

PsicoUSF

Psico-USF

ISSN: 1413-8271

revistapsico@usf.edu.br

Universidade São Francisco

Brasil

Batistuta Manzi-Oliveira, Alana; Belinassi Balarini, Fernanda; da Silva Marques, Letícia
Aparecida; Pasian, Sonia Regina

Adaptação transcultural de instrumentos de avaliação psicológica: levantamento dos
estudos realizados no Brasil de 2000 a 2010

Psico-USF, vol. 16, núm. 3, diciembre, 2011, pp. 367-381

Universidade São Francisco

São Paulo, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=401036087013>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica

Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal

Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

Adaptação transcultural de instrumentos de avaliação psicológica: levantamento dos estudos realizados no Brasil de 2000 a 2010

Alana Batistuta Manzi-Oliveira – Universidade de São Paulo, Ribeirão Preto, Brasil

Fernanda Belinassi Balarini – Universidade de São Paulo, Ribeirão Preto, Brasil

Letícia Aparecida da Silva Marques – Universidade de São Paulo, Ribeirão Preto, Brasil

Sonia Regina Pasion¹ – Universidade de São Paulo, Ribeirão Preto, Brasil

Resumo

Diante da busca de aprimoramento técnico-científico nos processos de avaliação psicológica, este trabalho apresenta um panorama das investigações científicas brasileiras acerca da adaptação de instrumentos de avaliação psicológica. Com base em levantamento bibliográfico dos estudos publicados no período de 2000 a 2010, foram selecionados 24 estudos que constituíram o *corpus* analisado. A maioria dos trabalhos seguiu diretrizes internacionais e nacionais de adaptação de instrumentos, sendo que 24 estudos realizaram processo de tradução, 18 apresentaram análise de fidedignidade e 16, evidências de validade. Pôde-se depreender, dentre os estudos analisados, que há uma variedade de procedimentos empregados para a adaptação de instrumentos de avaliação psicológica no contexto brasileiro. Ainda, puderam ser observadas algumas diferenças quanto à utilização de termos referentes à taxonomia psicométrica. Nesse sentido, pretende-se oferecer elementos para reflexão de pesquisadores e profissionais no que diz respeito à adaptação de instrumentos de avaliação psicológica no Brasil.

Palavras-chave: Avaliação psicológica, Adaptação transcultural, Revisão da literatura, Validade, Precisão.

Cross-cultural adaptation of psychological assessment instruments: Brazilian literature review from 2000 to 2010

Abstract

Given the search for technical-scientific improvement in the psychological assessment processes, this paper presents an overview of the Brazilian research about the cross-cultural adaptation of psychological assessment instruments. Based in a literature review of studies published between the period from 2000 to 2010, 24 studies were selected that formed the corpus analyzed. Most studies have followed international and national guidelines, among which 24 studies performed translation process, 18 showed analysis of reliability and 16 studies presented validity evidences. It can be concluded, among the analyzed studies, that there are a variety of procedures employed for the adaptation of psychological assessment instruments in the Brazilian context. Still, some differences were observed regarding the use of terms referring to the psychometric taxonomy. It is intended to provide reflection elements to researchers and professionals regarding the adaptation of psychological assessment instruments in Brazil.

Keywords: Psychological assessment, Cross-cultural adaptation, Literature review, Validity, Reliability.

A adaptação de instrumentos de avaliação psicológica tem ganhado espaço no âmbito da pesquisa científica atual, como uma importante área para o desenvolvimento da prática e da ciência psicológica. A literatura internacional salienta que hoje existem evidências suficientes para afirmar a necessidade de versões multilíngues de instrumentos, entre elas o aumento do número de pesquisas multinacionais e multiculturais (Beaton & cols., 2002; Hambleton, 2005). Além disso, é possível enumerar vantagens para adaptação de testes em detrimento da construção de instrumentos, como a economia de tempo e de recursos financeiros, a possibilidade de comparação de estudos entre grupos de várias culturas e diferentes linguagens e o alcance na equidade de avaliação, em

termos de métodos e de comparabilidade entre os escores (Hambleton, 2005).

No Brasil, a maior parte dos instrumentos utilizados na área clínica e de pesquisa em psicologia é oriunda de outras culturas, despontando a adaptação de instrumentos avaliativos em detrimento de sua construção (Cassepp-Borges, Balbinotti & Teodoro, 2010). Espera-se ainda um aumento substancial da adaptação de instrumentos no futuro, tendo em vista que os intercâmbios internacionais de investigação científica são cada vez mais comuns e o interesse em pesquisas interculturais é gradativamente crescente (Hambleton, 2005).

Com esse crescente intercâmbio de informações e as diferenças internacionais quanto a políticas sobre o uso e a comercialização de instrumentos de avaliação psicológica, a *International Test Commission* (ITC) criou diretrizes para uso correto dos mesmos, buscando encorajar melhores práticas no processo de avaliação psicológica. As chamadas *Diretrizes para o uso de testes* (ITC, 2003) buscam conciliar diferentes práticas internacionais e foram elaboradas a

¹ Endereço para correspondência:
Departamento de Psicologia – FFLRP – USP
Av. Bandeirantes, 3900 – Monte Alegre – 14040-901 – Ribeirão Preto (SP)
E-mail: spasion@ffclrp.usp.br

partir de espaços criados com o objetivo de promover a prática adequada na adaptação de instrumentos de avaliação psicológica em diferentes contextos culturais. No Brasil, a Resolução 02/2003 do Conselho Federal de Psicologia (CFP) (2003) teve como um de seus documentos de referência o material produzido pela ITC e, na direção do que propõem as diretrizes, salienta que os requisitos mínimos para instrumentos de avaliação psicológica construídos no Brasil também se aplicam àqueles advindos do exterior. Isso pode ser visto no seguinte trecho desta citada resolução:

Art. 7º – Também estão sujeitos aos requisitos estabelecidos na presente Resolução os testes estrangeiros de qualquer natureza, traduzidos para o português, que devem ser adequados a partir de estudos realizados com amostras brasileiras, considerando a relação de contingência entre as evidências de validade, precisão e dados normativos com o ambiente cultural onde foram realizados os estudos para sua elaboração. (CFP, 2003)

As exigências abrangem, de maneira geral, a apresentação da fundamentação teórica do construto, das características psicométricas do instrumento, dos procedimentos de aplicação, correção e das referências para a interpretação dos escores (CFP, 2003). Isso significa que os instrumentos de avaliação psicológica advindos de outras culturas, ainda que atendam a tais critérios na população original, devem passar por procedimentos científicos e técnicos que garantam suas qualidades no uso com a população brasileira. Nesse sentido, o maior desafio da adaptação de instrumentos de avaliação diz respeito à equivalência do construto medido pelo teste, em ambas as versões (Hambleton, 2005; Pedroso, 2004, citado por Cassepp-Borges, Balbinotti & Teodoro, 2010).

O conjunto das considerações precedentes circunscreve uma área necessária de investimento científico nos processos de avaliação psicológica: os cuidados metodológicos no processo de adaptação de seus instrumentais. Essa necessidade já se encontra sedimentada internacionalmente, de modo que uma tentativa de retrato da realidade nacional nessa direção se configura como possibilidade de contribuição a este campo de pesquisa e de atuação profissional. Nesse sentido, este trabalho procura apresentar um panorama das investigações científicas brasileiras voltadas aos processos de adaptação de instrumentos de avaliação psicológica, de modo a esclarecer as estratégias utilizadas e os alcances dessas iniciativas nacionais, abrindo possibilidades para futuros trabalhos nessa área, de maneira a assegurar sua qualidade técnica e o compromisso de oferecimento de melhores alternativas de ação profissional.

Dessa forma, o presente trabalho tem como objetivo apresentar e debater os achados de um levantamento bibliográfico, pautado na literatura científica nacional, acerca dos processos de adaptação de instrumentos de avaliação psicológica. Desse modo, buscou-se verificar e mapear as práticas utilizadas no Brasil, relativas à adaptação desses instrumentos, com base em trabalhos publicados no período de 2000 a 2010.

Método

Procedimento

A revisão da literatura científica nacional sobre adaptação de instrumentos de avaliação psicológica foi realizada a partir do estabelecimento de critérios de inclusão e exclusão dos estudos e da análise crítica de seus resultados. Para operacionalizar essa revisão, as seguintes etapas foram seguidas: (a) seleção da questão temática; (b) escolha da base de dados; (c) estabelecimento dos critérios para a seleção/exclusão dos estudos; (d) análise e interpretação dos resultados; (e) apresentação da revisão bibliográfica em si.

A fim de se elaborar um levantamento bibliográfico abrangente, foi realizada uma busca no portal de pesquisa da Biblioteca Virtual em Saúde (BVS), que reúne 14 bases de dados bibliográficas em Ciências da Saúde, com acesso livre e gratuito. A busca eletrônica foi efetivada usando das palavras-chave “instrumentos psicológicos” e “adaptação”, e o operador booleano “e” (and), no período de janeiro de 2000 a outubro de 2010. Antes da realização dessa busca, foi feita uma pesquisa na lista de descritores que o portal disponibiliza, de modo a examinar quais seriam os termos mais adequados para atender ao objetivo do estudo. Os descritores encontrados tendiam a restringir a busca de estudos como, por exemplo, no caso de “validade de testes”, “estudos de validação” ou “tradução”, que não alcançavam, necessariamente, a identificação de trabalhos voltados à adaptação de instrumentos de avaliação psicológica. Sendo assim, optou-se pela utilização das palavras-chave “instrumentos psicológicos” e “adaptação”, que permitiriam a recuperação de grande número de trabalhos no tema em foco, procurando mapear os estudos publicados na última década.

Em seguida, as publicações identificadas nessa busca inicial foram submetidas a critérios de inclusão e de exclusão para o presente trabalho. Nesta revisão foram selecionados exclusivamente artigos científicos publicados em periódicos indexados nas bases bibliográficas e excluídos trabalhos como teses, dissertações, livros e capítulos de livro. Os artigos que apareceram mais de uma vez, em diferentes

combinações de palavras-chave ou em mais de uma base consultada, foram computados uma única vez.

Além dos artigos repetidos, foram excluídos artigos que não tratavam de estudos de adaptação transcultural de instrumentos de avaliação psicológica para a população brasileira; estudos que utilizavam instrumentos de avaliação, porém sem o foco em aspectos metodológicos (não eram sobre o instrumento); estudos ligados a outras áreas do saber (como engenharia, administração), trabalhos de tradução e adaptação cultural para o português de Portugal, estudos de revisão de literatura, propostas de operacionalização de procedimentos de adaptação transcultural e estudos que se propunham à construção de instrumentos de avaliação psicológica.

Os artigos científicos assim selecionados e recuperados foram integralmente lidos e analisados, tomando por base um roteiro organizado nos seguintes eixos: (a) justificativa, objetivos e procedimento realizado; (b) características do instrumento; (c) procedimentos de tradução; (d) procedimentos de verificação das qualidades psicométricas e (e) controles metodológicos. A seguir, os achados foram sistematizados em tabelas descritivas, constituindo-se nos resultados a serem apresentados e comentados.

Resultados e discussão

A partir da busca bibliográfica efetuada foram identificados 93 trabalhos científicos. Os respectivos resumos desses 93 artigos foram estudados e neles aplicados os critérios de inclusão e exclusão determinados. Assim, ao final desse processo, 26 artigos foram selecionados e 24 foram integralmente recuperados, compondo os resultados básicos aqui analisados, conforme itens avaliativos a seguir apresentados.

Justificativa, objetivos e procedimentos

Os 24 estudos recuperados neste levantamento bibliográfico tiveram como objetivo adaptar para o Brasil instrumentos psicológicos de avaliação já existentes em outros contextos, sendo em sua maioria originalmente apresentados em língua inglesa. A Tabela 1 apresenta e caracteriza, em termos gerais, os estudos aqui identificados, relatando brevemente o instrumento de avaliação psicológica em foco, os objetivos e os procedimentos técnicos desenvolvidos nessas pesquisas.

Tabela 1. Estudos brasileiros de adaptação transcultural de instrumentos de avaliação psicológica no período de 2000 a 2010 (integralmente recuperados, n=24)

					Continua
Autoria	Ano	Instrumento	Tradução e Retrotradução	Fidedignidade	Validade
Leme, Saccol, Barbosa, Eijnisman, Faloppa & Cohem	2010	Athletic Shoulder Outcome Rating Scale	2 Traduções 1 Retrotradução Comitê de Juízes Estudo Piloto	Teste-Retest	Construto (convergente)
Moriguchi, Alem, Veldhoven & Coury	2010	Need for Recovery Scale	3 Traduções 2 Retrotraduções Comitê de Juízes	Teste-Retest e Consistência Interna	Construto (convergente)
Sardinha, Levitan, Lopes, Perna, Esquivel, Griez & Nardi	2010	Questionário de Atividade Física Habitual (QAFH)	2 Traduções 2 Retrotraduções Equivalência Semântica	Não realizada	Não realizada
Medeiros & Guerra	2009	Activities of Daily Living Questionnaire (ADLQ)	2 Traduções 1 Retrotradução	Consistência Interna	Construto (convergente) e Conteúdo (juízes)
Neves, Muniz - Caldas, Medeiro, Moraes & Gonçalves	2009	Behçet's Disease Current Activity Form (BDCAF)	2 Traduções 2 Retrotraduções	Reprodutibilidade Inter-observador	Critério (concorrente)
Gambaro, Santos, Thé, Castro & Cendoroglo	2009	The Geriatric Pain Measure	2 Traduções 2 Retrotraduções Comitê de Juízes	Não realizada	Conteúdo (juízes)
Gonçalves & Pilon	2009	Spirituality Self Rating Scale (SSRS)	2 Traduções 1 Retrotradução Comitê de Juízes	Consistência Interna	Construto (convergente)
Mengarda, Passos, Picon, Costa & Picon	2008	Endometriosis Health Profile Questionnaire – EHP-30	2 Traduções 1 Retrotradução	Teste-Retest e Consistência Interna	Construto (convergente), Conteúdo e Face

Tabela 2. Estudos brasileiros de adaptação transcultural de instrumentos de avaliação psicológica no período de 2000 a 2010 (integralmente recuperados, n=24)

Continuação					
Autoria	Ano	Instrumento	Tradução e Retrotradução	Fidedignidade	Validade
Sardinha, Nardi & Eifert	2008	Cardiac Anxiety Questionnaire	2 Traduções 2 Retrotraduções Equivalência Semântica Estudo Piloto	Não realizada	Não realizada
Levitan, Nascimento, Freire, Mezzasalma & Nardi	2008	Social Avoidance And Distress Scale (SADS)	2 Traduções 2 Retrotraduções Equivalência Semântica Estudo Piloto	Não realizada	Não realizada
Mendes, Travassos, Martins & Marques	2008	Canadian Adverse Events Study (CAES)	2 Traduções 1 Retrotradução Comitê de Juízes Estudo Piloto	Não realizada	Não realizada
Sant'Anna, Blascovi-Assis & Magalhães	2008	Entrevista Inicial com os Pais (EIP) e Avaliação do Comportamento Lúdico da Criança (ACL)	2 Traduções 2 Retrotraduções Equivalência Conceitual Estudo Piloto	Fidedignidade do Avaliador	Não realizada
Torriani, Teixeira, Pinheiro, Goettems & Bonow	2008	Dental Anxiety Scale (DAS) e Behavior Rating Scale	2 Traduções 1 Retrotradução (comitê multidisciplinar) Estudo Piloto	Teste-Reteste	Face e Conteúdo
Novato, Grossi & Kimura	2008	Diabetes Quality of Life Measure for Youths	2 Traduções 2 Retrotraduções Equivalências Semântica, Idiomática, Cultural e Conceitual	Teste-Reteste e Consistência Interna	Construto (convergente), Critério (concorrente) e Conteúdo (juízes) Construto (convergente)
Pereira, Costa, Geocze, Borim & Ciconelli	2007	GERD-HRQL, HBQOL e GSAS	2 Traduções 1 Retrotradução Comitê de Juízes Estudo Piloto	Teste-Reteste e Reprodutibilidade Inter-observadores	
Goldfeld, Wiethaeuper, Terra, Baumgardt, Laueremann, Mardini, Abuchaim, Sordi, Soares & Ceitlin	2007	Inventory of Countertransference Behavior (ICB)	2 Traduções 1 Retrotradução Equivalências Conceitual e Semântica	Não realizada	Conteúdo (juízes) e Face
Ciconelli, Soárez, Kowalski & Ferraz	2006	Work Productivity and Activity Impairment - General Health (WPAI-GH) Questionnaire	1 Tradução 1 Retrotradução Comitê de Juízes Estudo Piloto	Teste-Reteste e Consistência Interna	Construto (convergente)
Echevarría-Guanilo, Rossi, Dantas & Santos	2006	Burns Specific Pain Anxiety Scale – BSPAS e Impact of Event Scale – IES	2 Traduções 2 Retrotraduções Comitê de Juízes Análise Semântica Estudo Piloto	Consistência Interna	Construto (convergente)
Torres, Hortale & Schall	2005	Diabetes Knowledge Scale (DKN-A) e Attitudes Questionnaires (ATT-19)	2 Traduções 2 Retrotraduções Comitê de Juízes Equivalência Semântica Estudo Piloto	Teste-Reteste e Consistência Interna	Não realizada
Bandeira, Calzavara & Varela	2005	Family Burden Interview Schedule/Short Form – FBIS	1 Tradução 1 Retrotradução Comitê de Juízes Estudo Piloto	Não realizada	Não realizada

Tabela 3. Estudos brasileiros de adaptação transcultural de instrumentos de avaliação psicológica no período de 2000 a 2010 (integralmente recuperados, n=24)

					Conclusão
Autoria	Ano	Instrumento	Tradução e Retrotradução	Fidedignidade	Validade
Camargo & Contel		Não cita o nome, mas diz que é composto por dois questionários	1 Tradução 1 Retrotradução Comitê de Juízes	Não realizada	Não realizada
Yusuf, Jorge, Habr-Gama, Kiss & Rodrigues	2004	FIQL – Fecal Incontinence Quality of Life	2 Traduções 2 Retrotraduções Comitê de Juízes Estudo Piloto	Teste-Retest e Reprodutibilidade entre examinadores	Construto (convergente) e Critério (concorrente)
Haase, Lacerda, Lima, Corrêa, Brito & Lana-Peixoto	2004	Inventário de Depressão de Beck, Questionário de Saúde Geral, Questionário de Fadiga Física e Cognitiva (CPF-MS) e Medida de Auto-Eficácia específica para Esclerose Múltipla (MSSE)	2 Traduções	Consistência Interna	Construto (convergente) e Critério (concorrente)
Gobitta & Guzzo	2002	Inventário de Auto-Estima (SEI) - Forma A	1 Tradução	Consistência Interna	

Em consonância com o que aponta a literatura internacional, a maioria dos estudos apresenta como justificativa para a adaptação de instrumentos de avaliação a economia de recursos e os ganhos com a possibilidade de comparação de resultados entre amostras de diferentes países. Esta análise comparativa, segundo Hambleton (2005), pode favorecer a compreensão de diferenças e de similaridades existentes entre grupos com características linguísticas ou culturais específicas.

A inexistência de instrumentos de avaliação nacionais com foco em construtos específicos de interesse dos pesquisadores é recorrente e, muitas vezes, não só explica a necessidade da adaptação do instrumento como também justifica a escolha de um método de avaliação estrangeiro. Ainda como justificativa para a adaptação de instrumentos de avaliação, também é citada a não-aplicabilidade dos mesmos à população brasileira, demonstrada por estudos anteriores.

A recomendação do Conselho Federal de Psicologia em relação à escolha de um instrumento a ser utilizado na avaliação psicológica é de que seja levada em consideração a segurança em sua utilização, de modo a garantir a legitimidade e a cientificidade dos dados obtidos (Werlang, Vilemor-Amaral & Nascimento, 2010). As evidências de validade e de fidedignidade, assim como a especificação do construto que o instrumento pretende avaliar e a literatura que fundamenta o desenvolvimento do instrumento, estão entre os requisitos obrigatórios que o Conselho Federal de Psicologia determina para todos os instrumentos de

avaliação psicológica. Esses requisitos acabam funcionando como critérios de escolha dos instrumentos a serem utilizados pelo profissional psicólogo, seja como pesquisador ou em sua prática clínica (CFP, 2003).

Observando o presente levantamento bibliográfico, foi possível perceber que, de forma geral, os autores justificam a escolha do instrumento psicológico a ser adaptado com base em suas características técnicas, relatadas na literatura internacional. Entre elas estão boas qualidades psicométricas (especialmente bons índices de validade e de fidedignidade), a forma de construção do instrumento ou a rapidez e facilidade na aplicação, e a capacidade de avaliar algum construto específico e significativo em determinados grupos clínicos, como, por exemplo, evolução do tratamento de afecções do ombro em atletas, qualidade de vida em mulheres com endometriose. Também são citadas como motivadores de seleção de um instrumento de avaliação psicológica (para estudo de sua adaptação ao contexto brasileiro), a possibilidade de sua aplicação em diferentes níveis de escolaridade e a grande utilização em práticas clínicas no país em que o mesmo foi desenvolvido.

Quando é referido que existem outros instrumentos que medem aquele construto específico, os autores justificam a adaptação com base em qualidades técnicas do material em relação aos demais, como sua aplicabilidade a determinada população ou as características da população (idade, comorbidades, especificidades do diagnóstico). É importante ressaltar, no entanto, que nem sempre se encontram descrições

completas das características do instrumento, como seu referencial teórico de base ou a população a que se aplica, ou mesmo os índices psicométricos observados em estudos anteriores.

Baseando-se nessas justificativas, nos artigos recuperados, os objetivos foram descritos principalmente utilizando as expressões “traduzir e realizar a adaptação transcultural” ou similares (traduzir e adaptar, por exemplo). Os estudos descreveram suas metas como: “adaptar e verificar as qualidades psicométricas” ou “analisar as qualidades psicométricas”. Alguns restringiram o objetivo à realização da “adaptação transcultural”, englobando procedimentos de tradução, verificação da adequação semântica/conceitual e análise das propriedades psicométricas.

Desse modo, o termo “adaptação transcultural” ora é utilizado como um procedimento mais abrangente, seguindo referências como Guillemín, Bombardier e Beaton (1993) e Hambleton (2005), ora pareceu não incluir os procedimentos de tradução. Nestes casos, em que os procedimentos de tradução não são incluídos, a “adaptação transcultural” tende a compreender a avaliação da versão por comitê de juízes, modificação dos itens, verificação da adequação semântica com o grupo-alvo da população (validade de face) e testagem piloto.

Parece haver um consenso em relação à diferenciação entre a simples tradução do instrumento avaliativo e sua adaptação transcultural, na direção do que a literatura internacional aponta e recomenda. Hambleton (2005) ressalta que a adaptação de técnicas de exame psicológico abarca todas as atividades que envolvem a verificação da capacidade do instrumento de medir o mesmo construto em idioma e cultura diferentes do contexto original onde foi desenvolvido. Por sua vez, o processo de traduzir um instrumento de avaliação psicológica é descrito como uma tradução literal do conteúdo avaliativo do instrumento, constituindo-se apenas como uma das etapas do processo de adaptação.

No entanto, na prática profissional brasileira, historicamente, nem sempre os procedimentos de verificação das qualidades psicométricas de um instrumento de avaliação psicológica de origem estrangeira fizeram parte do denominado “processo de adaptação sociocultural” do instrumento em questão. Alguns estudos, embora tenham como objetivo realizar a adaptação de instrumentos, não verificaram sua validade e fidedignidade para o grupo-alvo, como recomendado por Hambleton (2005). Dos 24 trabalhos recuperados, todos realizaram procedimentos de tradução, acompanhados da retrotradução e da

avaliação da pertinência da versão produzida, com ajustes semânticos e conceituais. Deste conjunto, foram 16 os estudos que realizaram análises de validade e 18 avaliaram a precisão dos instrumentos de avaliação adaptados ao Brasil.

As recomendações do Conselho Federal de Psicologia (Anache & Reppold, 2010), amparadas pelas Diretrizes da *Internacional Test Commission* (2003), deixam clara a importância das evidências de validade e precisão dos instrumentos a serem utilizados, com implicações diretas na prática do profissional psicólogo. Alguns estudos recuperados no presente levantamento, como, por exemplo, o de Sardinha e cols. (2010), reconhecem a importância desses procedimentos, apontando-os como indicações para futuras investigações científicas. No entanto, o conjunto de resultados deste levantamento tende a indicar que a prática da adaptação de instrumentos de avaliação psicológica para o contexto brasileiro ainda não tem contemplado a totalidade de recomendações e esforços no sentido de garantir o rigor teórico e metodológico no emprego de instrumentos de avaliação psicológica na prática profissional. Atenta-se para o fato de que esses limites na revisão metodológica dos instrumentos em si podem vir a causar impacto negativo na qualidade do trabalho do psicólogo, atingindo um de seus deveres fundamentais imposto pelo Código de Ética Profissional.

Neste levantamento não foram encontrados trabalhos que se propuseram a estabelecer as normas avaliativas do instrumento em uso para a população brasileira como parte do procedimento de adaptação transcultural. É importante ressaltar que a construção de normas é reconhecida pelo Conselho Federal de Psicologia como requisito mínimo e obrigatório para a adequação do uso de instrumentos de avaliação psicológica. As normas oferecem referenciais de comparação de desempenho e permitem a avaliação correta do respondente em relação à população geral (Pasquali, 2001). Pode-se questionar, inclusive, se um processo de adaptação transcultural está completo sem o estabelecimento de padrões normativos, tendo em vista que a utilização do instrumento a partir de referenciais de outras populações pode não embasar adequada interpretação dos resultados. O uso de padrões normativos de outros grupos populacionais, na verdade, contradiz diretrizes técnicas determinadas no *Standards for Educational and Psychological Testing*, que pressupõem a elaboração de normas desenvolvidas para a população na qual o instrumento está sendo utilizado (Geisinger, 1994). No entanto, há que ressaltar que as publicações examinadas talvez retratem uma etapa ou parcelas dos estudos desenvolvidos no

contexto nacional sobre aqueles instrumentos de avaliação em foco, relativizando inicialmente o argumento crítico, exigindo investigação complementar sobre esses materiais.

Tradução

Do total amostral de 24 estudos, todos realizaram procedimentos relacionados à tradução do instrumento de avaliação original. Em relação aos procedimentos utilizados, a análise aqui realizada considerou o número e a qualificação dos profissionais que participaram de cada etapa do processo (tradução, retrotradução, análises posteriores), além de quais foram os procedimentos adotados para verificar a versão produzida.

A primeira fase dos estudos correspondeu à tradução em si do instrumento de avaliação. A maioria dos estudos (15) descreve sua tradução como “conceitual”, realizada por dois tradutores profissionais e independentes, com língua materna portuguesa. Dentre estes, dois trabalhos também realizaram a tradução literal por dois professores de língua inglesa. Os demais trabalhos utilizaram diversas formas para esse processo, 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 mais um professor de língua inglesa; um tradutor bilíngue; dois pesquisadores da própria equipe do estudo. Um dos trabalhos não mencionou a forma como conduziu esse processo. Além disso, também um dos artigos referiu ter realizado pré-teste após essa primeira tradução de instrumento de avaliação.

Os trabalhos também se preocuparam com a retrotradução (versão do português do Brasil para a língua original), sendo que, em 13 casos, ela foi elaborada por um ou dois brasileiros bilíngues. Em outros oito estudos, a retroversão foi feita por tradutores ingleses (seguindo estudo original), cuja língua materna era o inglês, enquanto um estudo contou com a participação de uma equipe multidisciplinar nessa fase. Novamente, um dos artigos avaliados não mencionou como realizou essa parte do processo de adaptação do instrumento de avaliação.

As etapas técnicas posteriores foram constituídas, basicamente, de avaliações por comitês multiprofissionais, com análises semânticas, comparações de versões preliminares e aplicações de pré-teste em cinco dos trabalhos. Em outros três estudos, foram realizadas comparações das versões, análises de sua equivalência e consenso entre os tradutores. Em outros quatro casos, além disso, foram incluídas aplicações de pré-teste do instrumento. Os

demais estudos se preocuparam em examinar as versões em comitês multiprofissionais, comparando-as e examinando seu conteúdo (análise de face); ou apenas realizaram uma revisão do material, não oferecendo detalhes informativos.

Segundo Beaton e cols. (2002), algumas etapas devem ser seguidas para que ocorra uma tradução e adaptação transcultural válida e eficaz de qualquer instrumento de avaliação ligado à área da saúde. Ainda de acordo com esses autores, essas etapas teriam início com (pelo menos) duas traduções feitas do material original para a língua-alvo, garantindo, dessa forma, que possíveis discrepâncias e ambiguidades pudessem ser identificadas. Quinze estudos preocuparam-se com essa primeira etapa técnica proposta por Beaton e cols. (2002), produzindo duas traduções independentes por tradutores bilíngues que tinham a língua-alvo como língua materna.

Após esse primeiro estágio, ainda segundo Beaton e cols. (2002), uma síntese da tradução deveria ser feita por um terceiro profissional, tendo por base as diferenças eventuais das versões traduzidas. Esse profissional deveria trabalhar de modo imparcial, sem conhecimento prévio dos conteúdos abordados no estudo. Nessa análise, apenas o estudo de Neves, Muniz, Medeiros, Moraes, e Gonçalves (2009) descreveu ter realizado esse estágio técnico de modo completo.

Em conformidade com as orientações de Beaton e cols. (2002), a etapa seguinte consistiria na retrotradução da versão-síntese alcançada. Nesse momento o instrumento seria vertido para o idioma original, totalmente às cegas, visando testar se os itens traduzidos refletiriam o conteúdo da versão original, verificando assim sua validade. Essa etapa também deveria ser realizada por dois tradutores com língua materna igual a da origem do instrumento de avaliação em foco. Dentre os estudos analisados neste atual levantamento, 22 se propuseram a realizar a retrotradução, porém somente 13 deles se preocuparam em incluir um profissional que realizasse essa etapa totalmente às cegas, sendo que em oito estudos os tradutores tinham a língua materna equivalente à original do instrumento.

Para que se possa alcançar uma equivalência nas traduções e retrotraduções, faz-se necessário, conforme Beaton e cols. (2002), que se forme um comitê de avaliadores, composto tanto por membros da equipe de pesquisa quanto por tradutores que participaram da primeira e da segunda fase. Apenas seis trabalhos se preocuparam em formar um comitê de especialistas (multidisciplinar e que também abarcasse os profissionais responsáveis pelas traduções), enquanto

12 dos trabalhos referiram ter recorrido a um comitê multidisciplinar para examinar a qualidade das versões do instrumento trabalhadas até esse momento.

O trabalho desse comitê de avaliação deveria se pautar por quatro tipos de equivalências necessárias e importantes: semântica (que analisa se as palavras têm o mesmo significado), idiomática (que avalia eventuais coloquialismos), cultural (que avalia se as experiências diárias podem ser traduzidas da mesma forma em diferentes culturas) e conceitual (que examina se o construto estudado permaneceu com o significado original). Dentre os trabalhos analisados neste levantamento, 12 citaram ter realizado estudo de equivalência entre as versões do instrumento, sem, contudo, especificar em qual nível se pautaram. Outros nove trabalhos avaliaram a equivalência semântica, três a idiomática, três a conceitual e apenas um citou ter examinado a equivalência cultural. Além disso, dois estudos falaram em validade de conteúdo e de face a partir dessa etapa de tradução e de adaptação do instrumento, sem maiores informações.

A última fase do processo de tradução e de adaptação de um instrumento avaliativo, segundo Beaton e cols. (2002), deveria ser um pré-teste ou um teste piloto com a nova versão do material, em um grupo de 30 a 40 participantes. Estes, por sua vez, deveriam ser entrevistados logo após responder ao instrumento, com o objetivo de informar sobre sua apreciação no tocante à qualidade do item e do material em si, oferecendo ricos dados sobre a aplicabilidade do mesmo. Essa fase final foi descrita por 12 dos estudos apresentados. Ainda assim, os estudos que utilizaram o pré-teste do material variaram sua amostra entre nove e 24 participantes. Apenas um estudo apresentou amostra compatível com o que é indicado por Beaton e cols. (2002), composta por 40 participantes (Pereira, Costa, Geocze, Borim & Ciconelli, 2007).

No levantamento bibliográfico realizado não foram encontradas informações específicas sobre a utilização de bilíngues para testar os instrumentos. Sireci (2010) propõe uma metodologia de testagem que utiliza bilíngues, buscando se aproximar cada vez mais da equivalência entre as duas versões do instrumento de avaliação (a original e a adaptada). O estudo conduzido por Ribas, Moura e Hutz (2004), no entanto, é um exemplo de trabalho brasileiro que utilizou bilíngues na adaptação de um instrumento de avaliação psicológica e que, entretanto, não foi captado no presente trabalho. Trabalhos dessa natureza podem não ter sido recuperados por conta das palavras-chave utilizadas na presente busca bibliográfica. Além disso, os unitermos empregados pelos pesquisadores em suas publicações científicas podem limitar a visibilidade

desse tipo de cuidado técnico, o que merece a devida reflexão pelos profissionais da área.

Fidedignidade

De acordo com *The Standards for Educational and Psychological Testing* (AERA, APA & NCME, 1999), todo aquele que se propõe a desenvolver um instrumento de avaliação tem a responsabilidade de investigar a fidedignidade do mesmo. Considerando que essa propriedade psicométrica diz respeito aos escores obtidos pelo instrumento, o que é suscetível às mudanças de contexto e dos indivíduos avaliados, torna-se relevante que também seja investigada a precisão durante os processos de adaptação transcultural do instrumento avaliativo.

Considerando os estudos identificados neste atual levantamento bibliográfico, conforme dados sistematizados na Tabela 1, pôde-se notar que 18 deles realizaram algum procedimento relativo à análise da fidedignidade na adaptação de seus instrumentos para o contexto brasileiro. Dentre estes estudos, nove realizaram somente um método de avaliação da fidedignidade e oito trabalhos fizeram uso de dois procedimentos técnicos nessa direção. Somente o estudo de Pereira, Costa, Geocze, Borim, Ciconelli e Camacho-Lobato (2007) recorreu a três métodos distintos para avaliar a fidedignidade em seus instrumentos de avaliação (teste-reteste, inter-observador e consistência interna).

Para analisar a precisão em um instrumento de avaliação, diferentes métodos podem ser utilizados a fim de estimar possíveis fontes de erros, expressos pelos coeficientes de fidedignidade (Urbina, 2007). Os delineamentos mais empregados nessa direção, segundo Nunes e Primi (2010), são: equivalência de formas paralelas, consistência interna, estabilidade temporal (também conhecido como método de teste-reteste) e precisão de avaliadores.

Dentre esses diferentes procedimentos, o método de teste-reteste e o de consistência interna foram os mais utilizados nas pesquisas aqui analisadas. Destes, 11 fizeram uso do teste-reteste. Segundo Anastasi (1977), esse é o método que mais evidencia a precisão de um instrumento de exame psicológico, implicando uma segunda aplicação do mesmo, em momento posterior, guardando-se tempo de dias a semanas entre as avaliações. Para a autora, “a precisão do reteste mostra até que ponto os resultados em um teste podem ser generalizados para outras ocasiões” (p. 93). Vale destacar que o procedimento teste-reteste foi denominado em vários dos trabalhos analisados como “reprodutibilidade intra-avaliador/intraobservador”.

Urbina (2007) e Anastasi e Urbina (2000) ressaltam a importância da especificação do intervalo de tempo transcorrido entre as aplicações, quando este procedimento de teste-reteste é usado. Segundo as autoras, não há um período de intervalo fixo que possa ser preconizado, no entanto, deve-se atentar às possíveis consequências inerentes à escolha do intervalo entre teste e reteste. Essa escolha deve-se basear nas características da população investigada (estágio de desenvolvimento, presença de patologias, por exemplo), assim como nos aspectos do instrumento utilizado (objetivos do teste, comportamento que pretende medir ou prever, entre outras características).

Todos os estudos aqui examinados e que utilizaram esse procedimento na análise da fidedignidade especificaram o tempo do intervalo entre teste e reteste variando de uma hora a trinta dias. De acordo com Urbina (2007), intervalos muito curtos podem ser suscetíveis a efeitos de aprendizagem (associados à memória), podendo afetar os escores na repetição da medida. Alguns dos estudos que utilizaram intervalo bastante curto na reaplicação referem-se a instrumentos que avaliam presença ou intensidade de sintomas, ou construtos relacionados à qualidade de vida. Segundo os pesquisadores, o tema em foco nesses estudos voltava-se a condições temporárias ou vivências situacionais. Nesse sentido, o curto intervalo de tempo entre teste e reteste poderia estar justificado, uma vez que abordaram construtos ou estados mais propensos a flutuações do que aqueles relacionados, por exemplo, a habilidades e aptidões (Urbina, 2007).

Dos trabalhos analisados neste levantamento, 11 recorreram também ao método da consistência interna. De acordo com a definição de Urbina (2007) “medidas de consistência interna são procedimentos estatísticos que procuram avaliar a extensão da inconsistência entre os itens de um teste” (p. 133). Esse método avalia a precisão baseando-se nos dados de apenas uma aplicação e uma única forma do instrumento.

O método de consistência interna é um dos mais conhecidos e utilizados, principalmente na avaliação da fidedignidade de escalas e instrumentos com um número variado de itens. O fato de a maioria dos estudos analisados tratarem da adaptação de escalas, questionários e inventários (que são instrumentos compostos por vários itens), pode justificar o amplo uso desse método dentre eles.

Por fim, cabe destacar que cinco dos estudos aqui analisados apontaram a utilização da reprodutibilidade inter-avaliador/interobservador ou “fidedignidade do avaliador” como método para

análise da fidedignidade. Ambas as nomenclaturas foram utilizadas pelos estudos para denominação do método em questão. De acordo com Anastasi e Urbina (2000), a fidedignidade do avaliador é realizada a partir da avaliação do mesmo instrumento por dois ou mais profissionais, independentemente. Assim, os resultados obtidos são comparados e o coeficiente de correlação resultante constitui-se na medida de fidedignidade do avaliador. Esse método é geralmente utilizado quando há indícios de que fatores subjetivos possam influenciar na avaliação de um instrumento (Urbina, 2007). Dessa maneira, a fidedignidade do avaliador permite indicar a coerência entre avaliações de determinada atividade, realizadas por profissionais diferentes (Nunes & Primi, 2010).

Dentre as pesquisas analisadas neste atual levantamento bibliográfico, nenhuma examinou a equivalência de formas paralelas como estratégia de análise da fidedignidade do instrumento avaliativo. Ainda em relação aos possíveis métodos de análise da precisão dos instrumentos de avaliação, um dos estudos analisados citou que foi utilizado o coeficiente de correlação de duas metades, porém sem esclarecer, com detalhes, os procedimentos empregados.

A análise da fidedignidade em instrumentos de avaliação é de suma importância. Como enfatiza Thompson (2003), ela deve ser testada em todos os estudos, principalmente pelo fato de que a fidedignidade é uma questão de precisão dos escores e não do instrumento em si, havendo grande chance de que sofra variações entre grupos com características distintas. Além disso, o autor acrescenta que escores pouco precisos comprometem resultados de estudos e avaliações, uma vez que podem inviabilizar análises de validade no mesmo instrumento e fornecer evidências pouco confiáveis, tanto estatisticamente, quanto para a prática.

Validade

A validade, em um teste psicológico, refere-se àquilo que é medido por ele e o quão eficazmente ele o faz (Anastasi & Urbina, 2000). Nesse sentido, deve-se levar em conta o que o teste se propõe a medir, além de considerar os conceitos implicados nessa medida (Fachel & Camey, 2000). Messick (1995) faz uma ressalva de que a validade não diz respeito a uma qualidade do instrumento de avaliação psicológica em si, mas de seus escores e das inferências realizadas a partir de suas interpretações.

Ao se retomarem, mais uma vez, as evidências sistematizadas na Tabela 1, sobre os estudos avaliados na presente revisão, nota-se que 16 deles realizaram algum procedimento relativo à análise da validade na

adaptação de seus instrumentos de avaliação para o contexto brasileiro. Dez desses estudos analisaram somente um tipo de evidência de validade, mas outros cinco trabalhos recorreram a dois processos de validação. Somente dois estudos realizaram mais de três procedimentos na análise da validade de seu instrumento adaptado para o Brasil.

A edição de 1985 do manual *The Standards for Educational and Psychological Testing* (AERA, APA & NCME, 1985) agrupou as evidências de validade em três categorias principais: validade relacionada a conteúdo, validade relacionada a critério e validade relacionada a construto. Já na edição de 1999 (AERA, APA & NCME, 1999) houve a preocupação de salientar que o conceito de validade é unitário, ou seja, há diferentes processos de validação que permitem demonstrar diversos aspectos desse conceito, mas não tipos diferentes de validade. Essa nova classificação engloba os tipos de evidência de validade descritos na edição de 1985 (AERA, APA & NCME, 1985), porém mediante nova nomenclatura e novo agrupamento. Vale ressaltar que o crivo avaliativo utilizado pelo Sistema de Avaliação dos Testes Psicológicos (SATEPSI – Conselho Federal de Psicologia) para a verificação da qualidade dos testes psicológicos baseia-se nessa nova classificação de validade (Nunes & Primi, 2010).

Observou-se, no presente levantamento bibliográfico, que a maioria dos estudos analisados utilizou a nomenclatura de classificação dos tipos de evidência de validade descrita na edição de 1985 do manual *The Standards for Educational and Psychological Testing* (AERA, APA & NCME, 1985). Portanto, optou-se por adotá-la nas análises descritas a seguir.

Dentre as diferentes categorias de evidências de validade, a validade relacionada a construto, do tipo convergente, esteve presente em 11 dos estudos analisados. Segundo Fachel e Camey (2000), a validade relacionada a construto verifica o alcance do instrumento em relação ao domínio de investigação, ou seja, o quanto ele mede o traço ou conceito teórico que se propõe a medir. Para isso, todos os estudos analisados que fizeram uso desse tipo de evidência de validade realizaram-na pela comparação do instrumento em adaptação com outros tipos de medidas já validadas e que examinam característica similar, caracterizando a validade convergente (Nunes & Primi, 2010).

Segundo Urbina (2007), os estudos que realizam correlações entre resultados de diferentes instrumentos de avaliação são bastante frequentes, uma vez que podem se pautar em amostras de conveniência, facilitando o processo de coleta de dados. Isso pode

justificar o fato de a maioria dos estudos analisados terem utilizado esse método para avaliar a validade.

Posteriormente, a validade relacionada a conteúdo, assim classificada pelos estudos analisados, foi efetivada em sete dos trabalhos aqui examinados, recorrendo a análises por comitês de especialistas. Como ressaltam Cassep-Borges e cols. (2010), a validação de conteúdo compreende “estudos que investigam a clareza, a representatividade e a relevância dos itens” (p. 511) de um instrumento de avaliação. Normalmente, esse tipo de validação não exige cálculo estatístico adicional, resultando, assim, do julgamento de pessoas com amplo conhecimento na área do conceito teórico investigado, denominadas juízes (Fachel & Camey, 2000).

A validade relacionada a critério, por sua vez, esteve presente em cinco dos estudos examinados no presente levantamento. Segundo Nunes e Primi (2010), esse processo de validade caracteriza-se como um dos métodos mais importantes na avaliação da aplicabilidade do instrumento, no que diz respeito, sobretudo, a diagnósticos. A validade relacionada a critério permite diferenciar grupos com características distintas, constituindo-se em um procedimento bastante relevante para a demonstração da utilidade do instrumento de avaliação. Esse tipo de validação pode ser do tipo preditivo ou concorrente (Fachel & Camey, 2000). Todos os estudos aqui examinados que apresentaram validade relacionada a critério utilizaram o delineamento do tipo concorrente, ou seja, a verificação do diagnóstico ou da característica clínica particular de cada grupo avaliado foi realizada simultaneamente à administração do instrumento em adaptação, ou em períodos próximos (Nunes & Primi, 2010).

O predomínio da validade concorrente sobre a preditiva, dentre estudos analisados, pode ser explicado por critérios práticos. Anastasi (1977) salienta que estudos de validade concorrente são mais utilizados e, algumas vezes, até substituem a validade preditiva, já que esta se torna frequentemente impraticável pela questão do tempo entre a administração do instrumento e o critério preditivo, além da difícil obtenção de amostra pré-selecionada adequada para essa finalidade.

Por fim, a validade de face, como foi classificada nos estudos analisados, foi a menos frequente nesta revisão, observada em três estudos. Também conhecida como “validade aparente”, esta avalia, de modo superficial, se o instrumento parece ou não válido à população à qual será administrado (Anastasi & Urbina, 2000). Refere-se, principalmente, à forma de apresentação de seu conteúdo, à linguagem

utilizada e avalia se os termos usados são apropriados ao grupo específico ao qual o instrumento será aplicado (Fachel & Camey, 2000). Da mesma forma que a validade de conteúdo propriamente dita, a validade de face não requer análise estatística, sendo, de modo geral, determinada por especialistas da área (Fachel & Camey, 2000). Vale ressaltar que os estudos na presente revisão consideraram a avaliação de especialistas, assim como a de indivíduos componentes da população-alvo à qual o instrumento seria aplicado. Dessa forma, pretendia-se garantir a adequação dos itens do instrumento baseada na compreensão e no entendimento dos mesmos por pessoas com as mesmas características do público-alvo.

Além disso, todos esses estudos fizeram uso, concomitantemente à validade de face, de outro procedimento de validação (de conteúdo propriamente dito ou de construto), o que concorda com as recomendações da *International Test Commission* (2003). Segundo suas diretrizes, dever-se-ia evitar o julgamento de um instrumento de avaliação unicamente com base em sua validade aparente. Nesse mesmo sentido, Anastasi e Urbina (2000) apontam que esse tipo de validade não pode substituir aquela determinada de forma objetiva, sendo necessário, além da validade de face, avaliar a validade do instrumento em sua forma final.

Embora cerca de 60% da totalidade dos estudos analisados nesta revisão tenham realizado algum tipo de análise de validade na adaptação de seus instrumentos, esse número pode ser considerado pequeno diante da relevância das evidências de validade nos processos de adaptação de instrumentos de avaliação. Segundo o manual *The Standards for Educational and Psychological Testing* (AERA, APA & NCME, 1999), a validade constitui-se em análise fundamental no desenvolvimento de instrumentos de avaliação, sendo responsabilidade de quem cria a técnica, daquele que o adapta para outra cultura, assim como do profissional que a utiliza, uma vez que os dados de validade fornecem evidências e fundamentação teórica para seu uso pretendido.

Nesse sentido, considerando a validade como conceito único e seus diferentes processos como vários aspectos desse mesmo conceito (AERA, APA & NCME, 1999), torna-se relevante, dentre os estudos de adaptação cultural de instrumentos de avaliação psicológica, colecionar o maior número possível dessas evidências empíricas e técnicas, garantindo a esses instrumentos graus satisfatórios da validade de seus escores e das inferências realizadas a partir de suas interpretações. Para tanto, deve-se sempre ter em foco

os objetivos para os quais o instrumento de avaliação foi elaborado e será utilizado.

Considerações finais

O levantamento bibliográfico realizado fornece um panorama geral de como têm sido efetuadas as adaptações transculturais de instrumentos de avaliação psicológica no Brasil. Podem ser apreendidas tendências que representam a prática em relação à adaptação de instrumentos de avaliação não somente no âmbito da ciência psicológica, mas também em outras áreas do conhecimento. No entanto, como já foi mencionado, alguns estudos que também realizaram adaptações podem não ter sido recuperados na presente revisão, especialmente pelas palavras-chave utilizadas na busca bibliográfica, dada a falta de uniformidade entre os termos que se referem a esta metodologia. É importante considerar também que, por se tratar de uma revisão abrangente, a análise realizada pode não ter considerado peculiaridades de cada estudo que mereceriam destaque em análises mais específicas, ou mais pontuais, ultrapassando as possibilidades de uma abordagem panorâmica sobre o tema, como aqui adotada.

Foi possível observar que a maioria dos estudos recuperados na presente revisão seguem diretrizes internacionais e nacionais de adaptação de instrumentos. Entre os mais citados estão Bunchaft e Cavas (2002), Guillemín, Bombardier e Beaton (1993), Vallerand (1989) e aquelas da Organização Mundial da Saúde (Beaton & cols., 2002). Essas diretrizes técnicas norteiam a utilização dos instrumentos de avaliação psicológica, bem como a qualidade dos processos de adaptações em todo o mundo, contribuindo para a validade de pesquisas transculturais (Hambleton, 2005).

Do ponto de vista dos controles metodológicos, alguns estudos se destacam em relação a diferentes critérios utilizados em seus procedimentos técnicos. Entre eles estão estudos cuja escolha criteriosa e a qualificação dos juízes e profissionais que participaram dos procedimentos de adaptação dos instrumentos valorizam e fortalecem o rigor metodológico do trabalho, como, por exemplo, o trabalho desenvolvido por Mendes, Travassos, Martins e Marques (2008). Há estudos que esclarecem quais foram os critérios de inclusão e exclusão dos participantes e apresentam a caracterização da amostra (por exemplo, de Mengarda, Passos, Picon, Costa & Picon, 2008) e estudos que deixam claros os critérios levados em consideração para escolha do instrumento a ser adaptado, como o realizado por Levitan, Nascimento, Freire, Mezzasalma e Nardi (2008).

Caspe-Borges e cols. (2010) ressaltam que, levando em consideração aspectos éticos, o primeiro passo para realizar a adaptação de um instrumento é solicitar a autorização formal do autor para sua utilização. Os trabalhos de Sant'Anna, Blascovi-Assis e Magalhães (2008) e Camargo e Contel (2004) contaram não somente com a autorização para o trabalho com o instrumento, mas também descrevem contatos posteriores com os autores (envio da retrotradução, participação na versão final). Além de demonstrar rigor e cuidados em relação à versão original produzida, o contato com os autores é fonte de possibilidades de trocas científicas enriquecedoras entre pesquisadores.

É preciso destacar também a importância da descrição detalhada do método. Como salientam Van de Vijver e Poortinga (2010), a descrição da ordem dos instrumentos aplicados ou da forma de aplicação de cada um deles, bem como das análises dos dados realizadas são fundamentais para garantir a replicabilidade e a confiabilidade do estudo, além de evitar vieses metodológicos. A descrição dos instrumentos utilizados nos procedimentos de validade convergente ou divergente, por exemplo, é fundamental nos procedimentos, ou os índices apresentados como resultados não podem ser considerados suficientes e significativos pelo pesquisador que examina o tema.

No que diz respeito à análise da fidedignidade e da validade, essas são qualidades atribuídas aos escores e à interpretação de seus resultados, não sendo propriedades do instrumento em si. Isso implica que, em diferentes grupos, com características e contextos distintos, a precisão, assim como a validade, deveriam ser verificadas, sendo imprescindíveis nos processos de adaptação transcultural de instrumentos de avaliação.

Como enfatiza Thompson (2003), a análise da fidedignidade dos escores é tão significativa que sua imprecisão pode comprometer, ou até mesmo inviabilizar os estudos de validade no mesmo instrumento. Escores pouco precisos dificilmente garantem a avaliação, de forma adequada, de construtos específicos. Dessa forma, pode-se considerar que a precisão e a validade são análises complementares, e que ambas devem estar presentes nos processos de adaptação transcultural de instrumentos de avaliação (CFP, 2010).

Diante de algumas divergências referentes à taxonomia das propriedades psicométricas (fidedignidade e validade), seria de grande utilidade técnica e clareza científica uma padronização desses termos, tendo em vista facilitar o uso e a identificação dos mesmos dentro do contexto brasileiro. Dentro desta perspectiva pode-se compreender o esforço e o

empenho de vários pesquisadores da área, resultando na publicação do CFP (2010) relativa à avaliação psicológica no Brasil.

As análises realizadas na atual revisão bibliográfica pretendem fornecer parâmetros de ações para pesquisadores e para profissionais que realizam avaliações psicológicas, tendo sido efetivadas sob a luz de importantes publicações científicas que norteiam a prática da adaptação de instrumentos de avaliação psicológica no Brasil. Dessa maneira, espera-se que estudos futuros que se proponham a adaptar instrumentos de avaliação para o contexto brasileiro, tenham a preocupação de incluir em seus procedimentos, além da tradução, as análises da precisão e da validade nessas medidas, com o intuito de disponibilizarem instrumentos suficientemente pautados em termos científicos e adequados para o uso, proporcionando expansão dos estudos e qualificação profissional nessa área.

Referências

- AERA, APA & NCME – American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1985). *The standards for educational and psychological testing*. Washington: American Educational Research Association.
- AERA, APA & NCME – American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1999). *The standards for educational and psychological testing*. Washington: American Educational Research Association.
- Anache, A. A. & Reppold, C. T. (2010) Avaliação psicológica: implicações úteis. Em CFP, *Avaliação Psicológica: diretrizes na regulamentação da profissão*. (pp. 57-85). Brasília: CFP.
- Anastasi, A. & Urbina, S. (2000). *Testagem Psicológica*. (M. A. V. Veronese, Trad.). Porto Alegre: Artes Médicas. (Original publicado em 1997).
- Bandeira, M., Calzavara, M. G. P. & Varella, A. A. B. (2005). Escala de sobrecarga dos familiares de pacientes psiquiátricos: adaptação transcultural para o Brasil (FBIS-BR). *Jornal Brasileiro de Psiquiatria*, 54(3), 206-214.
- Beaton, D., Bombardier, C., Guillemin, F. & Ferraz, B. M. (2002). *Recommendations for the cross-cultural adaptation of health status measures*. American Academy of Orthopaedic Surgeons. Institute for Work & Health.

- Bunchaft, G. & Cavas, C. S. T. (2002). Sob medida: um guia sobre a elaboração de medidas do comportamento e suas aplicações. São Paulo, Vetor.
- Camargo, I. B. & Contel, J. O. B. (2004). Tradução e adaptação de questionários norte-americanos para a avaliação de habilidades e conhecimentos na prática psiquiátrica brasileira. *Revista de Psiquiatria do Rio Grande do Sul*, 26(3), 288-299.
- Cassep-Borges, V., Balbinotti, M. A. A. & Teodoro, M. L. M. (2010). Tradução e validação de conteúdo: uma proposta para a adaptação de instrumentos. Em L. Pasquali e cols., *Instrumentação psicológica: fundamentos e práticas*. (pp. 506-520). Porto-Alegre: Artmed.
- Ciconelli, R. M., Soárez, P. C., Kowalski, C. C. G. & Ferraz, M. B. (2006). The Brazilian Portuguese version of the Work Productivity and Activity Impairment: General Health (WPAI-GH) Questionnaire. *São Paulo Medical Journal*, 124(6), 325-33.
- CFP – Conselho Federal de Psicologia. (2003). *Resolução CFP 02/2003: define e regulamenta o uso, a elaboração e a comercialização de testes psicológicos*. Brasília: CFP.
- Echevarría-Guanilo, M. E., Rossi, L. A. Dantas, Rosana A. S. & Santos, C. B. (2006). Adaptação transcultural da “Burns Specific Pain Anxiety Scale – BSPAS” para ser aplicada em pacientes queimados brasileiros. *Revista Latino-Americana de Enfermagem*, 14(4), 526-533.
- Fachel, J. M. G. & Camey, S. (2000). Avaliação psicométrica: a qualidade das medidas e o entendimento dos dados. Em J. A. Cunha (Org.). *Psicodiagnóstico – V* (pp. 158-176). Porto Alegre: Artmed.
- Gambaro, R. C., Santos, F. C., Thé, K. B., Castro, L. A. & Cendoroglo, M. S. (2009). Avaliação de dor no idoso: proposta de adaptação do Geriatric Pain Measure para a língua portuguesa. *Revista Brasileira de Medicina*, 66(3), 62-65.
- Geisinger, K. F. (1994). Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6(4), 304-312.
- Giusti, E. & Befi-Lopes, D. M. (2008). Tradução e adaptação transcultural de instrumentos estrangeiros para o Português Brasileiro (PB)/ translation and cross-cultural adaptation of instruments to the Brazilian Portuguese language. *Pro Fono*, 20(3), 207-210.
- Gobitta, M. & Guzzo, R. S. L. (2002). Estudo inicial do Inventário de Auto-Estima (SEI): Forma A/ initial study of the Self-Esteem Inventory (SEI): Form A. *Psicologia: Reflexão e Crítica*, 15(1), 143-150.
- Goldfeld, P. R. M., Wiethaeuper, D., Terra, L., Baumgardt, R., Lauermann, M., Mardini, V., Abuchaim, C., Sordi, A., Soares, L. & Ceitlin, L. H. F. (2007). Adaptação transcultural do Inventory of Countertransference Behavior (ICB) para o português brasileiro. *Revista de Psiquiatria do Rio Grande do Sul*, 29(1), 56-62.
- Gonçalves, A. M. S. & Pillon, S. C. (2009). Adaptação transcultural e avaliação da consistência interna da versão em português da Spirituality Self Rating Scale (SSRS). *Revista de Psiquiatria Clínica*, 36(1), 10-15.
- Guillemin, F. Bombadier, C. & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology*, 46(12), 1417-1432.
- Haase, V. G., Lacerda, S. S., Lima, E. P., Corrêa, T. D., Brito, D. C. S. & Lana-Peixoto, M. A. (2004). Avaliação do funcionamento psicossocial na esclerose múltipla: características psicométricas de quatro medidas de auto-relato. *Arquivos de Neuropsiquiatria*, 62(2A), 282-291.
- Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting test into multiple languages and cultures. Em R. K. Hambleton, P. F. Merenda & C. D. Spielberger. *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). Londres: LEA.
- ITC – International Test Commission (2003). *Diretrizes para o uso de testes* (Instituto Brasileiro de Avaliação Psicológica, Trad.). Campinas. (Original publicado em 2000).
- Leme, L., Saccol, M., Barbosa, G., Eijnisman, B., Faloppa, F. & Cohen, M. (2010). Validação, reprodutibilidade, tradução e adaptação cultural da escala Athletic Shoulder Outcome Rating Scale para a língua portuguesa. *Revista Brasileira de Medicina*, 67(supl.3), 29-38.
- Levitan, M. N., Nascimento, I., Freire, R. C., Mezzasalma, M. A. & Nardi, A. E. (2008). Equivalência semântica da versão brasileira da Social Avoidance and Distress Scale (SADS).

- Revista de Psiquiatria do Rio Grande do Sul*, 30(1), 49-58.
- Medeiros, M. E. & Guerra, R. O. (2009). Tradução, adaptação cultural e análise das propriedades psicométricas do Activities of Daily Living Questionnaire (ADLQ) para avaliação funcional de pacientes com a doença de Alzheimer. *Revista Brasileira Fisioterapia*, 13(3), 257-266.
- Mendes, W., Travassos, C., Martins, M. & Marques, P. M. (2008). Adaptação dos instrumentos de avaliação de eventos adversos para uso em hospitais brasileiros/ Adjustment of adverse events assessment forms for use in Brazilian hospitals. *Revista Brasileira de Epidemiologia*, 11(1), 55-66.
- Mengarda, C. V., Passos, E. P., Picon, P., Costa, A. F. & Picon, P. D. (2008). Validação de versão para o português de questionário sobre qualidade de vida para mulher com endometriose (Endometriosis Health Profile Questionnaire – EHP-30). *Revista Brasileira de Ginecologia e Obstetrícia*, 30(8), 384-392.
- Messick, S. (1995). Validity of psychological assesment: validation of inferences from persons' response and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.
- Moriguchi, C. S., Alem, M. E. R., Veldhoven, M. & Coury, H. J.C. G. (2010). Cultural adaptation and psychometric properties of Brazilian Need for Recovery Scale/ Adaptação cultural e parâmetros psicométricos da versão brasileira da "Need for Recovery Scale". *Revista de Saúde Publica*, 44(1), 131-13.
- Neves, F. S., Caldas, C. A. M., Medeiros, D. M., Moraes, J. C. B. & Gonçalves, C. R. (2009). Adaptação transcultural da versão simplificada (s) do Behçet's Disease Current Activity Form (BDCAF) e comparação do desempenho das versões brasileiras dos dois instrumentos de avaliação da atividade da Doença de Behçet: BR-BDCAF e BR-BDCAF(s). *Revista Brasileira de Reumatologia*, 49(1), 20-31.
- Novato, T. S., Grossi, S. A. A. & Kimura, M. (2007). Instrumento de qualidade de vida para jovens com diabetes (IQVJD). *Revista Gaúcha de Enfermagem*, 28(4), 512-519.
- Pasquali, L. (2001). *Técnicas de exame psicológico volume i: fundamentos das técnicas psicológicas*. São Paulo: Casa do Psicólogo.
- Pereira, G. I. N., Costa, C. D. S., Geocze, L., Borim, A. A., Ciconelli, R. M. & Camacho-Lobato, L. (2007). Tradução e validação para a língua portuguesa (Brasil) de instrumentos específicos para avaliação de qualidade de vida na doença do refluxo gastroesofágico. *Arquivos de Gastroenterologia*, 44(2), 168-177.
- Ribas Jr, R. C., Moura, M. L. S & Hutz, C. S. (2004). Adaptação brasileira da Escala de Desejabilidade Social de Marlowe Crowne. *Avaliação Psicológica*, 3(2), pp. 83-92.
- Saliba, V. A., Chaves Júnior, I. P., Faria, C. D. C. M. & Teixeira-Salmela, L. F. (2008). Propriedades psicométricas da *motor activity log*: uma revisão sistemática da literatura. *Fisioterapia em Movimento*, 21(3), 59-67.
- Sant'Anna, M. M. M., Blascovi-Assis, S. M. & Magalhães, L. C. (2008). Adaptação transcultural dos protocolos de avaliação do modelo lúdico. *Revista de Terapia Ocupacional*, 19(1), 34-47.
- Sardinha, A., Nardi, A. E. & Eifert, G. H. (2008). Tradução e adaptação transcultural da versão brasileira do Questionário de Ansiedade Cardíaca. *Revista de Psiquiatria do Rio Grande do Sul*, 30(2), 139-149.
- Sardinha, A., Levitan, M. N., Lopes, F. L., Perna, G., Esquivel, G., Griez, E. J. & Nardi, A. E. (2010). Tradução e adaptação transcultural do Questionário de Atividade Física Habitual. *Revista de Psiquiatria Clínica*, 37(1), 16-22.
- Sireci, S. G. (2010) Using bilinguals to evaluate the comparability of different language versions of a test. Em R. K. Hambleton, P. F. Merenda & C. D. Spielberger. (Eds.). *Adapting educational and psychological tests for cross-cultural assessment* (pp. 117-137). Londres: LEA.
- Thompson, B. (2003). *Score reliability: contemporary thinking on reliability issues*. Thousand Oaks, CA: Sage.
- Torres, H. C., Hortale, V. A. & Schall, V. T. (2005). Validação dos questionários de conhecimento (DKN-A) e atitude (ATT-19) de diabetes mellitus. *Revista de Saúde Publica*, 39(6), 906-911.
- Torriani, D. D., Teixeira, A. M., Pinheiro, R., Goettems, M. L. & Bonow, M. L. M. (2008). Adaptação transcultural de instrumentos para mensurar ansiedade e comportamento em clínica odontológica infantil. *Arquivos do Centro de Estudos do Curso de Odontologia*, 44(4), 17-23.

- Urbina, S. (2007). *Fundamentos de testagem psicológica* (C. Dornelles, Trad.). Porto Alegre: Artmed. (Original publicado em 2004).
- Vallerand, R. J. (1989). Vers une méthodologie de validation transculturelle de questionnaires psychologiques: implications pour la recherche en langue française. *Psychologie Canadienne*, 30, 662-680.
- Van de Vijver, F. J. R. & Poortinga, Y. H. (2010). Conceptual and methodological issues in adapting tests. Em R. K. Hambleton, P. F. Merenda & C. D. Spielberger. (Eds.). *Adapting educational and psychological tests for cross-cultural assessment*. (pp. 39-63). Londres: LEA.
- Werlang, B. S. G., Villemor-Amaral, A. E. & Nascimento, R. S. G. F. (2010). Avaliação psicológica, testes e possibilidades de uso. Em Conselho Federal de Psicologia. (Org.). *Avaliação psicológica: diretrizes na regulamentação da profissão* (pp. 87-100). Brasília: Conselho Federal de Psicologia.
- Yusuf, S. A. I., Jorge, J. M. N., Habr-Gama, A., Kiss, D. R. & Rodrigues, J. G. (2004). Avaliação da qualidade de vida na incontinência anal: validação do questionário FIQL (Fecal Incontinence Quality of Life). *Arquivos de Gastroenterologia*, 41(3), 202-208.

Recebido em 27/04/2011

Reformulado em 12/10/2011

Aprovado em 01/11/2011

Sobre as autoras:

Alana Batistuta Manzi-Oliveira é psicóloga (FFCLRP-USP), mestranda (FAPESP) em Psicologia pela Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto-USP e pesquisadora do Grupo de Estudos e Pesquisas em Desenvolvimento e Intervenção Psicossocial (GEPDIP). Atua na área de Avaliação Psicológica de Adolescentes em conflito com a Lei (adaptação transcultural de instrumentos de avaliação de personalidade).

Fernanda Belinassi Balarini é psicóloga (FFCLRP-USP), mestranda (FAPESP) em Neurologia, especialidade Neuropsicologia, pela Faculdade de Medicina de Ribeirão Preto – USP. Pesquisadora do Departamento de Neurociências e Ciências do Comportamento, atua na área de Avaliação Neuropsicológica em Idosos (adaptação transcultural de instrumentos de avaliação neuropsicológica).

Letícia Aparecida da Silva Marques é psicóloga (FFCLRP-USP), especialista técnica em Psicologia (FUNDHERP), mestranda (FAPESP) em Psicologia pela Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto-USP. Pesquisadora do Núcleo de Ensino e Pesquisa em Psicologia da Saúde. Atua na área de psicologia hospitalar (avaliação de qualidade de vida).

Sonia Regina Pasian é psicóloga (FFCLRP-USP), professora associada do Departamento de Psicologia da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto-USP. É, ainda, coordenadora do Centro de Pesquisas em Psicodiagnóstico – FFCLRP-USP, membro do Conselho Consultivo da Associação Brasileira de Rorschach e Métodos Projetivos (ASBRo) e editora associada da revista *Avaliação Psicológica* (IBAP).

