

## Propriedades psicométricas do instrumento adaptado *European Health Literacy Survey Questionnaire short-short form*

Fábio Luiz Mialhe<sup>1</sup>

 <https://orcid.org/0000-0001-6465-0959>

Katarinne Lima Moraes<sup>2</sup>

 <https://orcid.org/0000-0001-6169-0461>

Fernanda Maria Rovai Bado<sup>1,3</sup>

 <https://orcid.org/0000-0002-7974-5456>

Virginia Visconde Brasil<sup>4</sup>

 <https://orcid.org/0000-0002-0279-9878>

Helena Alves de Carvalho Sampaio<sup>5</sup>

 <https://orcid.org/0000-0001-5353-8259>

Flávio Rebutini<sup>6</sup>

 <https://orcid.org/0000-0002-3746-3266>

<sup>1</sup> Universidade Estadual de Campinas, Faculdade de Odontologia de Piracicaba, Piracicaba, SP, Brasil.

<sup>2</sup> Universidade Federal de Jataí, Curso de Enfermagem, Jataí, GO, Brasil.

<sup>3</sup> Secretaria Municipal de Saúde, Piracicaba, SP, Brasil.

<sup>4</sup> Universidade Federal de Goiás, Faculdade de Enfermagem, Goiânia, GO, Brasil.

<sup>5</sup> Universidade Estadual do Ceará, Benfica, Fortaleza, CE, Brasil.

<sup>6</sup> Universidade de São Paulo, Escola de Artes, Ciências e Humanidades, São Paulo, SP, Brasil.

**Objetivo:** investigar as propriedades psicométricas da versão adaptada ao português do Brasil do questionário de letramento em saúde *European Health Literacy Survey Questionnaire short-short form* (HLS-EU-Q6) em adultos brasileiros. **Método:** o instrumento foi traduzido e pré-testado em uma amostra de 50 indivíduos. Em seguida, ele foi aplicado em uma amostra com 783 indivíduos adultos. Os dados passaram por um processo adequado de testagem das propriedades, com a combinação de técnicas da Análise Fatorial Exploratória, Análise Fatorial Confirmatória e Teoria de Resposta ao Item. Para a avaliação da confiabilidade, foram utilizados os indicadores de alfa de Cronbach e o Ômega de McDonald. Foi aplicada a validação cruzada com análise integral dos dados. **Resultados:** a maioria dos participantes era do sexo feminino (68,1%), com idade média de 38,6 (dp=14,5) anos e 33,5% estudaram até o ensino fundamental. Os resultados indicaram um modelo unidimensional com variância explicada de 71,23%, níveis de cargas fatoriais, comunalidades e discriminação do item adequados, bem como estabilidade e a replicabilidade do instrumento para outras populações. **Conclusão:** a versão brasileira do HLS-EU-Q6 indicou que o instrumento está adequado para aplicação de forma indiscriminada na população a que se destina para aferir os níveis de letramento em saúde.

**Descritores:** Letramento em Saúde; Promoção da Saúde; Estudo de Validação; Inquéritos e Questionários; Sistema Único de Saúde; Adulto.

### Cómo citar este artículo

Mialhe FL, Moraes KL, Bado FMR, Brasil VV, Sampaio HAC, Rebutini F. Psychometric properties of the adapted instrument *European Health Literacy Survey Questionnaire short-short form*. Rev. Latino-Am. Enfermagem. 2021;29:e3436. [Access   ]; Available in: . DOI: <http://dx.doi.org/10.1590/1518-8345.4362.3436>.

## Introdução

O Letramento em Saúde (LS) é um constructo relacionado ao uso das múltiplas formas de informação em saúde nos mais variados contextos<sup>(1)</sup>. Apesar de haver várias definições, as quais abrangem as características pessoais, os recursos sociais e o papel dos serviços de saúde nesse processo<sup>(1-2)</sup>, para a Organização Mundial da Saúde (OMS), o LS diz respeito aos conhecimentos, motivações e competências das pessoas para acessar, compreender, julgar e aplicar as informações sobre saúde, de maneira a tomar decisões que as auxiliem a navegar pelos sistemas de saúde, bem como promover, prevenir e cuidar de sua saúde<sup>(1)</sup>.

Estudos indicam associações entre indivíduos com baixos níveis de LS e menor participação em atividades promotoras da saúde e relacionadas à prevenção de doenças, escolhas em saúde menos assertivas, pior autocontrole de doenças crônicas, maior frequência de hospitalizações e de casos de morbidade e mortalidade, com consequente aumento dos custos para os sistemas de saúde<sup>(1,3)</sup>. À vista disso, o LS é considerado pela OMS como um importante determinante social da saúde, influenciado por características socioeconômicas, culturais e do funcionamento dos sistemas de saúde<sup>(1)</sup>.

Diversos instrumentos já foram desenvolvidos para mensurar este constructo em indivíduos e populações<sup>(4-5)</sup>, entretanto, a maioria avalia apenas as características funcionais do LS, ou seja, as habilidades pessoais para ler e entender informações escritas e orais relacionadas à saúde<sup>(1,6)</sup>. A fim de superar tal lacuna, um consórcio europeu de instituições de pesquisa elaborou um modelo multidimensional e integrativo do LS e desenvolveu um instrumento para sua aferição constituído de 47 itens, denominado de HLS-EU-Q47<sup>(1,7)</sup>. O questionário HLS-EU-Q47 avalia as habilidades individuais em compreender, avaliar e aplicar informações relacionadas à saúde e foi elaborado a partir de um modelo conceitual de LS que integra três domínios: cuidados de saúde (16 questões), promoção da saúde (16 questões) e prevenção das doenças (15 questões). Suas opções de respostas estão dispostas segundo uma escala Likert de quatro pontos que varia entre 1 para muito difícil e 4 para muito fácil<sup>(7-9)</sup>. Por despender cerca de 10 minutos para o seu preenchimento, versões mais curtas foram desenvolvidas, ou seja, o HLS-EU-Q16 (*short form*) e o HLS-EU-Q6 (*short short form*), os quais apresentam 16 e seis questões, respectivamente<sup>(9)</sup>. Entretanto, até o momento, poucos estudos utilizaram o HLS-EU-Q6<sup>(7,9-12)</sup> e/ou avaliaram suas propriedades psicométricas<sup>(13-14)</sup>, evidenciando-se a importância de testá-lo de forma mais robusta em outras populações.

Apesar de a maioria das pesquisas sobre o LS estar concentrada no continente europeu, América do Norte e Austrália<sup>(1-2)</sup>, tem-se observado, na última década, uma expansão dos estudos em outras partes do mundo, como no Brasil<sup>(15-16)</sup>, inclusive com a criação da Rede Brasileira de Letramento em Saúde (REBRALS). Tendo em conta o baixo nível de escolaridade e as dificuldades de compreensão das recomendações profissionais pela população brasileira<sup>(15-16)</sup>, é importante que existam instrumentos simples e curtos para mensurar o constructo do LS nesse contexto, a fim de torná-lo aplicável na prática dos serviços.

No nosso país, o termo *literacy* tem sido traduzido como alfabetização, literacia e letramento<sup>(17)</sup>. Entretanto<sup>(17)</sup>, apesar de ambos serem processos inseparáveis, a alfabetização deve ser entendida como o "processo de aquisição e apropriação do sistema da escrita, alfabético e ortográfico" enquanto o letramento como "o desenvolvimento de habilidades de uso da leitura e da escrita nas práticas sociais que envolvem a língua escrita, e de atitudes positivas em relação a essas práticas".

Apesar desse contexto profícuo de pesquisas, nota-se ainda que, até o momento, foram validados poucos instrumentos para aferir o LS na população brasileira que agregam aspectos mais amplos do constructo para além de mensurar seus aspectos funcionais, e que sejam de fácil e rápida aplicação<sup>(18)</sup>. Assim, o objetivo do presente estudo foi analisar as evidências das propriedades psicométricas do instrumento HLS-EU-Q6, validado para o português do Brasil.

## Método

O projeto de pesquisa foi submetido e aprovado pelo Comitê de Ética em Pesquisa (CAAE: 58131216.5.0000.5418). Inicialmente, foi solicitada à Prof<sup>a</sup> Kristine Sørensen, autora responsável pelo instrumento, a autorização para sua tradução ao português do Brasil.

O instrumento foi traduzido e adaptado segundo as recomendações da literatura<sup>(19-20)</sup>. Para tal, a versão original do questionário HLS-EU-Q47 foi traduzida do inglês para o português do Brasil por dois professores de inglês e um pesquisador da área da saúde com conhecimento da língua inglesa.

A versão de consenso foi então retraduzida para o inglês (*back-translation*) por dois tradutores nativos de língua inglesa que não participaram da primeira etapa da tradução. Um comitê de especialistas, composto por seis *experts* na área da saúde com experiência no campo do Letramento em Saúde e com alto nível de proficiência em língua inglesa, bem como um professor de português e

outro de linguística, avaliou todo o processo de tradução e retrotradução e propôs uma versão final do instrumento, sendo que a versão em inglês foi submetida à avaliação e aprovação da autora responsável. A versão do HLS-EU-Q47 foi então aplicada em 50 indivíduos adultos, usuários dos serviços de saúde dos municípios de Piracicaba/SP; São Paulo/SP, Aparecida de Goiânia/GO e Fortaleza/CE, selecionados aleatoriamente. Não se verificou a necessidade de mudanças no instrumento após esta fase de pré-teste.

A avaliação das propriedades psicométricas do HLS-EU-Q6 foi realizada por meio de um estudo transversal com uma amostra de 783 indivíduos adultos. Dentre eles, 320 residiam em uma área adscrita a uma Unidade Básica de Saúde (UBS) situada no município de São Paulo/SP e 293 residiam em áreas próximas a três Unidades de Saúde da Família (USF) localizadas no município de Piracicaba/SP. As residências foram sorteadas de forma aleatória. De posse do cadastro dos usuários, e com a ajuda de um programa de celular, os indivíduos foram convidados a participar da pesquisa e entrevistados em suas residências. Ademais, participaram do estudo 50 indivíduos acompanhantes de uma instituição de referência em tratamento do câncer em Fortaleza/CE, que aguardavam na sala de espera e foram aleatoriamente convidados para participar da pesquisa. Por último, 120 usuários de uma USF localizada em Aparecida de Goiânia/GO que aguardavam por atendimento na sala de espera e aceitaram participar da pesquisa foram aleatoriamente convidados a participar.

Para o cálculo do tamanho da amostra, considerou-se uma proporção de, pelo menos, 15 adultos para cada pergunta do questionário, superior a recomendação geral encontrada na literatura de 10:1, a qual possibilita análises mais precisas<sup>(21)</sup>.

O questionário HLS-EU-Q6 é denominado de *short-short form* e é composto por seis questões do HLS-EU-Q47<sup>(7,9)</sup> [Em uma escala que vai de "muito fácil" a "muito difícil", com que facilidade você consegue: 1. avaliar quando você precisa de uma segunda opinião de outro médico?; 2. usar as informações que seu médico passa a você para tomar decisões sobre a sua doença?; 3. encontrar informações sobre como lidar com problemas de saúde mental, como o estresse ou depressão?; 4. avaliar se as informações sobre os riscos à saúde disponíveis nos meios de comunicação são confiáveis? (por ex. TV, internet ou outros meios de comunicação); 5. encontrar informações sobre as atividades que são boas para o seu bem-estar mental? (por ex. meditação, exercício, caminhada, pilates etc.); 6. entender as informações disponíveis nos meios de comunicação sobre como ficar mais saudável? (por ex. internet, jornais, revistas)]. As questões 1 e 2 estão relacionadas à avaliação e aplicação

de informações relevantes para a saúde no campo do cuidado à saúde enquanto as questões 3 e 4 tratam sobre encontrar/acessar e avaliar as informações no campo da prevenção das doenças. Por último, as questões 5 e 6 investigam a capacidade do indivíduo para encontrar/acessar e entender as informações relevantes à saúde no campo da promoção da saúde<sup>(7,9)</sup>.

O escore final individual é uma média calculada por meio da soma das respostas das seis questões dividida pelo número de itens respondidos. O escore é calculado desde que pelo menos cinco das seis questões sejam respondidas de forma diferente de 1, e varia entre 1 e 4, sendo que valores maiores indicam melhores níveis de LS. De acordo com os autores do instrumento, o valor final dos escores classifica os indivíduos segundo três níveis de LS: inadequado ( $\leq 2$ ); problemático ( $> 2$  e  $\leq 3$ ); e suficiente ( $> 3$ )<sup>(7,9-10)</sup>.

Para a análise estatística, os dados passaram por um processo extenso e robusto de testagem das propriedades, com a combinação de técnicas da Análise Fatorial Exploratória (AFE), Análise Fatorial Confirmatória (AFC) e Teoria de Resposta ao Item (TRI), visando a busca de fortes evidências de validação na etapa de construto e sua estabilidade para outras subamostras. A AFE exige o cumprimento de diversas etapas, tais como: técnicas de inspeção dos dados, do método de análise do fator, da técnica de retenção e de rotação e os índices de qualidade das cargas fatoriais<sup>(22)</sup>.

A testagem da dimensionalidade foi realizada pela Análise Paralela Robusta por meio da *Optimal implementation of Parallel Analysis* com a *Minimum rank factor analysis* que minimiza a variância comum dos resíduos<sup>(23)</sup>. A robustez do teste foi determinada a partir da associação de um *bootstrap* com uma extrapolação amostral para 5000. A estimação da matriz policórica foi realizada por meio do *Bayes Modal Estimation*<sup>(24)</sup>.

A dimensionalidade, na análise fatorial exploratória (modelo irrestrito), foi testada pela Análise Paralela, que tem sido considerada uma das técnicas mais efetivas e precisas para testagem do número de fatores/dimensionalidade<sup>(25-27)</sup>. A extração dos fatores foi feita pela técnica RULS (*Robust Unweighted Least Squares*), que reduz os resíduos das matrizes<sup>(27)</sup>.

Como análise complementar para a testagem do número de fatores, foram aplicadas as técnicas de unidimensionalidade/multidimensionalidade<sup>(28)</sup> UNICO (*Unidimensional Congruence*  $> 0,95$ ), ECV (*Explained Common Variance*  $> 0,80$  - QUINN, 2014) e MIREAL (*Mean of Item Residual Absolute Loadings*  $< 0,30$ ). Estas técnicas foram aplicadas ao instrumento e aos itens. No caso dos itens, foram usadas para garantir e avaliar se o item aderiria de forma unidimensional ou

multidimensional, isto é, se existia a possibilidade de o item carregar de forma significativa em mais de uma dimensão. A variância explicada do instrumento deve ficar por volta de 60% e as cargas fatoriais iniciais de 0,30<sup>(22)</sup>. Além disso, buscam-se comunalidades médias com valores entre 0,40 e 0,60<sup>(29)</sup>. A manutenção ou retirada de um item do modelo dependerá da magnitude das comunalidades, das cargas fatoriais, do tamanho da amostra e do grau que o item consegue mensurar o fator e da inexistência de *cross-loading*.

Para a confirmação do ajustamento das cargas fatoriais foi adotada a técnica de *Normal-Ogive Graded Response Model*<sup>(30)</sup> para estrutura politômica, por meio da Teoria de Resposta ao Item. Adotou-se o índice de discriminação do item ( $a$ ), que mede a força de associação entre o item e a variável latente e tem interpretação similar às cargas fatoriais da análise fatorial exploratória<sup>(31)</sup> para complementá-la. Adotou-se a recomendação de Baker<sup>(32)</sup> de que " $a$ " <0,65 é considerada baixo poder de discriminação; entre 0,65 e 1,34 discriminação moderada, entre 1,35 e 1,69 alta discriminação e acima de 1,70 discriminação muito alta.

Para os índices de ajustamento na AFC consideraram-se as cargas fatoriais maiores do que 0,50 e os seguintes índices mínimos para a adequação, considerando o número de participantes e variáveis: NNFI (*Non-Normed Fit Index* >0,95); CFI (*Comparative Fit Index* >0,95); GFI (*Goodness Fit Index* >0,95); AGFI (*Adjusted Goodness Fit Index* >0,95); RMSEA (*Root Mean Square Error of Approximation* <0,08) e RMSR (*Root Mean Square of Residuals* <0,08)<sup>(22)</sup>.

A confiabilidade do instrumento foi avaliada a partir de dois indicadores: Alfa<sup>(33)</sup> e Ômega<sup>(34)</sup>. A adoção de dois indicadores buscou aumentar a confiabilidade da interpretação, pois têm sido reportadas inúmeras inconsistências da confiabilidade por meio do alpha de Cronbach<sup>(35-36)</sup>.

A replicabilidade do construto foi avaliada pelo *Generalized G-H Index*<sup>(37)</sup> com índice maior do que 0,80<sup>(28)</sup> e, para a qualidade e a efetividade da estimativa dos fatores, foram utilizados o Factor Determinacy Index, apontando para uma estimativa adequada valores maiores do 0,90, *EAP marginal reliability* (>0,80), *sensibility ratio* (SR>2) e *Expected percentage of true differences* (EPTD>90%). A aplicação de múltiplos indicadores decorre da necessidade de atestar por diversas técnicas as evidências de validade do instrumento. Além disso, a aplicação e interpretação dos índices de ajustamento do modelo (*goodness-of-fit* - GOF), por si, não garantem que a solução da análise fatorial seja boa ou útil na prática, pois é possível obter índices satisfatórios de solução baseados em itens de baixa qualidade<sup>(38-39)</sup>.

Como forma de aumentar a confiabilidade e a replicabilidade do modelo proposto, aplicou-se a validação cruzada<sup>(40)</sup> e foi aplicada a técnica de *Holdout*<sup>(41)</sup>. A referida técnica divide o banco em uma amostra de treinamento que pode variar de 10%, 30% e 50% e outro conjunto de dados, denominado banco de teste<sup>(41)</sup>. O banco de dados foi dividido em 50/50 com escolha aleatória dos itens. Foi utilizado o *site* Random.org ([www.random.org](http://www.random.org)) e a técnica de *random sequence generator* para a divisão dos grupos. Os bancos foram denominados da seguinte maneira: amostra completa (AC com 783 casos); amostra 1 (A1 - banco de treinamento com 392 casos) e amostra 2 (A2 - banco de teste com 391 casos). Outra modificação é que, usualmente, em casos de aplicação de AFE e AFC, tende-se a utilizar o primeiro banco de treinamento na AFE e o banco de teste na AFC<sup>(22)</sup>. Nesse estudo, optou-se pela aplicação dos procedimentos de análises com o intuito de ampliar as evidências de validade e a qualidade do instrumento. A análise se estendeu para a amostra completa (AC) caso ocorresse o ajuste nas duas amostras. As análises partiram do banco de treinamento (amostra 1) e, em seguida, nos dois outros conjuntos de dados para cada conjunto de técnicas.

As análises foram efetuadas por meio do SPSS 23, AMOS 23 e o Factor 10.8.01.

## Resultados

Participaram do estudo 783 indivíduos, cuja idade média foi de 38,6 (14,5) anos, sendo que 68,1% (n=533) eram do sexo feminino. Além disso, 262 (33,5%) estudaram até o ensino fundamental e 52,7% utilizavam o sistema público de saúde como única forma de acesso aos serviços de saúde. Sobre as características das subamostras, 32,5%, 33,7%, 38% e 39% dos participantes eram do sexo masculino, respectivamente, em Piracicaba, São Paulo, Aparecida de Goiânia e Fortaleza. As médias das idades dos participantes foram de 41,6; 39,6; 38,2 e 38,1 anos, também nesta ordem, em Piracicaba, São Paulo, Aparecida de Goiânia e Fortaleza. Sobre o nível de escolaridade, 82,6%; 78,3%; 78% e 81,6% dos participantes apresentavam até o ensino médio completo, respectivamente, em Piracicaba, São Paulo, Aparecida de Goiânia e Fortaleza. Dos participantes, 52,7% utilizavam o sistema público de saúde como única forma de acesso aos serviços de saúde.

Apenas 91 (1,9%) das 4.698 respostas possíveis foram *missing* e o próprio *software* (*Factor*) faz as simulações de efeitos dos *missing* para corrigir o modelo<sup>(42)</sup>.

Sobre as análises psicométricas do instrumento HLS-EU-Q6, os índices de adequação da amostra baseados

em correlação policóricas indicaram bons níveis de fatorabilidade para as três configurações do banco. Amostra 1: índice de Kaiser-Meyer-Olkin (KMO = 0,82), esfericidade de Bartlett = 314,5 (df = 15;  $P < 0,0001$ ) e o determinante da matriz = 0,19 ( $<0,0001$ ). Amostra 2: KMO = 0,82, esfericidade de Bartlett = 342,8 (df = 15;  $P < 0,0001$ ) e o determinante da matriz = 0,17. Para a Amostra Completa o KMO foi de 0,84, esfericidade de Bartlett = 636,6 (df = 15;  $p < 0,0001$ ) e o determinante da matriz = 0,19 ( $p < 0,0001$ ).

A primeira análise foi centrada no estudo da dimensionalidade/fatores do instrumento e a análise paralela (AP) indicou a existência de apenas uma dimensão para o instrumento com uma variância explicada de 69,92% da variável latente, acima do mínimo recomendado em modelos iniciais<sup>(34)</sup>. Os autovalores também apontaram

para apenas uma dimensão com um autovalor de 3,62. Não houve qualquer indicador de que este conjunto de itens pudesse se alinhar em um modelo multidimensional.

Na amostra 2 a AP indicou a existência de apenas uma dimensão para o instrumento com variância explicada de 68,95%. A mesma unidimensionalidade ocorreu pelo critério dos autovalores (3,76). A análise com o banco completo demonstrou a unidimensionalidade pela AP com variância explicada de 71,23%, e o mesmo ocorreu pelos critérios de Kaiser (autovalor = 3,69). A unidimensionalidade foi confirmada pelos valores de Unico (A1 = 0,98; A2 = 0,98 e AC = 0,99), para o ECV (A1 = 0,86; A2 = 0,87 e AC = 0,89) e o MIREAL (A1 = 0,24; A2 = 0,29 e AC = 0,23). Como forma extensiva de testagem da unidimensionalidade, os índices foram aplicados aos itens e os resultados podem ser visualizados na Tabela 1.

Tabela 1 - Valores de Unico, ECV e MIREAL dos itens da versão brasileira do instrumento HLS-EU-Q6 para as três amostras analisadas. Piracicaba/SP; São Paulo/SP, Aparecida de Goiânia/GO e Fortaleza/CE, Brasil, 2018

ITEM	Unico*			ECV†			MIREAL‡		
	Amostra 1	Amostra 2	Amostra completa	Amostra 1	Amostra 2	Amostra completa	Amostra 1	Amostra 2	Amostra completa
ITEM 1	0,93	0,98	0,98	0,72	0,83	0,85	0,52	0,28	0,30
ITEM 2	0,99	0,99	0,98	0,95	0,87	0,86	0,15	0,33	0,29
ITEM 3	1,00	0,98	0,99	0,99	0,86	0,96	0,05	0,28	0,13
ITEM 4	0,99	0,99	1,00	0,92	0,91	0,98	0,21	0,21	0,10
ITEM 5	0,99	0,98	0,99	0,93	0,85	0,90	0,17	0,30	0,22
ITEM 6	0,98	0,99	0,98	0,84	0,88	0,85	0,35	0,32	0,35

\*Unico = Unidimensional Congruence; †ECV = Explained Common Variance; ‡MIREAL = Mean of item residual absolute loadings

Os resultados da dimensionalidade dos itens indicaram que todos os itens em todas as amostras apresentaram o I-UNICO unidimensional. O I-ECV apresentou uma pequena violação do item 1 para a A1, com todos os outros itens unidimensionais.

No I-REAL houve violação do item 1 na A1 e do item 6, de forma marginal, nas três amostras. Na AC, apenas o item 6 mostrou-se com os resíduos da carga fatorial levemente acima de 0,30. Novamente, os indicadores predominantemente apontaram para a

unidimensionalidade, não havendo qualquer violação desta nessa primeira fase das análises.

O fato de o instrumento ser unidimensional dispensou a realização de técnicas rotacionais da matriz fatorial e indicou a aplicação da técnica de *Normal-ogive graded response model* na TRI, adequada para o modelo unidimensional politômico.

A Tabela 2 apresenta os valores das cargas fatoriais, comunalidades e discriminação do item para as três amostras.

Tabela 2 - Cargas fatoriais, comunalidades e discriminação do item para a versão brasileira do instrumento HLS-EU-Q6. Piracicaba/SP; São Paulo/SP, Aparecida de Goiânia/GO e Fortaleza/CE, Brasil, 2018

ITEM	Carga fatorial ( $\lambda$ )			Comunalidades ( $h^2$ )			Discriminação do item ( $a$ )		
	Amostra 1	Amostra 2	Amostra completa	Amostra 1	Amostra 2	Amostra completa	Amostra 1	Amostra 2	Amostra completa
ITEM 1	0,76	0,64	0,72	0,58	0,41	0,52	1,17	0,83	1,04
ITEM 2	0,67	0,83	0,72	0,45	0,68	0,52	0,91	1,48	1,04
ITEM 3	0,67	0,68	0,70	0,45	0,47	0,49	0,91	0,94	0,99
ITEM 4	0,75	0,72	0,75	0,56	0,52	0,56	1,13	1,04	1,14
ITEM 5	0,68	0,72	0,68	0,46	0,52	0,46	0,92	1,04	0,92
ITEM 6	0,80	0,84	0,83	0,64	0,72	0,67	1,34	1,61	1,44

As cargas fatoriais estabeleceram-se entre 0,64 e 0,84 nas amostras, o que indica níveis satisfatórios e adequados. Não foram encontrados problemas de colinearidade/multicolinearidade e *Heywood Cases*. A colinearidade/multicolinearidade, quando as cargas fatoriais ficam acima de 0,85, pode indicar redundância dos itens, problemas de distribuição dos dados e gerar distorções na mensuração da variável latente. Da mesma forma, não foram encontradas violações do limite das cargas fatoriais (-1 a +1). Este tipo de violação é denominado *Heywood Cases* e é um indicador de possíveis inadequações da amostra, de estimativas impróprias da variância do erro e indefinição do modelo. A ausência desses problemas nos permite afirmar que a solução do modelo foi apropriada e não houve efeitos deletérios decorrentes da amostra e, principalmente, do modelo estabelecido.

Todas as comunalidades foram acima de 0,40, com variação entre 0,41 e 0,72. Todos os itens apresentaram valores de discriminação entre 0,83 a 1,61, variando entre moderada e alta discriminação. Desta forma, as cargas fatoriais ( $\lambda$ ), comunalidades ( $h^2$ ) e discriminação do item ( $a$ ) apresentaram níveis adequados e consistentes para o modelo unidimensional.

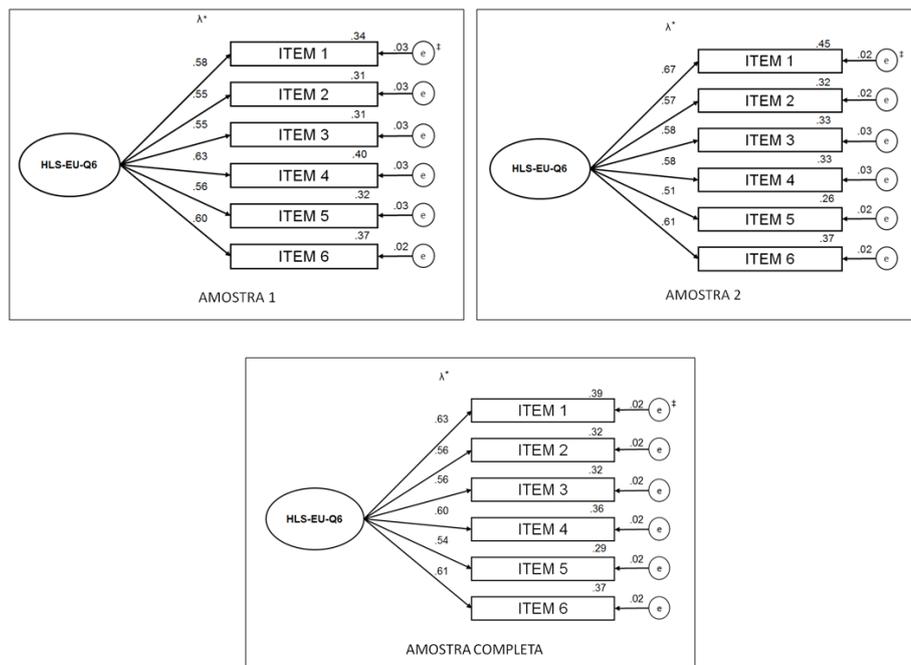
Os valores para os indicadores de confiabilidade para a A1 para o alfa de Cronbach e o ômega de McDonald foram de 0,86. Para A2 foi de 0,87 para o alfa e de 0,88 para o ômega. Na AC foi de 0,87 para ambos os índices.

Quanto à replicabilidade do constructo pelo índice G-H latente e observado, os escores ficaram em 0,87, 0,89 e 0,88 para o G-H latente, respectivamente, para as amostras A1, A2 e AC. O G-H observado ficou em 0,78, 0,77 e 0,79, também na devida ordem, para as amostras A1, A2 e AC. Como houve uma diferença pequena entre os níveis de G-H latente e observado, pode-se inferir uma estabilidade do modelo mesmo quando aplicado a outras amostras populacionais e sua consequente generalização.

Para as medidas de qualidade e efetividade dos escores do modelo, o FDI apresentou um nível elevado e adequado (A1 = 0,93; A2 = 0,95 e AC = 0,88) para avaliar a relação entre a estimativa dos escores da solução e a variável latente a qual eles estimam. O EAP (A1 = 0,87; A2 = 0,89 e AC = 0,93), o SR (A1 = 2,64; A2 = 2,92 e AC = 2,72), e o EPTD (A1 = 91,4%; A2 = 92,3% e AC = 91,6%) também indicaram uma qualidade e efetividade da solução do modelo.

A Análise Fatorial Exploratória e os índices de TRI apontaram para um modelo unidimensional consistente e com manutenção dos níveis satisfatórios em todas as etapas de análises.

Sobre a AFC, o diagrama de caminho foi estabelecido para cada uma das amostras. A Figura 1 apresenta os resultados das cargas fatoriais, poder de predição do item ( $R^2$ ) e erro padrão para a A1, A2 e modelo completo (Figura 1).



\*Carga fatorial; †Erro padrão

Figura 1 - Diagrama de caminhos para cada uma das amostras.

A AFC da A1 apresentou os itens com cargas fatoriais variando entre 0,55 e 0,63, ou seja, acima do mínimo recomendado de 0,50. Os valores preditivos dos itens R2 estabeleceram-se entre 0,31 e 0,40. O diagrama de caminhos da amostra 2 mostrou níveis similares ao do modelo1 (A1) para os indicadores avaliados. As cargas fatoriais variaram de 0,51 a 0,67 com valores preditivos dos itens variando de 0,26 a 0,45, demonstrando, portanto, níveis satisfatórios. O diagrama de caminhos

para a amostra completa também estabeleceu um modelo com níveis adequados para as cargas fatoriais, as quais variaram de 0,54 a 0,63 e com níveis de predição dos itens de 0,29 a 0,37. Tal fato indicou resultados estáveis, satisfatórios e consistentes para as três amostras testadas.

A Tabela 3 apresenta os valores de autovalores pela correlação e covariância, as cargas fatoriais, nível de predição dos itens (R2), resíduos e erro padrão para as três amostras.

Tabela 3 - Autovalores, cargas fatoriais, R2<sup>2</sup>, resíduos e erro padrão dos modelos avaliados para a versão brasileira do instrumento HLS-EU-Q6 para as três amostras avaliadas. Piracicaba/SP; São Paulo/SP; Aparecida de Goiânia/GO e Fortaleza/CE, Brasil, 2018

Índices	Amostra 1	Amostra 2	Amostra completa
Autovalor (r)*	2,70	2,72	2,70
Autovalor (cov) <sup>†</sup>	1,11	1,14	1,12
Cargas fatoriais	0,55 a 0,60	0,51 a 0,67	0,53 a 0,64
R2 <sup>2</sup>	0,30 a 0,37	0,26 a 0,45	0,28 a 0,41
Resíduos	-0,018 a 0,019	-0,033 a 0,051	-0,021 a 0,022
Erro padrão	0,029 a 0,036	0,019 a 0,030	0,021 a 0,026

\*r = Autovalores pela correlação; <sup>†</sup>cov = Autovalores pela variância; <sup>2</sup>R2 = Nível de predição dos itens

É fundamental apontar que nas três amostras testadas na AFC, tanto o cálculo dos autovalores pela correlação quanto da covariância, indicaram que os modelos são unidimensionais. Verificou-se, por meio do conceito de correlação, que os autovalores foram de 2,70, 2,72 e 2,70, respectivamente para a A1, A2 e a AC. Além disso, pelo conceito de covariância, os valores foram de

1,11, 1,14 e 1,12 para as três amostras. Não houve o surgimento de outros autovalores acima de 1.

Na Tabela 4 verificam-se todos os índices globais da AFE, AFC, TRI, confiabilidade, replicabilidade, qualidade e efetividade do modelo. Todos os indicadores analisados apontaram para um modelo unidimensional, consistente, preciso e estável.

Tabela 4 - Síntese do modelo para a versão brasileira do instrumento HLS-EU-Q6. Piracicaba/SP; São Paulo/SP; Aparecida de Goiânia/GO e Fortaleza/CE, Brasil, 2018

	Índice	Técnica	Amostra 1	Amostra 2	Amostra completa
Exploratória	<i>Adequacy of correlation matrix</i>	<i>Determinant of the matrix Bartlett</i>	0,19	0,17	0,19
			314,0 (df = 15)	342,8 (df = 15)	636,3 (df = 15)
		KMO (Kaiser-Meyer-Olkin)	0,82	0,82	0,84
	<i>Explained Variance (Kaiser Criterion)</i>		60,39%	62,68%	61,63%
	<i>Explained Variance (AP)</i>		69,92%	68,95%	71,23%
Confirmatória	<i>Polychoric Correlation (rp = )</i>		0,39 a 0,68	0,37 a 0,69	0,40 a 0,65
	<i>Robust Mean-Scaled Chi Square (X2/ df = 56)</i>		20,19 (df = 9)*	22,69 (df = 9)*	22,54 (df = 9)*
	<i>Non-Normed Fit Index (NNFI)</i>		0,97	0,97	0,98
	<i>Comparative Fit Index (CFI)</i>		0,98	0,98	0,99
	<i>Goodness of Fit Index (GFI)</i>		0,99	0,98	0,99
	<i>Adjusted Goodness of Fit Index (AGFI)</i>		0,98	0,97	0,99
	<i>Root Mean Square Error of Approximation (RMSEA)</i>		0,08	0,08	0,06

	Índice	Técnica	Amostra 1	Amostra 2	Amostra completa
Confiabilidade	<i>Root Mean Square of Residuals (RMSR)</i>		0,05	0,08	0,05
	<i>Standardized Cronbach's Alpha</i>		0,86	0,87	0,87
	<i>McDonald's Omega</i>		0,86	0,88	0,87
	<i>Construct Reliability - Index G H (Latente e observada)</i>		(0,87; 0,78)	(0,89; 0,77)	(0,88; 0,79)
Unidimensionalidade	<i>Unidimensional Congruence (UNICO)</i>		0,98	0,98	0,99
	<i>Explained Common Variance (ECV)</i>		0,86	0,87	0,89
	<i>Mean of item residual absolute loading (MIREAL)</i>		0,24	0,29	0,23
Qualidade efetividade	<i>Factor Determinacy Index (FDI)</i>		0,93	0,94	0,88
	<i>EAP Marginal Reliability</i>		0,87	0,89	0,93
	<i>Sensitivity Ratio (SR)</i>		2,64	2,92	2,72
	<i>Expected percentage of true differences (EPTD)</i>		91,4%	92,3%	91,6%

De acordo com os resultados dos escores do HLS-EU-Q6, apenas 2% dos participantes foram classificados como apresentando níveis suficientes de LS; 51,7% com níveis problemáticos e 46,3% com níveis inadequados.

## Discussão

Os resultados do presente estudo demonstraram que a versão em português do Brasil do HLS-EU-Q6 apresentou característica unidimensional, cargas fatoriais satisfatórias e bons níveis de confiabilidade, os quais apontam para um instrumento com evidências de estrutura interna consistente e confiável para a mensuração do constructo desejado.

Até o presente momento, esse é o primeiro estudo a avaliar a validade do HLS-EU-Q6 por meio de múltiplas técnicas de dimensionamento e de índices de ajustamento do modelo. No *European Health Literacy Study* (HLS-EU), o ajustamento do modelo para o HLS-EU-Q6 foi avaliado pela AFC em subamostras que responderam ao instrumento completo, e observou-se uma estrutura fatorial satisfatória na maior parte das amostras dos países participantes<sup>(7,9)</sup>. A versão brasileira do HLS-EU-Q6, por outro lado, demonstrou características de unidimensionalidade e bom ajustamento na AFE, TRI e AFC em todas as amostras avaliadas na validação cruzada. As técnicas psicométricas aplicadas no estudo são muito mais extensivas e contemporâneas do que nos estudos em outros países.

As análises trazem ainda uma série de indicadores que são raros de serem realizados em estudos psicométricos, alguns por serem recentes, e não estão disponíveis nos *softwares* comerciais e o uso de técnicas mais extensivas de análises dos dados, que incorporam o conceito de evidências de validade da estrutura externa por múltiplos indicadores. Há um avanço nos últimos anos da ampliação de múltiplas técnicas. Conforme apontado<sup>(22)</sup>, poucos

estudos aplicaram múltiplas técnicas para análises de validação. Portanto, há um avanço substancial, mas ainda insuficiente dessa prática, principalmente quando muitos estudos ainda utilizam como critério para a retenção do modelo os Critérios de Kaiser e Scree-Test. Outro fator é a aplicação de uma validação cruzada completa, em vez de parcial, quando se utiliza o banco de treinamento na AFE e o banco de teste na AFC, isto no intuito de que possíveis erros existentes na análise da AFE sejam transpostos para a AFC. Subentende-se aos que não têm profundo conhecimento de psicométrica que a AFC (modelo restrito) é superior a AFE (modelo irrestrito). Deve-se ter claro que a AFC também pode não levar ao ajustamento, em razão de imprecisão do modelo, erro no número de fatores, a omissão de *cross-loading* e erros de correlação<sup>(43)</sup>. A aplicação completa da técnica, assemelhada ao que se faz nos *K-folds*, assegura que o modelo, em razão das subamostras, possa ser extrapolado para populações mais heterogêneas. Ademais, a aplicação do Índice G-H avalia quão bem definida a variável latente está a partir dos itens do instrumento, ou seja, a viabilidade de um modelo de medição dado por um conjunto de itens. Tais análises permitem avaliar a probabilidade do modelo ser estável entre estudos, populações ou subpopulações<sup>(28,44)</sup>. Assim, mesmo na amostra desse estudo, composta por indivíduos de quatro cidades, não houve instabilidade no instrumento, assegurando a qualidade, efetividade, estabilidade e replicabilidade do modelo final em diversos contextos. Além disso, todos os indicadores da TRI nas três amostras estabeleceram-se em níveis adequados, reforçando e legitimando os resultados obtidos nos indicadores primários das análises fatoriais. Estes indicadores fazem parte de um conjunto de análises que atestam a confiabilidade do instrumento.

Neste mesmo caminho, sobre a confiabilidade do instrumento, os valores de alfa encontrados para as três

amostras da versão brasileira demonstraram bons valores, e superiores aos encontrados na amostra total do estudo europeu ( $\alpha=0,803$ )(7,9), em adultos na Itália ( $\alpha=0,672$ )<sup>(13)</sup> e na França ( $\alpha=0,83$ )(14), bem como um estudo com diabéticos na Bélgica ( $\alpha=0,797$ )<sup>(12)</sup>. Além do mais, os valores de ômega corroboraram a confiabilidade da versão brasileira do instrumento.

Dentre as dimensões do LS possíveis de se avaliar com a versão brasileira do HLS-EU- Q6 estão: a avaliação e aplicação das informações de saúde gerais; encontrar, acessar e avaliar as informações para prevenção de doenças e promoção da saúde<sup>(7,9)</sup>. A aplicação de instrumentos para o rastreamento de habilidades de LS no acolhimento nos serviços de saúde tem sido medida recomendada para qualificar a coleta de dados dos usuários e como um meio para que os profissionais de saúde, inclusive os enfermeiros, norteiem seus cuidados, sendo considerado como o "sexto sinal vital"<sup>(1,45)</sup>.

A porcentagem de indivíduos com níveis inadequados de LS no presente estudo (46,3%) foi superior à média encontrada nos países do estudo europeu (9%), bem como na França (5%), Itália (8,9%) e Bélgica (9,8%)<sup>(7,9,12-14)</sup>. Este fato pode ter ocorrido por diferentes características socioeconômicas das populações analisadas, já que os níveis de escolaridade e de renda dos indivíduos nos estudos supracitados eram bem superiores aos do presente estudo. Adicionalmente, as características cognitivas, culturais, da organização do sistema educacional e de saúde podem ter contribuído para estas diferenças<sup>(7,9,13)</sup>. Embora não seja o objetivo desse estudo, a identificação de que apenas 2% dos participantes possuíam níveis suficientes de LS, ou seja, eram capazes de encontrar, acessar, compreender, avaliar e usar as informações de saúde assinala para a importância de os profissionais conhecerem quais as limitações individuais que impactam diretamente no cuidado em saúde.

O presente estudo traz importantes avanços no conhecimento científico relacionado ao processo de validação de instrumentos para a mensuração do LS e, ainda, a disponibilização de um instrumento com extensas evidências de validade para avaliar o LS da população brasileira de rápida e fácil aplicação, fato que poderá viabilizar sua inserção na rotina dos serviços de saúde.

A mensuração do LS por meio do HLS-EU-Q6 poderá auxiliar os profissionais da saúde a redirecionarem as intervenções da área de maneira a identificar as reais necessidades dos usuários dos serviços de saúde, tornando uma nova possibilidade de se pensar e executar o cuidado centrado no paciente. A OMS recomenda que a quantificação das limitações relacionadas ao LS é um importante passo para tal ação.

Dessa forma, recomenda-se que futuros estudos ampliem a aplicação desse instrumento para outras amostras e populações, visando conhecer a classificação dos níveis de letramento (inadequado, problemático e suficiente) para as diversas regiões do país.

## Conclusão

A versão brasileira do instrumento HLS-EU-Q6 indicou evidências de validade de estrutura interna adequadas para aferição dos níveis de letramento em saúde de adultos brasileiros. Configura, portanto, uma ferramenta que pode ser facilmente utilizada na prática clínica capaz de mensurar de maneira rápida e objetiva as limitações no acesso, compreensão e uso das informações em saúde, seja para a prevenção de doenças ou para a promoção de saúde.

## Referências

1. Okan O, Bauer U, Levin-Zamir D, Pinheiro P, Sorensen K, editors. *International Handbook of Health Literacy: Research, practice and policy across the lifespan*. [Internet]. Bristol: Policy Press; 2019 [cited Apr 2, 2020]. Available from: <https://library.oapen.org/handle/20.500.12657/24879>
2. Liu C, Wang D, Liu C, Jiang J, Wang X, Chen H et al. What is the meaning of health literacy? A systematic review and qualitative synthesis. *Fam Med Community Health*. 2020;8(2):e000351. doi:10.1136/fmch-2020-000351
3. Palumbo R. Examining the impacts of health literacy on healthcare costs. An evidence synthesis. *Health Serv Manage Res*. 2017;30(4):197-212. doi: 10.1177/0951484817733366
4. Liu H, Zeng H, Shen Y, Zhang F, Sharma M, Lai W, et al. Assessment Tools for Health Literacy among the General Population: A Systematic Review. *Int J Environ Res Public Health*. 2018;15(8). pii: E1711. doi: 10.3390/ijerph15081711
5. Nguyen TH, Paasche-Orlow MK, McCormack LA. The State of the Science of Health Literacy Measurement. *Stud Health Technol Inform*. 2017;240:17-33. doi: 0.3233/ISU-170827
6. Nutbeam D, McGill B, Premkumar P. Improving health literacy in community populations: a review of progress. *Health Promot Int*. 2018;33(5):901-11. doi: 10.1093/heapro/dax015
7. Pelikan JM, Ganahl K. Measuring Health Literacy in General Populations: Primary Findings from the HLS-EU Consortium's Health Literacy Assessment Effort. In: Logan RA, Siegel ER, editors. *Health Literacy: New Directions in Research, Theory and Practice*. 1st ed. Amsterdam: IOS Press; 2017.
8. Duong TV, Aringazina A, Baisunova G, Nurjanah, Pham TV, Pham KM, et al. Measuring health literacy in Asia:

- Validation of the HLS-EU-Q47 survey tool in six Asian countries. *J Epidemiol.* 2017;27(2):80-6. doi: 10.1016/j.je.2016.09.005
9. Pelikan JM, Röthlin F, Ganahl K, Boltzmann L. Measuring comprehensive health literacy in general populations – the HLS-EU instruments. [Internet]. 2nd International Conference of Health Literacy and Health Promotion; Taipei, Taiwan; 6-8 October 2014 [cited Apr 2, 2020]. Available from: <https://www.bumc.bu.edu/healthliteracyconference/files/2014/06/Pelikan-et-al-HARC-2014-fin.pdf>
10. Amoah PA, Phillips DR, Gyasi RM, Koduah AO, Edusei J. Health literacy and self-perceived health status among street youth in Kumasi, Ghana. *Cogent Med.* 2017;4(1):1275091. doi: <https://doi.org/10.1080/2331205X.2016.1275091>
11. Vandebosch J, den Broucke SV, Schinckus L, Schwarz P, Doyle G, Pelikan J, et al. The impact of health literacy on diabetes self-management education. *Health Educ J.* 2018;77(3):349-62. doi: [doi: 10.1177/0017896917751554](https://doi.org/10.1177/0017896917751554)
12. Schinckus L, Dangoisse F, Van den Broucke S, Mikolajczak M. When knowing is not enough: Emotional distress and depression reduce the positive effects of health literacy on diabetes self-management. *Patient Educ Couns.* 2018;101(2):324-30. doi: 10.1016/j.pec.2017.08.006
13. Lorini C, Lastrucci V, Mantwill S, Vettori V, Bonaccorsi G; Florence Health Literacy Research Group. Measuring health literacy in Italy: a validation study of the HLS-EU-Q16 and of the HLS-EU-Q6 in Italian language, conducted in Florence and its surroundings. *Ann Ist Super Sanita.* 2019;55(1):10-8. doi: 10.4415/ANN\_19\_01\_04
14. Rouquette A, Nadot T, Labitrie P, Van den Broucke S, Mancini J, Rigal L, et al. Validity and measurement invariance across sex, age, and education level of the French short versions of the European Health Literacy Survey Questionnaire. *PLoS One.* 2018;13(12):e0208091. doi: 10.1371/journal.pone.0208091
15. Rigolin CCD, Bastos Jr JC, Mello LC, Carvalho CCB. The Brazilian scientific production of theses and dissertations on health literacy. *R Tecnol Soc.* 2018;14(34):178-95. doi: <http://dx.doi.org/10.3895/rts.v14n34.7599>
16. Maragno CAD, Mengue SS, Moraes CG, Rebelo MVD, Guimarães AMM, Pizzol TDS. Test of health literacy for Portuguese-speaking Adults. *Rev Bras Epidemiol.* 2019;22:e190025. doi: 10.1590/1980-549720190025
17. Soares M. Multiple facets of literacy and initial reading instruction. *Rev Bras Educ.* 2004;(25):5-17. doi: <https://doi.org/10.1590/S1413-24782004000100002>
18. Health Literacy Tool Shed. [Homepage]. [cited Jun 29, 2020]. Available from: [https://healthliteracy.bu.edu/language\\_of\\_validated\\_version=20](https://healthliteracy.bu.edu/language_of_validated_version=20)
19. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine.* 2001;25:3186-91. doi: 10.1097/00007632-200012150-00014
20. Reichenheim ME, Moraes CL. Operationalizing the cross-cultural adaptation of epidemiological measurement instruments. *Rev Saúde Pública.* 2007;41(4):665-73. doi: [doi.org/10.1590/S0034-89102006005000035](https://doi.org/10.1590/S0034-89102006005000035)
21. Hair JR, Black WC, Babin BJ, Anderson R., Tatham RL. *Multivariate data analysis.* 8th ed. Upper Saddle River: Intl Thomson Business; 2018.
22. Goretzko D, Pham TTH, Bühner M. Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Curr Psychol.* 2019;1-12. doi: <https://doi.org/10.1007/s12144-019-00300-2>
23. Timmerman ME, Lorenzo-Seva U. Dimensionality Assessment of Ordered Polytomous Items with Parallel Analysis. *Psychol Methods.* 2011;16(2):209-20. doi: 10.1037/a0023353
24. Choi J, Kim S, Chen J, Dannels S. A comparison of maximum likelihood and Bayesian estimation for polychoric correlation using Monte Carlo simulation. *J Educ Behav Stat.* 2011;36(4):523-49. doi: 10.3102/1076998610381398
25. Finch WH. Using Fit Statistic Differences to Determine the Optimal Number of Factors to Retain in an Exploratory Factor Analysis. *Educ Psychol Meas.* 2020;80(2):217-41. doi: 10.1177/0013164419865769
26. Dobriban E, Owen AB. Deterministic Parallel Analysis: An Improved Method for Selecting Factors and Principal Components. *J Royal Stat Soc-B.* 2019;81(1):163-83. doi: 10.1111/rssb.12301
27. 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;24(4):468-91. doi: 10.1037/met0000200
28. 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;78(5):762-80. doi: [doi: 10.1177/0013164417719308](https://doi.org/10.1177/0013164417719308)
29. Hattori M, Zhang G, Preacher KJ. Multiple Local Solutions and Geomin Rotation. *Multiv Behav Res.* 2017;52(6):720-31. doi: 10.1080/00273171.2017.1361312
30. Samejima, F. Estimation of Latent Ability Using a Response Pattern of Graded Scores. Richmond: Psychometric Society; 1969.
31. Jordan P, Spiess M. Rethinking the interpretation of item discrimination and factor loadings. *Educ Psychol Measur.* 2019;79(6):1103-32. doi: <https://doi.org/10.1177/0013164419843164>
32. Baker FB. *The basics of item response theory.* 2nd ed. ERIC Clearinghouse on Assessment and Evaluation; 2001.

33. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16(3):297-334. doi: 10.1007/BF02310555
34. McDonald RP. *Test theory: A unified treatment*. Mahwah: Lawrence Erlbaum; 1999.
35. Hoekstra R, Vugteveen J, Warrens MJ, Kruijen PM. An empirical analysis of alleged misunderstandings of coefficient alpha. *Int J Soc Res Methodol*. 2019;22(4):351-64. doi: 10.1080/13645579.2018.1547523
36. McNeish D. Thanks coefficient alpha, we'll take it from here. *Psychol Methods*. 2018;23(3):412-33. doi: 10.1037/met0000144
37. Hancock GR, Mueller RO. Rethinking construct reliability within latent variable systems. In: Cudek R, duToit SHC, Sorbom DF, editors. *Structural equation modeling: Present and future*. Lincolnwood: Scientific Software; 2000.
38. Ferrando PJ, Navarro-Gaonzález D, Lorenzo-Seva U. Assessing the quality and effectiveness of the factor score estimates in psychometric factor-analytic applications. *Methodology (Gott)*. 2019;15(3):119-27. doi: 10.1027/1614-2241/a000170
39. Fokkema M, Greiff S. How Performing PCA and CFA on the Same Data Equals Trouble. *Eur J Psychol Assess*. 2017;33(6):399-402. doi: 10.1027/1015-5759/a000460
40. Lee LC, Liang C-Y, Jemain AA. Validity of the best practice in splitting data for hold-out validation strategy as performed on the ink strokes in the context of forensic science. *Microchem J*. 2018;139:125-33. doi: 10.1016/j.microc.2018.02.009
41. Gütlein M, Helma C, Karwath A, Kramer S. A Large-Scale Empirical Evaluation of Cross-Validation and External Test Set Validation in (Q)SAR. *Mol Inform*. 2013;32:516-28. doi: 10.1002/minf.201200134
42. Lorenzo-Seva U, Van Ginkel JR. Multiple imputation of missing values in exploratory factor analysis of multidimensional scales: estimating latent trait scores. *Ann Psychol*. 2016;32(2):596-608. doi: <https://doi.org/10.6018/analesps.32.2.215161>
43. Bollen KA. When Good Loadings Go Bad: Robustness in Factor Analysis. *Struct Equ Model*. 2019;1-10. doi: 10.1080/10705511.2019.1691005
44. Rodriguez A, Reise SP, Haviland MG. Applying Bifactor Statistical Indices in the Evaluation of Psychological Measures. *J Pers Assess*. 2016;98(3):223-37. doi: 10.1080/00223891.2015.1089249
45. Ingram RR, Kautz DD. Creating "Win-Win" Outcomes for Patients with Low Health Literacy: A Nursing Case Study. *Med Surg Nursing*. 2018;27(2):132-4.

---

### Contribuição dos autores:

**Concepção e desenho da pesquisa:** Fábio Luiz Mialhe, Katarinne Lima Moraes, Fernanda Maria Rovai Bado, Virginia Visconde Brasil, Helena Alves De Carvalho Sampaio, Flávio Rebutini. **Obtenção de dados:** Katarinne Lima Moraes, Fernanda Maria Rovai Bado, Virginia Visconde Brasil, Helena Alves De Carvalho Sampaio. **Análise e interpretação dos dados:** Fábio Luiz Mialhe, Flávio Rebutini. **Análise estatística:** Flávio Rebutini. **Redação do manuscrito:** Fábio Luiz Mialhe, Katarinne Lima Moraes, Fernanda Maria Rovai Bado, Virginia Visconde Brasil, Helena Alves De Carvalho Sampaio, Flávio Rebutini. **Revisão crítica do manuscrito quanto ao conteúdo intelectual importante:** Fábio Luiz Mialhe, Katarinne Lima Moraes, Fernanda Maria Rovai Bado, Virginia Visconde Brasil, Helena Alves De Carvalho Sampaio, Flávio Rebutini. **Todos os autores aprovaram a versão final do texto.** **Conflito de interesse: os autores declararam que não há conflito de interesse.**

Recebido: 02.04.2020  
Aceito: 12.09.2020

Editor Associado:  
Ricardo Alexandre Arcêncio

**Copyright © 2021 Revista Latino-Americana de Enfermagem**  
Este é um artigo de acesso aberto distribuído sob os termos da Licença Creative Commons CC BY.

Esta licença permite que outros distribuam, remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito pela criação original. É a licença mais flexível de todas as licenças disponíveis. É recomendada para maximizar a disseminação e uso dos materiais licenciados.

---

Autor correspondente:

Fábio Luiz Mialhe

E-mail: [mialhe@unicamp.br](mailto:mialhe@unicamp.br)

 <https://orcid.org/0000-0001-6465-0959>