

Autora | Author

Nayara D. M. A. Agapito Khouri\*  
Jonas Carvalho e Silva\*\***REVISÃO NARRATIVA: METODOLOGIAS DE ADAPTAÇÃO  
E VALIDAÇÃO DE INSTRUMENTOS PSICOLÓGICOS****NARRATIVE REVIEW: METHODOLOGIES FOR ADAPTATION  
AND VALIDATION OF PSYCHOLOGICAL INSTRUMENTS**

**Resumo:** O artigo descreve os procedimentos metodológicos utilizados na adaptação e validação de instrumentos psicológicos, e exemplifica com resultados de pesquisas. Foram abordadas as metodologias que preconizam: a avaliação de seis equivalências - conceitual, de itens, semântica, operacional, de mensuração e funcional; apontam os cinco passos para a adaptação: (1) tradução para o idioma alvo, (2) síntese de versões traduzidas, (3) avaliação da versão síntese por especialistas, (4) avaliação da versão síntese pelo público-alvo e (5) estudo-piloto; e ainda as três etapas fundamentais para a validação: mensuração da validade convergente, análise fatorial e análise da fidedignidade por meio do coeficiente Alfa de Cronbach. Esse complexo processo é apresentado nas metodologias de quatro estudos de adaptação e validação dos instrumentos no Brasil, Portugal, Estados Unidos e Canadá com idiomas similares e distintas. As discussões articuladas contribuirão para a capacitação de estudantes, pesquisadores e profissionais da saúde.

**Palavras chave:** adaptação, validação, instrumentos psicológicos.

**Abstract:** *The paper describes the methodological procedures used for adaptation and validation of psychological instruments and provides examples in previous studies. We present the methodologies that preach the need to : assess six equivalences - conceptual, items, semantic, operational, measurement and functional; point out five steps to the adaptation process: (1) translation to the target language; (2) translated versions synthesis; (3) evaluation version synthesis by experts; (4) evaluation summary version by the target audience and (5) pilot study; and further, three key steps for validation: measurement of convergent validity, factor analysis and reliability analysis through Cronbach's alpha coefficient. Such a complex process is described in the method of four studies of adaptation and validation of instruments in Brazil, Portugal, USA and Canada with similiar and diferent languages. The discussion produced is important to instruct students, researchers and clinical professionals.*

**Keywords:** adaptation, validation, instruments for management of scientific activity.

## INTRODUÇÃO

Em processos de investigação de um construto psicológico é comum a utilização de instrumentos específicos, sendo os inventários de autorrelato, questionários e escalas os procedimentos mais referidos, principalmente em contextos de intervenção clínica ou avaliação para procedimentos na área da saúde. Entretanto, muitas vezes, uma ferramenta proposta para medir um construto não está disponível naquele contexto ou cultura e uma alternativa viável é a utilização de versões de instrumentos, já devidamente testados em outros países e que se propõem a executar o mesmo tipo de investigação (CASSEPP-BORGES; BALBINOTTI; TEODORO, 2010; DUARTE; BORDIN, 2000). Com a intenção de melhor conhecer esse processo, o objetivo desta revisão narrativa da literatura científica é descrever os procedimentos metodológicos mais utilizados para adaptação e validação dos instrumentos estrangeiros de avaliação psicológica para a cultura brasileira.

Os psicólogos clínicos que atuam em unidades de saúde, ambulatorial e hospitalar, utilizam os instrumentos padronizados em programas de triagem, para identificar a gravidade dos casos e, desta forma, possibilitar melhor direcionamento dos recursos disponíveis (DUARTE; BORDIN, 2000). Além disso, a testagem psicológica, dentro da avaliação clínica, auxilia o processo do diagnóstico, definição de condutas terapêuticas e o planejamento da intervenção mais sistematizada, o que leva às decisões mais prudentes sobre o contexto avaliado (CFP, 2013; ANADRADE; SALES, 2017). Por outro lado, a carência de instrumentos formais, objetivos e adaptados a determinados contextos, prejudica e compromete a eficiência de uma intervenção, ou tratamento em geral (GIUSTI; BEFILLI, 2008).

O processo de adaptação de um instrumento consiste em um conjunto de etapas, previamente planejadas, que avaliam se um dado instrumento consegue medir o mesmo construto, em línguas e contextos culturais diferentes. Nesse processo, tradutores e pesquisadores devem identificar conceitos, termos e expressões que sejam culturalmente, psicologicamente e linguisticamente equivalentes, em uma segunda língua. Além disso, a adaptação inclui outras avaliações sobre a equi-

valência da forma adaptada, a testagem semântica e de compreensão da população a ser avaliada (HAMBLETON, 2005).

No Brasil, a maior parte dos testes e escalas, adotados em situações de avaliação clínica e em pesquisas, é oriunda de outras culturas (CASSEPP-BORGES; BALBINOTTI; TEODORO, 2010; SILVA et al., 2014). A aplicação de instrumentos produzidos e adaptados em contextos culturais diferentes, sem o uso adequado das técnicas de tradução, adaptação cultural e equivalência semântica, pode provocar mensurações distorcidas as condições sócio-emocionais dos sujeitos. Por essa razão, todo o processo de adaptação e tradução deve seguir diretrizes e procedimentos específicos (CASSEPP-BORGES et al., 2010).

## DIRETRIZES PARA UTILIZAÇÃO DE TESTES PSICOLÓGICOS ESTRANGEIROS

O Conselho Federal de Psicologia (CFP), na “Resolução 002 de março de 2003”, definiu os testes psicológicos como procedimentos sistemáticos para observação e registro, que tem como objetivo mensurar características e processos psicológicos, tais como emoção/afeto, cognição/inteligência, motivação, personalidade, psicomotricidade, atenção, memória e percepção, dentre outros. Essa Resolução prevê que a utilização dos testes estrangeiros, de qualquer natureza, demandam estudos com as populações brasileiras, em respeito as evidências de validade, precisão e dados normativos com o ambiente original do instrumento (CFP, 2003).

Segundo o *International Test Commission* (ITC, 2005) as diretrizes para a adaptação dos testes e de outros instrumentos psicológicos devem ser consideradas aos termos de comparação para os códigos nacionais já existentes, conferindo-lhes reconhecimento e consistência internacional. O processo de tradução e adaptação deve compreender os contextos linguísticos e culturais da população alvo, mantendo qualidade comparável ao original. Dentre as diretrizes destacam-se: (a) o uso da língua deve ser apropriado para todas as populações culturais e linguísticas às quais o instrumento se destina; (b) a escolha da técnica, formato dos itens e dos procedimentos, deve ser familiar a todas as populações a que se destinam; (c) o conteúdo dos itens e dos materiais de estímulo devem ser familiares a todas as populações destinadas; e (d) deve haver a garantia de que o projeto de coleta de dados permite o uso das técnicas estatísticas adequadas para estabelecer uma equi-

valência de itens entre as diferentes versões linguísticas do(s) teste(s) ou instrumento(s) (ITC, 2005).

Antes de executar o procedimento técnico chama-se a atenção para as questões éticas que devem ser consideradas (CASSEPP-BORGES; BALBINOTTI; TEODORO, 2010), entre elas a questão de direitos autorais. O(s) autor(es) do instrumento original deve(m) ser consultado(s) previamente e conceder autorização formal aos autores da adaptação. Aos participantes de pesquisa devem ser garantidos o sigilo dos dados de identificação, a utilização das informações exclusivamente para a pesquisa e a isenção de qualquer responsabilidade com relação às suas opiniões expressas. Além disso, o projeto de pesquisa deve ser aprovado por um Comitê de Ética em Pesquisa credenciado pela Comissão Nacional de Ética em Pesquisa do Ministério da Saúde.

## ETAPA DA ADAPTAÇÃO CULTURAL DOS INSTRUMENTOS

Mesmo com as diretrizes estabelecidas para o processo formal de adaptação transcultural, não há consenso quanto às estratégias de execução (REICHENHEIM; MORAES, 2007). Verifica-se um conjunto de recortes nos procedimentos oriundos das diversas fontes e da experiência dos pesquisadores. Os autores propõem uma sistemática operacional para o processo de adaptação transcultural dos instrumentos desenvolvidos em outros contextos linguísticos e socioculturais, com prioridade para avaliação de seis tipos de equivalência: conceitual, de itens, semântica, operacional, de mensuração e funcional.

Observa-se também que, além da psicologia, outras áreas das ciências da saúde adotam a mesma metodologia nos processos de validação e adaptação dos instrumentos psicológicos. Em concordância, foi testada a equivalência conceitual dos itens e semântica na versão brasileira do instrumento *Dysfunctional Voiding Scoring System* (DVSS) para a avaliação de disfunção do trato urinário inferior em crianças (RIZZINI et al., 2009). Nesse estudo a primeira etapa do processo de adaptação, após a tradução, foi a avaliação da equivalência conceitual, na qual a pertinência dos conceitos e a capacidade de assimilação do instrumento na cultura alvo foram avaliados. Em seguida, indicaram a equivalência dos itens com o objetivo de verificar se eles representam as mesmas dimensões na população alvo. O procedimento ocorreu por meio de dis-

cussão com um grupo de profissionais que atuam na área da nefropediatria (RIZZINI et al., 2009).

No mesmo estudo, a etapa subsequente foi a avaliação da equivalência semântica, que verificou se a tradução mantinha significado do conteúdo de cada item e se provocou efeitos semelhantes nos respondentes das duas culturas. Para tanto, submeteram o instrumento original e a versão traduzida à avaliação de dois tradutores com formação em Letras e especialização em língua inglesa e a outros dois profissionais da área de saúde com fluência em inglês (RIZZINI et al., 2009).

A equivalência operacional (REICHENHEIM; MORAES, 2007) está relacionada ao formato dos itens e das questões, contextos e ambientes da administração, modo(s) de aplicação e da quantificação dos dados. Deve-se e garantir o respeito aos padrões e formatos básicos mesmo que não sejam absolutamente idênticos ao original. Por conseguinte, a avaliação de equivalência verifica as propriedades psicométricas do instrumento adaptado em relação ao original. Nessa etapa, a avaliação explora a estrutura dimensional e a adequação dos itens por meio de métodos multivariados tais como a análise fatorial exploratória (AFE). A equivalência de mensuração se adequa aos conceitos de validação e por essa razão, optou-se por detalhá-la no tópico correspondente. Por último está a equivalência funcional, que revela se a eficácia geral do instrumento é satisfatória em mais de uma cultura (REICHENHEIM; MORAES, 2007).

Outra proposta para adaptação cultural de instrumentos psicológicos (BORSA; DAMÁSIO; BANDEIRA, 2012) contempla seis etapas: 1) tradução do instrumento por pelo menos dois tradutores bilíngues e independentes, com as devidas considerações linguísticas, culturais, contextuais e científicas sobre o constructo; 2) comparação entre as versões traduzidas por especialistas e pesquisadores da área. Nessa fase, há consenso com a proposta anterior (REICHENHEIM; MORAES, 2007) pois são feitas as avaliações das equivalências semântica e conceitual. Ademais, são incluídas duas equivalências, a idiomática (a tradução de itens de dificuldade sem alteração do significado cultural do mesmo) e a experiencial (quando se avalia se determinado item de um instrumento é aplicável à nova cultura).

Na continuidade das seis etapas, 3) avaliação da síntese por juízes especialistas, que analisam aspectos ligados à estrutura, *layout*, instruções e abrangência. Os juízes devem possuir conhecimento e experiência na área de abrangência do constru-

to; 4) avaliação pelo público-alvo para verificar se os itens, as instruções e a escala de resposta são compreensíveis para a população local ao qual se destina. Solicita-se a leitura em voz alta das questões e uma breve explicação do significado de cada um dos itens. Outra opção é o preenchimento independente do instrumento para futura discussão sobre a compreensão de cada item com os pesquisadores (BORSA; DAMÁSIO; BANDEIRA, 2012).

O quinto passo é a tradução reversa (*Back-translation*), que é a devolução da versão sintetizada e revisada do instrumento para o idioma de origem a fim de verificar se a nova versão espelha o conteúdo dos itens, conforme propõe a versão original e, ainda, apresentar a versão para a avaliação do autor do instrumento original. Essa etapa também deve ser realizada por dois tradutores bilíngues e independentes (GIUSTI; BEFI-LOPES, 2008). E por último, 6) acontece a realização do estudo piloto, quando o instrumento é aplicado em uma pequena amostra proporcional e similar à população-alvo. Deve-se avaliar a compreensão dos participantes em relação às instruções, preenchimento, significado de cada item e tempo de aplicação. Novas sugestões de alteração podem ser discutidas entre os pesquisadores e os especialistas até a proposta da versão final do instrumento (BORSA; DAMÁSIO; BANDEIRA, 2012).

Pacico (2015) segue diretrizes similares às descritas acima, porém considera necessário um número maior de juízes no processo de adaptação, principalmente na fase de comparação da versão traduzida com a original. Outra diferença está na etapa de avaliação pelo público alvo, onde a autora sugere que o procedimento deva envolver grupos focais.

Em concordância com os esses procedimentos de validação, 24 estudos foram selecionados no período de 2000 a 2010, sobre adaptação de instrumentos de avaliação psicológica para o Brasil (MANZI-OLIVEIRA et al., 2011). Todos os estudos realizaram procedimentos de tradução, acompanhados da retrotradução e de avaliação da pertinência da versão produzida, com ajustes semânticos e conceituais. Pequenas variações de procedimentos foram descritas. Na fase de tradução, por exemplo, a maioria dos estudos (15) tiveram dois tradutores profissionais e independentes com língua materna portuguesa. Os demais trabalhos adotaram diversas formas, tais como: tradução por dois profissionais brasileiros e um nativo da língua inglesa; dois tradutores bilíngues e um professor de língua portuguesa; dois profissionais da equipe de

pesquisa mais um professor de língua inglesa; um tradutor bilíngue; e, ainda, dois pesquisadores da própria equipe do estudo. Para a retrotradução, em 13 estudos, um ou dois brasileiros bilíngues foram responsáveis. Em outros oito estudos, a retrotradução foi feita por tradutores ingleses (seguindo o estudo original), cuja língua materna era o inglês, enquanto um estudo contou com a participação de uma equipe multidisciplinar e bilíngue.

As etapas técnicas posteriores foram constituídas, basicamente, de avaliações por comitês multiprofissionais, com análises semânticas, comparações de versões preliminares e aplicações de pré-testes. Em três estudos foram realizadas comparações das versões, análises de sua equivalência e consenso entre os tradutores. Em quatro estudos, além disso, foram incluídas aplicações de pré-testes do instrumento. Os demais estudos se preocuparam em examinar as versões, utilizando comitês multiprofissionais, comparando-as e examinando seus conteúdos, ou apenas realizaram uma revisão de forma (MANZI-OLIVEIRA et al., 2011).

Podemos observar que as publicações convergem para as diretrizes indicadas (CFP, 2003; ITC, 2005), porém com variações nas aplicações, o que vai de encontro com as considerações sobre a falta de consenso, ou protocolo padrão, nas estratégias de execução do processo formal de adaptação transcultural.

## ETAPA DE VALIDAÇÃO DOS INSTRUMENTOS

O trabalho realizado por Imada et al. (2010), analisou a dinâmica familiar e identificou as mudanças na interação familiar, decorrentes do surgimento do câncer de mama, que poderiam prejudicar o ajustamento e a qualidade de vida das mulheres/pacientes e dos familiares. Para tanto, executaram a adaptação e a validação do instrumento *Family Dynamics Measure II* (FDM II), que identifica os dados individuais e familiares favoráveis ao reconhecimento das normas familiares e a identificação dos perfis das famílias, além de sugerir estratégias mais eficientes de intervenção. Na etapa da validação, os autores optaram por realizar a análise da validade do construto através da análise da validade convergente da FDM II com a Escala HADS (Escala de Ansiedade e Depressão Hospitalar) e da análise fatorial, além da análise da fidedigni-

dade, ou estabilidade da medida, por meio do cálculo do coeficiente Alfa de Cronbach.

Validade de construto diz respeito à capacidade do instrumento refletir de fato o domínio que pretende investigar (KAZDIN, 2010). Nesta perspectiva, uma das alternativas é utilizar outro instrumento capaz de avaliar o mesmo construto e mensurar correlação entre ambos. Os resultados obtidos por meio de outro teste válido predizem o mesmo desempenho que o teste a ser validado e servem de critério para determinar a validade do novo teste (PASQUALI, 2009). Assim, quando os dois instrumentos evidenciam a validade quando convergem ou apresentam indicadores elevados e significativos de correlação (SHAUGHNESSY; ZECHMEISTER; ZECHMEISTER, 2012).

No caso da pesquisa de Imada et al. (2010), os valores obtidos nas correlações entre as medidas de ansiedade e depressão da HADS, em comparação com as medidas dessas mesmas dimensões da versão adaptada da FDM II, apresentaram sentido inverso e intensidade de moderada a baixa, com  $p < 0,001$ . Sugerem, portanto, alta validade entre os instrumentos e apontam que quanto maior a ansiedade, mais desorganizada a família, mais distorcida e confusa a comunicação entre seus membros e maior a probabilidade de conflitos entre os papéis desempenhados por eles.

A outra fase do processo de validação é a análise fatorial, denominada de etapa da equivalência de mensuração (REICHENHEIM; MORAES, 2007). Análise fatorial e validade de construto têm sido associadas durante muito tempo, sendo imprescindível no processo da validação de instrumentos psicológicos. A validade do construto é reforçada se a estrutura fatorial da escala, isto é, as dimensões ou todas as questões que abordam um determinado construto/ fator, são consistentes com o que o instrumento se propõe a medir. A confirmação de uma estrutura fatorial hipotetizada é melhor estabelecida com técnicas de análise fatorial confirmatória, que permitem testar o ajuste relativo a modelos concorrentes (LAROS, 2005).

Para exemplificar a utilização da análise fatorial no processo de validação, a Escala *Center for Epidemiologic Studies Depression Scale* (CES-D) foi avaliada em uma amostra de estudantes universitários, partindo de uma versão brasileira, já existente da escala (FILHO; TEIXEIRA, 2011). A CES-D é um dos instrumentos mais referidos no mundo para a avaliação de sintomatologia depressiva em diversas faixas etárias. O estudo original desse instrumento revelou uma estrutura de quatro fatores oblíquos, sendo eles: afetos positivos (ava-

lia aspectos de otimismo, esperança e satisfação de vida), depressão (avalia afetos negativos), aspectos somáticos/atividade reduzida (contém indicadores relacionados a dificuldades de se engajar e manter atividades cotidianas) e problemas interpessoais (avalia crenças negativas que trazem dificuldades no funcionamento social). Cada fator aborda dimensões relativas ao construto Transtorno Depressivo, sendo que as questões específicas da estrutura de cada fator buscam abranger os principais aspectos afetivos, cognitivos, somáticos e comportamentais da depressão.

Os autores citam vários estudos realizados a fim de investigar a validade de construto e a estrutura fatorial subjacente à CES-D em uma variedade de contextos culturais, faixas etárias e versões traduzidas do instrumento. A principal divergência é a retirada ou junção de fatores, ou seja, a mudança na estrutura fatorial da escala, que sugere modelos unidimensionais, modelo de dois fatores ou estrutura de três fatores. Devido a tais diferenças, ao amplo uso e potencial do instrumento, o estudo testou estruturas fatoriais concorrentes e contribuiu com evidências de validade de construto para a CES-D.

O estudo utilizou análises fatoriais confirmatórias, através do *software* AMOS 18.0, para estimar o ajuste dos modelos concorrentes (modelos de um, dois, três e quatro fatores), relatados em estudos prévios. Foi utilizado um questionário sociodemográfico e os 20 itens da CES-D avaliados em escala Likert de quatro pontos. Os índices de ajuste eleitos para comparação dos modelos foram: (a) qui-quadrado com menores valores indicando menores diferenças entre o modelo e a matriz de correlações dos dados; (b) qui-quadrado normado; (c) *Akaike Information Criterion* (AIC) sem valores de referência, mas sendo o melhor modelo aquele com menor valor; (d) *TuckerLewis Index* (TLI) e *Comparative Fit Index* (CFI) com recomendação de valores  $\geq 0,90$  para um ajuste aceitável; e (e) *Root Mean Square Error of Approximation* (RMSEA), com valores aceitáveis abaixo de 0,08. Os resultados revelaram que o modelo original de quatro fatores obteve o melhor ajuste aos dados porque apresentou menos resíduos do que os demais (qui-quadrado e RMSEA), maior ajuste incremental (TLI e CFI) e maior parcimônia, considerando o número de parâmetros estimados e os graus de liberdade (AIC e qui-quadrado normado) (FILHO; TEIXEIRA, 2011).

Na análise fatorial da primeira pesquisa apresentada (IMADA et al., 2010) não apresentou concordância dos fatores obtidos com a proposta teórica das autoras sobre a dimensionalidade da escala. Alguns itens tiveram cargas fatoriais menores que 0,3 - sendo consideradas baixas por estarem próxi-



mas de zero. Com a exclusão desses itens, verificou-se uma melhora de desempenho no percentual de variância (36,7% x 39,4%), sugerindo a adaptação do instrumento sem os itens com cargas fatoriais menores que 0,3. As hipóteses para o resultado de falta de concordância são: (a) na análise fatorial da primeira versão (FDM) não houve confirmação da dimensionalidade teórica do instrumento, o que seria indicativo de uma característica do próprio instrumento; (b) o construto “dinâmica familiar” é complexo e envolve um conjunto de dimensões que se interrelacionam, sendo difícil distingui-los claramente, o que é comum em construtos psicológicos compostos por múltiplas facetas correlacionadas. Muitas vezes não se conformam aos modelos que requerem dessas facetas apresentar cargas em um único fator; e (c) o número elevado de itens por dimensão pode ser mais um dos motivos para a dificuldade em confirmar a dimensionalidade teórica proposta. Mesmo com as justificativas apresentadas, não foi confirmada a dimensionalidade do instrumento como um todo.

Dentro da etapa de validação de um instrumento psicológico, também se configura como essencial a avaliação da confiabilidade da estrutura fatorial, sendo o Alfa de Cronbach o método mais utilizado em estudos transversais dos quais as medições são realizadas em um único momento (DAMÁSIO, 2012).

A confiabilidade, fidedignidade ou precisão, de um teste diz respeito à capacidade de medir sem erros, ou seja, a capacidade de medir os mesmos participantes em ocasiões diferentes. Os testes equivalentes medem os mesmos participantes, na mesma ocasião, produzem os resultados idênticos, a correlação entre estas duas medidas deve ser igual a um. A análise da precisão de um instrumento psicológico quer mostrar o quanto ele se afasta do ideal da correlação igual a um, determinando um coeficiente que, quanto mais próximo de um, menos erro implica na aplicação (PASQUALI, 2009). Assim, os coeficientes de fidedignidade fornecem informações a respeito da magnitude do erro que pode influenciar os escores (URBINA, 2004).

No procedimento para o cálculo do coeficiente Alfa de Cronbach avalia-se a correlação entre respostas de um instrumento (correlação média entre perguntas), através da análise do perfil das respostas fornecidas pelos participantes da pesquisa. Considerando que todos os itens de um questionário utilizam a mesma escala de medição, o coeficiente é calculado através de uma equação, que utiliza a variância dos itens

individuais e a variância da soma dos itens de cada participante (HORA; MONTEIRO; ARICA, 2010). Geralmente, o índice obtido varia entre zero e um, sendo as diretrizes para interpretação dos valores do alfa mais comumente adotadas as que sugerem  $\alpha > 0,90$  = excelente;  $\alpha > 0,80$  = bom;  $\alpha > 0,70$  = aceitável;  $\alpha > 0,60$  = questionável;  $\alpha > 0,50$  = pobre; e  $\alpha < 0,50$  = inaceitável. É importante destacar que o valor do Alfa de Cronbach é influenciado tanto pelo valor das correlações dos itens, quanto pelo número de itens avaliados (DAMÁSIO, 2012).

Hora; Monteiro & Arica (2010) chamam a atenção para confusão entre validade e confiabilidade, presente na literatura científica. Enquanto a validade tenta verificar se um instrumento realmente mede aquilo que se propõe a medir, a confiabilidade está relacionada com a isenção de erros aleatórios. Como ainda não há um formalismo matemático para afirmar que uma escala é válida ou não os pesquisadores avaliam a validade da escala pelo nível de confiabilidade. Todavia um alto grau de confiabilidade não significa, necessariamente, que o instrumento seja válido. Esta distinção é clara para o cálculo do Alfa de Cronbach, que contempla três pressupostos básicos: (a) o questionário deve estar dividido e agrupado em dimensões, questões que tratam de um mesmo aspecto; (b) o questionário deve ser aplicado a uma amostra significativa e heterogênea; e (c) a escala já deve estar validada.

O coeficiente Alfa de Cronbach é bem aceito na comunidade científica, o que favorece sua adoção como ferramenta para estimação da confiabilidade. No Brasil, seu uso é frequente na área de saúde, principalmente devido à ampla utilização em questionários de avaliação de sintomas, percepção de problemas, efeitos de intervenção psicológica e/ou satisfação com serviços recebidos (HORA; MONTEIRO; ARICA, 2010).

Além da análise fatorial, análise da validade convergente e do cálculo do coeficiente de Alfa de Cronbach, referidos neste texto, outros testes estatísticos também são utilizados como métodos de validação de um instrumento psicológico, dependendo dos objetivos da pesquisa.

Uma dessas técnicas, comumente utilizada no processo de validação de um instrumento, é a de Correlação de Pearson, que visa avaliar a consistência interna e/ou a análise de convergência entre medidas diferentes de um mesmo construto (WECHSLER; SCHELINI, 2006). Correlação pode ser entendida como a relação linear que mantém uma ligação entre

duas variáveis. Ela mensura a direção e o grau da relação entre as duas variáveis quantitativas. Assim, o coeficiente de correlação de Pearson ( $r$ ) é uma medida de associação linear entre variáveis e exige um compartilhamento de variância (FILHO; JÚNIOR, 2009). O coeficiente de correlação Pearson ( $r$ ) varia de menos um (-1) a um (1), sendo que uma correlação perfeita (-1 ou 1) indica que o escore de uma variável pode ser determinado, exatamente, ao se saber o escore da outra, assim, quanto mais perto de um, ou de menos um, maior é o grau de dependência estatística linear, ou força da relação, entre as variáveis (FILHO; JÚNIOR, 2009).

Outra medida utilizada no processo de validação é o coeficiente de correlação intraclasse (*Intraclass Correlation Coefficient* - ICC), uma ferramenta estatística utilizada para a mensuração da confiabilidade de medidas. O ICC avalia a homogeneidade de duas ou mais medidas e é interpretado como a proporção da variabilidade total atribuída ao objeto medido, sendo muito utilizado para mensurar a precisão de instrumentos de medida. Ela avalia, através da variação e em relação ao grau de semelhança existente entre indivíduos de origens diferentes, ou desempenhos de origens diferentes, o grau de semelhança entre o desempenho de indivíduos interligados por uma origem comum (LAUREANO, 2011).

Assim como no processo de adaptação, também não se observa significativo consenso sobre quais e quantas evidências de validade um instrumento deve demonstrar para ser considerado válido (BORSA; DAMÁSIO; BANDEIRA, 2012).

## OUTROS EXEMPLOS DE ADAPTAÇÃO E VALIDAÇÃO DE INSTRUMENTOS

Para melhor ilustrar o complexo processo de adaptação e validação de instrumentos propostos originalmente em línguas e culturas distintas da brasileira, são descritas as metodologias de outros quatro estudos. Dois se referem à adaptação e validação de instrumentos de língua inglesa para o Brasil. O terceiro exemplo traz um estudo aplicado a Portugal, que foi incluído com a intenção de descrever o mesmo processo em outro país de língua portuguesa. E o último exemplo ilustra como se dá o processo entre os Estados Unidos da América e o Canadá, dois países de língua inglesa.

### Exemplo I

O estudo para traduzir, adaptar e identificar evidências da validade de construto do Communities That Care Youth

Survey (CTCYS), um instrumento que analisa fatores de risco e de proteção para o envolvimento com drogas, violência e delinquência (CORRÊA, 2014). Participaram da pesquisa 917 pessoas.

O processo de adaptação do instrumento, seguiu cinco das seis etapas propostas (BORSA; DAMÁSIO; BANDEIRA, 2012), com exceção da tradução reversa: (1) três traduções do instrumento para o idioma alvo; (2) síntese das versões traduzidas; (3) avaliação da versão síntese por especialistas na temática avaliada pelo instrumento; (4) avaliação da versão síntese pelo público-alvo (pré-teste cognitivo); (5) discussão dos termos traduzidos com um pesquisador integrante do grupo que desenvolveu o questionário; e (6) estudo-piloto.

Na etapa de validação, os dados foram submetidos a Análise Fatorial Exploratória (AFE) para verificar sua estrutura fatorial, bem como a análise da consistência interna, através do Alfa de Cronbach.

### Exemplo II

Reis et. al (2016) conduziram dois estudos com 561 sujeitos para adaptar e validar o *Emotion Regulation Checklist* (ERC), um instrumento de heterorrelato que se propõe avaliar o nível de regulação emocional de crianças. Na etapa de adaptação foram feitas três traduções por juízes independentes, do inglês para o português, e posteriormente duas autoras do estudo compararam as versões e a sintetizaram. A nova versão foi apresentada a um comitê de três especialistas e depois aplicada em um grupo focal. Finalmente, foi feita a retro-tradução e apresentada para os autores do instrumento.

Para verificar as propriedades psicométricas do ERC, consistência interna e evidências de validade convergente, os autores utilizaram análise fatorial exploratória (com o método Hull de retenção e interpretação dos fatores) e correlações de Pearson entre a escala do estudo e as escalas da SSRS-BR.

### Exemplo III

Pesquisa realizada em Portugal com 707 jovens adultos, traduziu, adaptou e descreveu as propriedades psicométricas do questionário “Estilo de Vida Fantástico”, um instrumento autoaplicável, que explora hábitos e comportamentos relacionados a estilos de vida saudável (SILVA; BRITO; AMADO, 2014).

Para o processo de adaptação o questionário foi traduzido para o português, a partir da sua versão em espanhol (do Chile), sendo posteriormente efetuada a tradução reversa. Este processo foi simplificado devido a existência de estudos

de validação em espanhol e português brasileiro, com indicação de semelhanças linguísticas e culturais. Assim, foi realizada uma análise comparativa das versões existentes nestes países para se alcançar a equivalência semântica. A versão piloto foi aplicada a um grupo de 12 pessoas, sendo solicitado que fizessem comentários e sugestões sobre os itens do questionário.

Na fase de validação os autores analisaram a validade convergente através da comparação como um questionário sobre características sócio demográficas (gênero, idade, graduação) e o inventário OMEV (O Meu Estilo de Vida). Utilizou-se correlação de Pearson e o coeficiente alfa de Cronbach para a mensuração da consistência interna. A reprodutibilidade do questionário e a consistência externa foram avaliadas pela correlação intraclasse (ICC), sendo testada item a item, domínio a domínio, e para o score geral do “Estilo de Vida Fantástico” (EVF) foram analisados dados da aplicação em dois momentos distintos, teste e reteste.

#### Exemplo IV

O *Psychosocial Assessment Tool* (PAT) é uma ferramenta para triagem de risco psicossocial, desenvolvida e validada com uma amostra de cuidadores de crianças recém-diagnosticadas com câncer nos Estados Unidos da América. O estudo teve como objetivo avaliar a adaptação cultural, a validade e confiabilidade do PAT com uma amostra canadense (BARRERA et al., 2014). Na fase de adaptação cultural, o instrumento foi aplicado a sete países, em formato de entrevista, que foram gravadas e transcritas. As transcrições foram revisadas por dois investigadores e sugestões feitas, por dois ou mais países, foram integradas ao instrumento.

Para avaliar a confiabilidade teste-reteste, 43 países completaram o questionário em dois momentos diferentes. Para avaliar a validade, 65 países completaram medidas adicionais, sendo dois questionários sobre a criança com câncer: o Sistema de Avaliação de Comportamento da Criança-2 (BASC-2) e o *Pediatric Quality of Inventory-Cancer Module* (PedsQL), e, ainda, um questionário sobre a sua própria ansiedade - Inventário de Ansiedade Traço-Estado (IDATE).

Para análises estatísticas, Barrera et al. (2013) optaram pelo cálculo do coeficiente de correlação intraclasse (ICC), no caso do teste de confiabilidade teste-reteste; a consistência interna foi calculada usando coeficientes e Kuder-Richardson KR-20, dada a natureza dicotômica da ferramenta. Para realizar as

análises de validação, foi utilizada a Correlação de Pearson, com os dados do PAT e os outros instrumentos listados.

Em todos os exemplos referidos os autores admitem a necessidade de mais estudos, porém, consideram os instrumentos culturalmente adequados e com boas evidências de validade para suas adaptações.

## CONSIDERAÇÕES FINAIS

Os exemplos referidos corroboram com a existência de extensa variabilidade entre os processos de adaptação e validação de um instrumento psicológico, principalmente na escolha das técnicas estatísticas durante a etapa de validação. Nas fases de tradução, a quantidade de tradutores envolvidos varia entre os estudos, a tradução reversa e a consulta a um comitê de especialistas não foi apontada como obrigatória por todos os autores. Nas etapas de validação, a comparação com outros instrumentos que avaliam construtos semelhantes (para mensurar a validade convergente) não foi referida em todos os estudos, nem foram apontados critérios para seleção de outros instrumentos, e, ainda, o tipo de análise estatística também apresentou grande variabilidade.

Essa variabilidade pode ser útil para a escolha do procedimento a ser utilizado, que deve considerar as características do instrumento, dos contextos de sua aplicação (tanto da versão original como da sua adaptação) e da população a quem se destina. Elas devem estar de acordo com a realidade da população e incluir ferramentas disponíveis ao pesquisador. Os contextos dos pesquisadores também são diversos, e a possibilidade de ajustar tal diversidade, de modo a permitir pesquisas semelhantes em diferentes culturas, é fundamental para a produção científica no contexto mundial, e, assim, democratizar o conhecimento científico.

Este manuscrito espera diminuir a carência de informações sistematizadas no processo de adaptação e validação de instrumentos psicológicos e ratificar a importância da formação em pesquisa nos currículos brasileiros de Psicologia. Replicar estudos dessa natureza; utilizar diferentes métodos para a adaptação e validação de um mesmo instrumento, e posteriormente compará-los; incentivar e divulgar estudos sobre adaptação e validação de instrumentos estrangeiros poderiam levantar questões importantes sobre quais procedimentos são mais adequados, além de melhorar o contexto de variabilidade metodológica deste processo.



## REFERÊNCIAS

- ANDRADE, J. M.; SALES, H. F. S. (2017). A diferença entre avaliação psicológica e testagem psicológica: questões emergentes . In M. R. C. Lins, J. C. Borsa (Orgs.), **Avaliação Psicológica: Aspectos Teóricos e Práticos** (p. 9-22). Petropolis: Vozes.
- BARKER, C.; PISTRANG, N.; & ELLIOT, R. (2015). **Research methods in clinical psychology: an introduction for students and practitioners**. West Sussex: John Wiley & Sons.
- BARRERA, M. et al. External validity and reliability of the psychosocial assessment tool (PAT) among Canadian parents of children newly diagnosed with cancer. **Pediatric Blood and Cancer**, v. 61, n. 1, p. 165–170, 2014.
- BORSA, J. C.; DAMÁSIO, B. F.; BANDEIRA, D. R. Cross-Cultural Adaptation and Validation of Psychological Instruments. **Paidéia**, v. 22, n. 53, p. 423–432, 2012.
- CASSEPP-BORGES, V.; BALBINOTTI, M. A.; TEODORO, M. M. Tradução e validação de conteúdo: Uma proposta para a adaptação de instrumentos. In: PASQUALI, L. (Ed.). **Instrumentação psicológica: Fundamentos e práticas**. Porto Alegre: Artmed, 2010. p. 506–520. CFP. Resolução 002 de março de 2003DOU 24 de Março de 2003, , 2003.
- CORRÊA, A. DE O. Adaptação e validação do communities that care youth survey (CTCYS) para uma comunidade brasileira: um estudo-piloto. [s.l.] Universidade de Brasília, 2014.
- DAMÁSIO, B. F. Uso da análise fatorial exploratória em psicologia. **Avaliação Psicológica**, v. 11, n. 2, p. 213–228, 2012.
- DUARTE, C. S.; BORDIN, I. A. S. Instrumentos de Avaliação. **Rev. Bras. Psiquiatr**, v. 22, n. Supl II, p. 55–58, 2000.
- FILHO, D. B. F.; JÚNIOR, J. A. D. S. Desvendando os mistérios do coeficiente de correlação de Pearson (r). **Revista Política Hoje**, v. 18, n. 1, p. 115–146, 2009.
- FILHO, N. H.; TEIXEIRA, M. A. P. A estrutura fatorial da Escala CES-D em estudantes universitários brasileiros. **Avaliação Psicológica**, v. 10, n. 1, p. 91–97, 2011.
- GIUSTI, E.; BEFI-LOPES, D. M. Tradução e adaptação transcultural de instrumentos estrangeiros para o Português Brasileiro (PB). **Pró-Fono Revista de Atualização Científica**, v. 20, n. 3, p. 207–210, 2008.
- HAMBLETON, R. K. Issues, Designs and Technical Guidelines for Adapting Tests Into Multiple Languages and Cultures. In: HAMBLETON, R. K.; MERENDA, P. F.; SPIELBERGER, C. D. (Eds.). **Adapting Psychological and Educational Tests for Cross-Cultural Assessment**. New Jersey: Lawrence Erlbaum, 2005. p. 03-37.
- HORA, H. R. M.; MONTEIRO, G. T. R.; ARICA, J. Confiabilidade em questionários para qualidade: um estudo com o coeficiente alfa de Cronbach. **Produto & Produção**, v. 11, n. 1973, p. 85–103, 2010.
- IMADA, T. C. M. L. et al. Adaptação e validação da Family Dynamics Measure II para familiares de mulheres com câncer de mama. **Psicologia: Teoria e Pesquisa**, v. 26, n. 3, p. 557, 2010.
- ITC. **ITC Guidelines for Translating and Adaptating Tests**. [s.l.: s.n.]. Disponível em: <<http://www.intestcom.org/upload/sitefiles/40.pdf>>.
- KAZDIN, A. E. **Research Design in Clinical Psychology**. 4. ed. New York: Allyn&Bacon, 2010.
- LAROS, J. A. O Uso da Análise Fatorial: Algumas Diretrizes para Pesquisadores. In: PASQUALI, L. (Ed.). **Análise fatorial para pesquisadores**. Brasília: LabPAM Saber e Tecnologia, 2005. p. 163–193.
- LAUREANO, G. H. DA C. **Coefficiente de Correlação Intraclasse: Comparação entre métodos de estimação clássico e bayesianos**. [s.l.] Universidade Federal do Rio Grande do Sul, 2011.
- MANZI-OLIVEIRA, A. B. et al. Adaptação transcultural de instrumentos de avaliação psicológica: levantamento dos estudos realizados no Brasil de 2000 a 2010. **Psico-USF**, v. 16, n. 3, p. 367–381, 2011.
- PASQUALI, L. Psicometria. **Rev Esc Enferm USP**, v. 43, p. 992–999, 2009.
- REICHENHEIM, M. E.; MORAES, C. L. Operacionalização de adaptação transcultural de instrumentos de aferição usados em

epidemiologia. **Revista de Saude Publica**, v. 41, n. 4, p. 665–673, 2007.

RIZZINI, M. et al. Equivalência conceitual , de itens e semântica da versão brasileira do instrumento Dysfunctional Voiding Scoring System ( DVSS ) para avaliação de disfunção do trato urinário inferior em crianças. **Cad. Saúde Publica**, Rio de Janeiro, v. 25, n. 8, p. 1743–55, 2009.

SHAUGHNESSY, J. J.; ZECHMEISTER, E. B.; ZECHMEISTER, J. S. **Metodologia de Pesquisa em Psicologia**. Porto Alegre: AMGH Editora, 2012.

SILVA, A. M. M.; BRITO, I. DA S.; AMADO, J. M. DA C. Tradução, adaptação e validação do questionário Fantastic Lifestyle Assessment em estudantes do ensino superior. **Ciência & Saúde Coletiva**, v. 19, n. 6, p. 1901–1909, 2014.

URBINA, S. **Fundamentos da Testagem Psicológica**. Porto Alegre: Artmed, 2004.

WECHSLER, S. M.; SCHELINI, P. W. Bateria de Habilidades Cognitivas Woodcock-Johnson III : Validade de constructo. **Psicologia: Teoria e Pesquisa**, v. 22, n. 3, p. 287–296, 2006.

## CURRÍCULOS

\* <http://lattes.cnpq.br/1731299679750211>

\*\* <http://lattes.cnpq.br/9382335692312792>