHO, Copenhagen, [cited 2020 Oct 13]; 1-14. Available from: <https://apps.who.int/iris/handle/10665/108545>.
4. Nóbrega AC, Freitas EV, Oliveira MA, Leitão MB, Lazzoli JK, Nahas RM, et al. Posicionamento oficial da Sociedade Brasileira de Medicina do Esporte e da Sociedade Brasileira de Geriatria e Gerontologia: atividade física e saúde no idoso. *Rev Bras Med Esporte*. 1999 [cited 2020 Oct 13]; 5:207-11. DOI: <https://doi.org/10.1590/S1517-86921999000600002>.
5. Ministério da Saúde. Guia de Atividade Física para a População Brasileira [Internet]. Brasília: 2021. Available from: <https://aps.saude.gov.br/biblioteca/visualizar/MjA1MA==>.
6. Mattos SM, Pereira DS, Moreira TM, Cestari VR, Gonzalez RH. Recomendações de atividade física e exercício físico durante a pandemia Covid-19: revisão de escopo sobre publicações no Brasil. *Rev Bras Ativ Fís Saúde*. 2020 [cited 2021 Sep 25]; 25:e0176. DOI: <https://doi.org/10.12820/rbafs.25e0176>.
7. Picha KJ, Lester M, Heebner NR, Abt JP, Usher EL, Capilouto G, et al. The self-efficacy for home exercise programs scale: development and psychometric properties. *J Orthop Sports Phys Ther*. 2019 [cited 2020 Oct 20]; 49(9):647-55. DOI: <https://www.jospt.org/doi/10.2519/jospt.2019.8779>.
8. Menold N. Rating-scale labeling in online surveys: An experimental comparison of verbal and numeric rating scales with respect to measurement quality and respondents' cognitive processes. *Sociol Methods Res*. 2020 [cited 2021 Apr 19]; 49(1):9-107. DOI: <https://doi.org/10.1177/0049124117729694>.
9. Menold N, Kemper CJ. The impact of frequency rating scale formats on the measurement of latent variables in web surveys-an experimental investigation using a measure of affectivity as an example. *Psihologija*. 2015 [cited 2021 Mar 10]; 48(4):431-49. DOI: <https://doi.org/10.2298/PSI1504431M>.
10. Robie C, Meade AW, Risavy SD, Rasheed S. Effects of response option order on Likert-type psychometric properties and reactions. *Educ Psychol Meas*. 2022 [cited 2022 Dec 15]; 82(6):1107-29. DOI: <https://doi.org/10.1177/00131644211069406>.
11. Zhang X, Noor R, Savalei V. Examining the effect of reverse worded items on the factor structure of the need for cognition scale. *PLoS one*. 2016 [cited 2021 Oct 10]; 11(6):e0157795. DOI: <https://doi.org/10.1371/journal.pone.0157795>.
12. Sonderen EV, Sanderman R, Coyne JC. Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PLoS one*. 2013 [cited 2021 Oct 15]; 8(7):e68967. DOI: <https://doi.org/10.1371/journal.pone.0068967>.

13. Wetzel E, Böhnke JR, Rose N. A simulation study on methods of correcting for the effects of extreme response style. *Educ Psychol Meas.* 2016 [cited 2021 Oct 20]; 76(2):304-24. DOI: <https://doi.org/10.1177/0013164415591848>.
14. Weijters B, Millet K, Cabooter E. Extremity in horizontal and vertical Likert scale format responses. Some evidence on how visual distance between response categories influences extreme responding. *Int J Res Mark.* 2021 [cited 2021 May 20]; 38(1):85-103. DOI: <https://doi.org/10.1016/j.ijresmar.2020.04.002>.
15. Association AER, Association AP, Association AE, Education NCoMi: Standards for Educational and Psychological Testing American Educational Research Association; 1999.
16. Association AER, Association AP, Association AE, Education NCoMi: Standards for Educational and Psychological Testing American Educational Research Association; 2014.
17. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine.* 2000 [cited 2021 May 10]; 25(24):3186-91. Available from: [https://journals.lww.com/spinejournal/Citation/2000/12150/Guidelines\\_for\\_the\\_Process\\_of\\_Cross\\_Cultural.14.aspx](https://journals.lww.com/spinejournal/Citation/2000/12150/Guidelines_for_the_Process_of_Cross_Cultural.14.aspx).
18. Gjersing L, Capplehorn JR, Clausen T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC Med Res Methodol.* 2010 [cited 2021 Apr 19]; 10:13. DOI: <https://doi.org/10.1186/1471-2288-10-13>.
19. Van de Vijver FJ, Leung K. *Methods and data analysis for cross-cultural research.* Cambridge University Press; 2021.
20. Iliescu D. *Adapting tests in linguistic and cultural situations.* Cambridge University Press; 2017.
21. Stewart AL, Thrasher AD, Goldberg J, Shea JA. A framework for understanding modifications to measures for diverse populations. *J Aging Health.* 2012 [cited 2021 Jul 15]; 24(6):992-1017. DOI: <https://doi.org/10.1177/0898264312440321>.
22. Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value Health.* 2009 [cited 2021 Jun 12]; 12(4):419-29. DOI: <https://doi.org/10.1111/j.1524-4733.2008.00470.x>.
23. Epstein J, Osborne RH, Elsworth GR, Beaton DE, Guillemin F. Cross-cultural adaptation of the Health Education Impact Questionnaire: experimental study showed expert committee, not back-translation, added value. *J Clin Epidemiol.* 2015 [cited 2021 Jun 15]; 68(4):360-9. DOI: <https://doi.org/10.1016/j.jclinepi.2013.07.013>.
24. Costa Palacio D, Rebutini F, Oliveira DB, Neto JP, Barbieri W, Sanchez TP, et al. Dental vulnerability scale in primary health care: evidence of content and structure internal validity. *BMC Oral Health.* 2021 [cited 2022 Sep 18]; 21:421. DOI: <https://doi.org/10.1186/s12903-021-01742-6>.
25. Nobile GG, Barrera SD, Rebutini F. Avaliação da alfabetização: elaboração e validação de conteúdo do IBALEC. *Rev. Psicopedagogia.* 2021 [cited 2023 Jan 11]; 38(117):333-45. DOI: <http://dx.doi.org/10.51207/2179-4057.20210028>.
26. Lawshe CH. A quantitative approach to content validity. *Pers Psychol.* 1975 [cited 2021 May 22]; 28(4):563-75. DOI: <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>.
27. Wilson FR, Pan W, Schumsky DA. Recalculation of the critical values for Lawshe's content validity ratio. *Meas Eval Couns Dev.* 2012 [cited 2021 Apr 08]; 45(3):197-210. DOI: <https://doi.org/10.1177/0748175612440286>.
28. Anastasi A, Urbina S. *Testagem psicológica.* Artmed; 2000.
29. Cizek GJ, Rosenberg SL, Koons HH. Sources of validity evidence for educational and psychological tests. *Educ Psychol Meas.* 2008 [cited 2021 Aug 18]; 68(3):397-412. DOI: <https://doi.org/10.1177/0013164407310130>.
30. Lorenzo-Seva U, Ferrando PJ. MSA: the forgotten index for identifying inappropriate items before computing exploratory item factor analysis. *Methodology.* 2021 [cited 2022 Dec 11]; 17(4):296-306. DOI: <https://doi.org/10.5964/meth.7185>.
31. Timmerman ME, Lorenzo-Seva U. Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychol Methods.* 2011 [cited 2022 Dec 11]; 16(2):209. DOI: <https://doi.org/10.1037/a0023353>.
32. Goretzko D, Bühner M. Robustness of factor solutions in exploratory factor analysis. *Behaviormetrika.* 2022 [cited 2022 Dec 11]; 49(1):131-48. DOI: <https://doi.org/10.1007/s41237-021-00152-w>.
33. Ferrando PJ, Lorenzo-Seva U. Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educ Psychol Meas.* 2018 [cited 2022 Dec 11]; 78(5):762-80. DOI: <https://doi.org/10.1177/0013164417719308>.
34. Lorenzo-Seva U. Promin: A method for oblique factor rotation. *Multivariate Behav Res.* 1999 [cited 2022 Dec 11]; 34(3):347-65. DOI: [https://doi.org/10.1207/S15327906MBR3403\\_3](https://doi.org/10.1207/S15327906MBR3403_3).
35. Wu AD, Zumbo BD, Marshall SK. A method to aid in the interpretation of EFA results: An application of Pratt's measures. *Int J Behav Dev.* 2014 [cited 2022 Dec 11]; 38(1):98-110. DOI: <https://doi.org/10.1177/0165025413506143>.
36. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika.* 1951 [cited 2022 Dec 11]; 16(3):297-334. DOI: <https://doi.org/10.1007/BF02310555>.
37. McDonald RP. *Test theory: A unified treatment.* Psychology Press; 2013.
38. Ferrando PJ, Lorenzo-Seva U. A note on improving EAP trait estimation in oblique factor-analytic and item response theory models. *Psicológica.* 2016 [cited 2022 Dec 11]; 37(2):235-47. Available from: <https://psycnet.apa.org/record/2016-34732-007>.
39. Bandalos DL. *Measurement theory and applications for the social sciences.* Guilford Publications; 2018.
40. Johnson RL, Morgan GB. *Survey scales: A guide to development, analysis, and reporting.* Guilford Publications; 2016.
41. Field A. *Discovering statistics using IBM SPSS statistics 5th ed.* 2018.
42. Auerswald M, Moshagen M. How to determine the number of factors to retain in exploratory factor analysis: a comparison of extraction methods under realistic conditions. *Psychol Methods.* 2019 [cited 2023 Jan 17]; 24(4):468. DOI: <https://doi.org/10.1037/met0000200>.

43. Schwarz G. Estimating the dimension of a model. *The annals of statistics*. 1978 [cited 2023 Feb 11]; 1:461-4. Available from: <https://www.jstor.org/stable/2958889>.
44. Lorenzo-Seva U, Timmerman ME, Kiers HA. The Hull method for selecting the number of common factors. *Multivariate Behav Res*. 2011 [cited 2023 Feb 11]; 46(2):340-64. DOI: <https://doi.org/10.1080/00273171.2011.564527>.
45. Hair J, Black W, Anderson R, Babin B. *Multivariate data analysis* (8<sup>th</sup> ed.). Cengage Learning EMEA. 2018.
46. He J, van de Vijver F. Bias and equivalence in cross-cultural research. *J Psychol*. 2012 [cited 2023 Feb 22]; 2(2):2307-19. DOI: <https://doi.org/10.1108/09504120910935093>.

**Contribuições dos autores:**

Concepção, D.R.V., K.J.P. e F.R.; metodologia, D.R.V. e F.R.; software, D.R.V. e F.R.; validação, D.R.V. e F.R.; investigação, D.R.V.; análise Formal, D.R.V. e F.R.; obtenção de recursos, D.R.V. e F.R.; curadoria de Dados, D.R.V. e F.R.; redação - preparação do manuscrito, D.R.V. e F.R.; redação – revisão e edição, D.R.V., K.J.P. e F.R.; visualização, D.R.V. e F.R.; supervisão, F.R.; administração do Projeto, D.R.V. e F.R. Todos os autores realizaram a leitura e concordaram com a versão publicada do manuscrito.