

Leandro Pernambuco<sup>1</sup>  
Albert Espelt<sup>2,3,4</sup>  
Hipólito Virgílio Magalhães Junior<sup>5</sup>  
Kenio Costa de Lima<sup>5</sup>

# Recomendações para elaboração, tradução, adaptação transcultural e processo de validação de testes em Fonoaudiologia

## *Recommendations for elaboration, transcultural adaptation and validation process of tests in Speech, Hearing and Language Pathology*

### Descritores

Estudos de Validação  
Validade dos Testes  
Fonoaudiologia  
Psicometria  
Avaliação

### Keywords

Validation Studies  
Validity of Tests  
Speech, Language and Hearing  
Sciences  
Psychometrics  
Evaluation

### Endereço para correspondência:

Leandro Pernambuco  
Departamento de Fonoaudiologia,  
Universidade Federal da Paraíba –  
UFPB Cidade Universitária, Bairro  
Castelo Branco, João Pessoa (PB),  
Brasil,  
CEP: 58051-900.  
E-mail: leandroape@globo.com

Recebido em: Outubro 25, 2016

Aceito em: Novembro 07, 2016

### RESUMO

**Objetivo:** apresentar um guia com recomendações para a tradução, adaptação, elaboração e processo de validação de testes em Fonoaudiologia. **Método:** as recomendações apresentadas foram baseadas em diretrizes internacionais tradicionais cujo enfoque está na elaboração, tradução, adaptação transcultural e processo de validação de testes. **Resultados:** as recomendações foram compiladas em dois quadros, sendo um deles referente aos procedimentos para tradução e adaptação transcultural e o outro à obtenção de evidências de validade, confiabilidade e medidas de acurácia dos testes. **Conclusão:** foi apresentado um guia com as principais recomendações para a organização e sistematização do processo de elaboração, tradução, adaptação transcultural e processo de validação de testes em Fonoaudiologia.

### ABSTRACT

**Purpose:** to present a guide with recommendations for translation, adaptation, elaboration and process of validation of tests in Speech and Language Pathology. **Methods:** the recommendations were based on international guidelines with a focus on the elaboration, translation, cross-cultural adaptation and validation process of tests. **Results:** the recommendations were grouped into two Charts, one of them with procedures for translation and transcultural adaptation and the other for obtaining evidence of validity, reliability and measures of accuracy of the tests. **Conclusion:** a guide with norms for the organization and systematization of the process of elaboration, translation, cross-cultural adaptation and validation process of tests in Speech and Language Pathology was created.

Trabalho realizado no Departamento de Fonoaudiologia, Universidade Federal da Paraíba – UFPB - João Pessoa (PB), Brasil.

<sup>1</sup> Universidade Federal da Paraíba – UFPB - João Pessoa (PB), Brasil.

<sup>2</sup> Agência de Salut Pública de Barcelona – ASPB - Barcelona, Spain.

<sup>3</sup> Universitat Autònoma de Barcelona – UAB - Bellaterra, Spain.

<sup>4</sup> Consorcio de Investigación Biomédica en Red de Epidemiología y Salud Pública – CIBERESP.

<sup>5</sup> Universidade Federal do Rio Grande do Norte – UFRN - Natal (RN), Brasil.

**Fonte de financiamento:** nada a declarar.

**Conflito de interesses:** nada a declarar

## INTRODUÇÃO

Na área da saúde, o adequado processo de avaliação depende da utilização de testes cujas interpretações dos resultados sejam válidas, confiáveis/precisas e equitativas. A validade refere-se à reunião de evidências que indicam se o teste realmente mensura o que propõe mensurar. Já a confiabilidade/precisão indica se o teste é reproduzível ao longo do tempo (estabilidade), se há controle dos erros de mensuração (precisão) e se o resultado do teste é dependente dos itens que o compõem (homogeneidade)<sup>(1-3)</sup>. A equidade permite analisar se o teste avalia as pessoas de forma imparcial, sem permitir que aspectos não relevantes exerçam influência acentuada no desfecho e gerem desigualdades<sup>(1,4)</sup>.

No Brasil, a aplicação desses princípios pela Fonoaudiologia é escassa ou executada de forma parcial, sendo ainda necessário sistematizar o percurso metodológico<sup>(5)</sup>. O processo de validação não é concluído apenas com a elaboração ou tradução e adaptação de um teste. Para que isso aconteça, é necessária a utilização de diretrizes internacionais que garantam a efetiva obtenção das propriedades psicométricas ou clinimétricas do teste.

Apesar das controvérsias existentes entre algumas diretrizes<sup>(6)</sup>, certos princípios metodológicos não podem ser desconsiderados. O objetivo deste manuscrito é apresentar um breve guia de recomendações para a tradução, adaptação, elaboração e processo de validação de testes em Fonoaudiologia.

## MÉTODO

As recomendações aqui apresentadas foram baseadas na experiência dos autores com as mais tradicionais e utilizadas diretrizes internacionais relacionadas à elaboração, tradução, adaptação transcultural e processo de validação de testes. No caso da tradução e adaptação transcultural, foram seguidas as recomendações de Beaton et al.<sup>(7)</sup> e as diretrizes da *International Test Commission* (ITC)<sup>(8)</sup>. No caso do processo de validação de testes, foram utilizados os princípios do *Standards for Educational and Psychological Testing* (SEPT)<sup>(1)</sup>, diretriz proposta por três organizações norte-americanas que compila as mais sólidas e utilizadas recomendações e definições relacionadas aos aspectos psicométricos envolvidos desde a elaboração até a interpretação dos testes.

Como este manuscrito é uma nota metodológica que não envolveu coleta em seres humanos, não foi necessário submeter o estudo à apreciação de comitê de ética.

## RESULTADOS

As recomendações para elaboração, tradução, adaptação transcultural e processo de validação de testes em Fonoaudiologia estão dispostas nos Quadros 1 e 2.

**Quadro 1.** Procedimentos para tradução e adaptação transcultural de testes

Etapa	Procedimentos
<b>a) Diretrizes prévias</b>	<b>a1)</b> Obtenção de permissão dos autores do teste original. <b>a2)</b> Formação de um comitê de especialistas (autores da nova versão e outros especialistas + pelo menos um autor da versão original, se possível) para debater os conceitos adjacentes ao teste a ser adaptado, considerando as características da população e cultura-alvo. Debatedores devem estar cientes dos objetivos do teste.
<b>b) Desenvolvimento do teste</b>	<b>b1) Tradução:</b> Dois tradutores habilitados, nativos no idioma alvo e fluentes no idioma e cultura fonte traduzem o teste de forma independente, considerando a equivalência conceitual e evitando a tradução literal. Recomenda-se que ambos os tradutores desconheçam o teste e que um deles não seja especialista no desfecho investigado pelo teste para que se preserve a representatividade da utilização popular do idioma alvo. <b>b2) Síntese das traduções:</b> feita de forma consensual*, de preferência pelo mesmo comitê mencionado no procedimento a2. O comitê irá construir uma versão única a partir da comparação das traduções e avaliação das discrepâncias semânticas, idiomáticas, conceituais, linguísticas e contextuais. <b>b3) Aplicabilidade da síntese das traduções/equivalência operacional:</b> verifica-se a adequação, estrutura e aplicação dos itens em um contexto real. Não há recomendação explícita para o tamanho amostral nesta etapa. Sugere-se a formação de estratos representativos da população-alvo, compostos com pelo menos dez indivíduos em cada estrato. Realiza-se entrevista cognitiva para verificar se a população compreende os itens do teste. Se o teste for um questionário, recomenda-se a estratégia de paráfrase, na qual o entrevistador faz a pergunta e pede para o entrevistado repeti-la imediatamente. O entrevistador também deve registrar possíveis dificuldades operacionais (por exemplo: tempo de aplicação longo; excessivo número de questões), reações não verbais do entrevistado (por exemplo: desinteresse; expressões faciais) e sugestões do entrevistado para melhorar a compreensão do item. O comitê mencionado no procedimento b2 será novamente convocado se ajustes no teste forem necessários. O teste será reaplicado em novos estratos e esse procedimento se repetirá até que o teste seja aplicado sem nenhuma dificuldade de compreensão. <b>b4) Retradução ou tradução reversa:</b> a versão obtida no procedimento b3 será traduzida para o idioma fonte para avaliar se os itens refletem o conteúdo da versão original. A versão no idioma alvo será encaminhada para pelo menos dois tradutores habilitados que desconheçam o teste e que sejam nativos no idioma fonte e fluentes no idioma e cultura alvo. <b>b5) Síntese das versões retraduzidas:</b> feita de forma consensual*, de preferência pelo mesmo comitê mencionado no procedimento a2. O comitê irá comparar as retraduições ao teste original e avaliar as discrepâncias semânticas, idiomáticas, conceituais, linguísticas e contextuais. <i>Nesta etapa, recomenda-se que os autores da versão original emitam opinião em relação à versão retraduzida.</i> <b>b6) Síntese final:</b> feita de forma consensual*, de preferência pelo mesmo comitê mencionado no procedimento a2. O comitê irá comparar a versão original à versão final no idioma alvo quanto às equivalências semântica, idiomática, conceitual, linguística, experiencial e contextual.

\*Nas avaliações consensuais podem ser adotadas estratégias de qualificação em relação às equivalências utilizando, por exemplo, escala analógica visual ou escalas do tipo Likert. O cálculo do Índice de Validade de Conteúdo geral (IVC) e por Item (IVC-I)<sup>(9)</sup> auxilia nesta etapa como indicador quantitativo de concordância entre os avaliadores

**Quadro 2.** Procedimentos para obtenção de evidências de validade, confiabilidade e medidas de acurácia dos testes

Etapa	Procedimentos
<b>Evidência de validade baseada no conteúdo do teste</b>	O conteúdo refere-se aos temas, redação, formato, tarefas ou questões de um teste, bem como as instruções para os procedimentos necessários para administrá-lo e pontuá-lo <sup>(1,3)</sup> . Algumas estratégias para elaboração de um teste são: extensa revisão de literatura; experiência empírica dos pesquisadores com o desfecho de interesse; entrevista com informantes-chave; realização de painel com especialistas e população-alvo; consulta à população-alvo por meio de grupos focais. Na elaboração dos itens, devem ser considerados os aspectos sintáticos e semânticos que contribuam para a clareza, pertinência, coerência e abrangência dos itens, além dos aspectos operacionais. A representatividade e relevância dos itens em relação ao desfecho devem ser avaliadas por um comitê de, no mínimo, dez juízes com expertise na temática do teste <sup>(2,3)</sup> . Recomenda-se o uso de escala analógica visual ou escalas do tipo Likert para avaliação dos itens do teste pelos juízes. Calcula-se o Índice de Validade de Conteúdo geral (IVC) e por Item (IVC-I) <sup>(9)</sup> para apresentação de um indicador quantitativo de concordância entre os juízes.
<b>Evidência de validade baseada nos processos de resposta</b>	Os processos de resposta resultam das observações ou julgamentos sobre o comportamento ou <i>performance</i> de diferentes estratos da população-alvo durante a aplicação do teste. Nesta etapa, busca-se compreender quais os processos psicológicos, cognitivos e sociais envolvidos na aplicação do teste. Os procedimentos são semelhantes aos descritos em b3 no Quadro 1.
<b>Evidência de validade baseada na consistência interna</b>	Verifica-se o grau de relação entre os itens do teste e o desfecho a partir da aplicação do teste em uma amostra da população-alvo. Caso as dimensões internas do teste estejam teoricamente definidas pelos autores, deve-se realizar a análise fatorial confirmatória (AFC). Se os autores necessitam explorar a possibilidade de reduzir itens e a existência de diferentes dimensões do teste, deve-se realizar análise de componentes principais (ACP) ou análise fatorial exploratória (AFE), com posterior AFC para validar o modelo encontrado. Caso o pesquisador siga a Teoria de Resposta ao Item (TRI) para investigação dos traços latentes, poderá executar essa análise após a AFC. Cada uma das análises mencionadas requer uma amostra diferente. Outros parâmetros adotados são, geralmente, a correlação item-total corrigida e a correlação interitens com valores acima de 0,3 <sup>(2)</sup> . Outra característica investigada nessa fase é o Funcionamento Diferencial dos Itens (FDI), que verifica se variáveis individuais ou de um grupo colocam uma ou mais categorias em vantagem ou desvantagem em relação à(s) outra(s) <sup>(4)</sup> .
<b>Evidência de validade baseada na relação com outras variáveis</b>	Verifica-se o quanto as relações entre o instrumento e outras variáveis externas são consistentes com o desfecho. 1) <i>Validade convergente</i> : análise das relações entre o resultado do teste e outras medidas que avaliam o mesmo desfecho ou desfechos similares; 2) <i>Validade discriminante</i> : o teste é avaliado quanto à sua capacidade de diferenciar grupos distintos da população-alvo ou compará-lo a outras medidas com desfechos diferentes; 3) <i>Validade de critério</i> : o critério é um atributo ou resultado operacionalmente distinto do desfecho do teste e é estabelecido por meio de hipóteses formuladas pelos pesquisadores. Na <i>validade de critério concorrente</i> , o resultado do teste é comparado ao desfecho de outros testes administrados paralelamente e, na <i>validade de critério preditiva</i> , avalia-se o resultado do teste em relação à condição do avaliado no futuro ou o resultado de avaliações posteriores.
<b>Confiabilidade/ Precisão</b>	O teste deve ser administrado em dois momentos distintos, geralmente 7 a 14 dias. Esse período é variável de acordo com as características do desfecho, especialmente a sua estabilidade no tempo. O teste deve ser aplicado pelo mesmo pesquisador e por pesquisadores distintos em ambos os momentos. Será calculado o coeficiente de correlação intraclassa (CCI) ou índice de Kappa para verificar a concordância interavaliadores e intravaliador. A consistência interna é obtida por meio dos valores do alpha de Cronbach, erro de mensuração, correlação item-total e correlação interitens.
<b>Equidade do teste</b>	Especialistas devem julgar se o resultado encontrado previamente na análise de FDI se reproduz de forma justa e garantir que o resultado no teste não seja influenciado por possíveis desvantagens provocadas por variáveis individuais.
<b>Acurácia</b>	Para obtenção dos indicadores sensibilidade, especificidade, valor preditivo positivo e negativo e razão de verossimilhança positiva e negativa, o resultado do teste será comparado ao resultado de outro teste considerado padrão de referência. O ponto de corte do instrumento e o seu poder em discriminar pessoas com e sem o desfecho serão obtidos por meio da curva ROC ( <i>receiver operator characteristic curve</i> ).
<b>Evidência de validade baseada nas consequências do teste</b>	Caracterizada pelo acúmulo de evidências relacionadas ao uso e eficácia do teste. Inclui análises de responsividade do teste a alguma intervenção e considera os estudos subsequentes de aplicação e reprodução do teste no contexto prático relacionado ao desfecho.

## DISCUSSÃO

A elaboração, tradução, adaptação transcultural e processo de validação são procedimentos metodológicos necessários para garantir que as interpretações dos resultados de um teste são válidas e confiáveis. No Brasil, esses procedimentos têm sido reproduzidos com frequência pela Fonoaudiologia, porém nem sempre com o imprescindível rigor científico preconizado pelas diretrizes internacionais.

No caso da tradução e adaptação transcultural, existem cerca de 30 diretrizes, mas nenhum consenso que aponte um único padrão de referência<sup>(6)</sup>. Recomendações propostas nos anos 1990 e atualizadas no ano 2000 são as mais disseminadas na literatura internacional<sup>(7)</sup>. Aqui foram seguidas estas recomendações, além das diretrizes da *International Test Commission Guidelines* (ITC)<sup>(8)</sup>.

Já o SEPT<sup>(10)</sup> existe desde a década de 1950 e especifica, além da confiabilidade/precisão e equidade, cinco fontes de evidência de validade, baseadas em: conteúdo, processos

de resposta, estrutura interna, relação com outras variáveis e consequências do teste. Na área da saúde, as propostas do *Scientific Advisory Committee of the Medical Outcomes Trust (SAC)*<sup>(10)</sup> e do *Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN)*<sup>(11)</sup> são referenciadas com frequência e também podem ser consultadas. Contudo, vale ressaltar que essas duas diretrizes baseiam-se em conceitos psicométricos já não recomendados pelo SEPT desde a edição de 1999, ou seja, são propostas que merecem ser analisadas com parcimônia, especialmente em virtude do marco conceitual.

## CONCLUSÃO

Foram apresentadas as principais recomendações para a organização e sistematização do processo de elaboração, tradução, adaptação transcultural e processo de validação de testes em Fonoaudiologia. Recomendamos a leitura mais aprofundada sobre cada um dos procedimentos aqui mencionados, mas esperamos que essas diretrizes auxiliem a condução de futuras pesquisas na área e possam ajudar o fonoaudiólogo a selecionar de forma mais criteriosa os testes que reúnem as melhores evidências para avaliar um determinado desfecho.

## REFERÊNCIAS

1. AERA: American Educational Research Association, APA: American Psychological Association, NCME: National Council on Measurement in Education. Standards for educational and psychological testing. New York: AERA; 2014.
2. Streiner DL, Norman GL. Health measurement scales: a practical guide to their development and use. 4th ed. New York: Oxford University Press; 2008.
3. Abad FJ, Olea J, Ponsoda V, Garcia C. Measurement in social sciences and health. Madrid: Sintesis; 2011.
4. Espelt A, Viladrich C, Doval E, Aliaga J, García-Rueda R, Tárrega S. Uso equitativo de tests en ciencias de la salud. *Gac Sanit*. 2014;28(5):408-10. PMID:24928357. <http://dx.doi.org/10.1016/j.gaceta.2014.05.001>.
5. Gurgel LG, Kaiser V, Reppold TZ. A busca de evidências de validade no desenvolvimento de instrumentos em Fonoaudiologia: revisão sistemática. *Audiol Commun Res*. 2015;20(4):371-83. <http://dx.doi.org/10.1590/2317-6431-2015-1600>.
6. Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol*. 2015;68(4):435-41. PMID:25698408. <http://dx.doi.org/10.1016/j.jclinepi.2014.11.021>.
7. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25(24):3186-91. PMID:11124735. <http://dx.doi.org/10.1097/00007632-200012150-00014>.
8. Muñiz J, Elosua P, Hambleton RK. Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*. 2013;25(2):151-7. PMID:23628527.
9. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health*. 2006;29(5):489-97. PMID:16977646. <http://dx.doi.org/10.1002/nur.20147>.
10. Aaronson N, Alonso J, Burnam A, Lohr KN, Patrick DL, Perrin E, et al. Assessing health status and quality of life instruments: attributes and review criteria. *Qual Life Res*. 2002;11(3):193-205. PMID:12074258. <http://dx.doi.org/10.1023/A:1015291021312>.
11. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-45. PMID:20494804. <http://dx.doi.org/10.1016/j.jclinepi.2010.02.006>.

## Contribuição dos autores

LP e HVMJ contribuíram com a concepção, escrita e revisão final do artigo; AE e KCL contribuíram com a revisão final do artigo.