

Fidelidade e validade na construção e adequação de instrumentos de medida

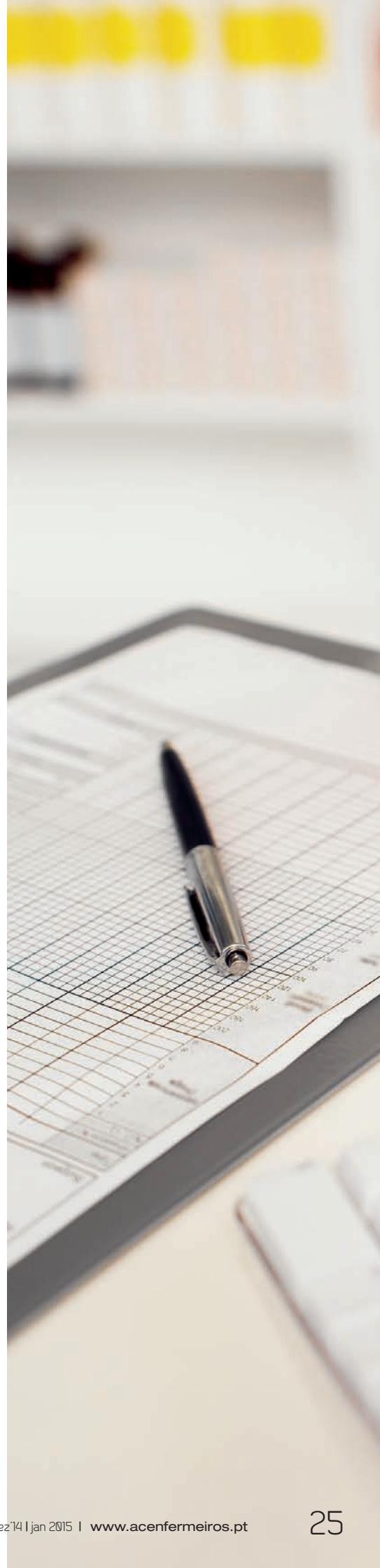
Fidelity and validity in construction and adequacy of measurement instruments

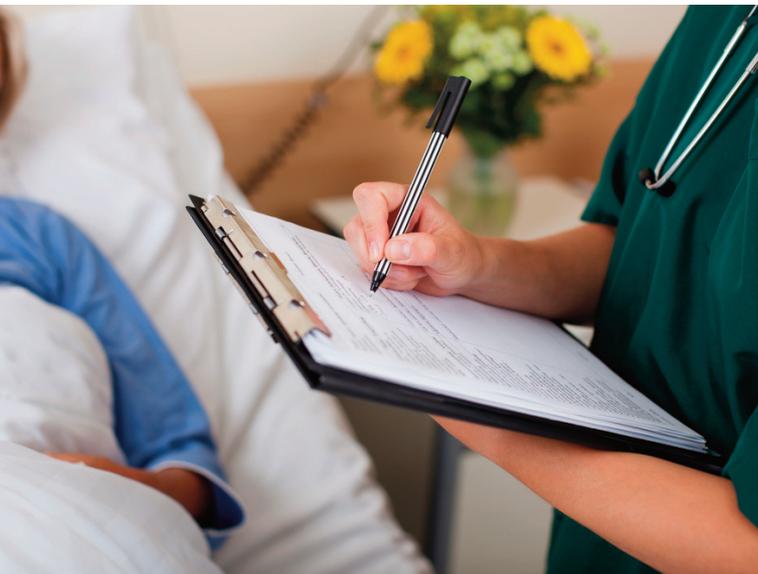
RESUMO

Objetivo: este artigo tem como objetivo sistematizar as primeiras etapas percorridas pelo investigador na elaboração, no processo de adequação e validação de instrumentos de medida de auto-preenchimento. **Método:** para o efeito foi realizada uma revisão da literatura, recorrendo à consulta de manuais e à pesquisa em bases de dados internacionais. **Resultados:** neste artigo dá-se especial enfoque à análise da fidelidade e validade dos referidos instrumentos de medida. Realça-se, também, a utilização da análise fatorial exploratória e confirmatória para a validação de construto. **Conclusão:** os instrumentos de medida devem ser precisos e validados, de modo a garantir a fiabilidade dos resultados. **Descritores:** validação; análise fatorial; psicometria.

ABSTRACT

Objective: the current article intends to systematize the first stages covered by the investigator in the elaboration, adequacy and validation of the instruments of measurement which are auto-complete. **Methodology:** a review of the literature was done, using manuals and international data bases. **Results:** in this paper special emphasis is given towards the construction and adequacy of the instruments of measurement. The use of exploratory and confirmatory factorial analysis to validate the construct, is also emphasized. **Conclusion:** the instruments of measurement must be precise and validated in order to guarantee the reliability of the results. **Descriptors:** validation; factor analysis; psychometrics.





INTRODUÇÃO

Atualmente podemos definir ciência por um conjunto organizado de conhecimento sobre a realidade, obtidos através do método científico (Almeida & Freire, 2003), para que possam ser reconhecidos pela comunidade acadêmica, prestadores de cuidados em saúde e decisores políticos (Ferreira & Marques, 1998).

Este domínio do conhecimento científico centra-se nos processos de recolha, análise e caracterização da informação, permitindo a tomada de decisão a partir dos dados recolhidos, mas a utilização de métodos de análise, em estatística, apenas assume significado quando os números são devidamente contextualizados, no fenómeno em estudo (Marôco & Bispo, 2003).

Dando continuidade ao artigo publicado anteriormente, com título “Construção, adaptação transcultural e adequação de instrumentos de medida”, prosseguiremos com a ideia de que os conhecimentos nesta área são fundamentais para a construção e adequação de instrumentos de medida de autopreenchimento, quer em situações de construção original de um instrumento de medida ou perante a utilização um instrumento de medida já construído, mas com necessidade de tradução e adaptação para a realidade linguística e cultural de um outro país.

A utilização de instrumentos de medida válidos conduz a uma quantificação de resultados passíveis de estudar o fenómeno em si, de definir novas linhas de investigação e mensurar ganhos em saúde para a população das intervenções implementadas (LoBiondo-Wood & Haber, 2001; Polit & Beck, 2010).

Pensamos que este artigo possibilitará a outros investigadores, agilizar procedimentos, uma vez que se sistematizam aqui, as etapas percorridas na elaboração, adequação e validação de instrumentos de medida, abordando os critérios estatísticos para a tomada de decisão, no que concerne à fiabilidade e validade das medidas.

FIDELIDADE OU CONFIABILIDADE

A fidelidade, ou confiabilidade, é a precisão e a consistência das medidas obtidas quando se utiliza um instrumento de medida. Significa que o instrumento é fiel, que

se obtém resultados semelhantes em situações comparáveis, mensurando-se sob a forma de coeficiente de correlação (r). O r operacionaliza-se num contínuo entre 0.00 (ausência de correlação) até 1.00 (correlação perfeita) (Almeida & Freire, 2003; Hill & Hill, 2002; Marôco, 2007; Pais-Ribeiro, 2007). Para a sua avaliação utilizam-se diversos critérios:

- **Consistência interna ou homogeneidade:** assegura a concordância entre os diferentes itens que pretendem medir distintos aspetos de um conceito. É usualmente avaliada através do cálculo do *alfa de Cronbach* (α) ou coeficiente *alfa* (Polit & Beck, 2010), que averigua se cada enunciado, de um instrumento de medida, mede da mesma forma que outros, um determinado conceito. O seu valor varia de 0.00 a 1.00 e quanto mais itens tiver, mais elevado será o seu valor (Pais-Ribeiro, 2007). A reavaliação do (α), sempre que o instrumento é utilizado, é considerada como uma etapa de rigor metodológico do estudo em desenvolvimento. Para melhor interpretação, operacionaliza-se os resultados do (α) da seguinte forma: <0.6 – Inaceitável; [0.6 e 0.7[- Fraco; [0.7 e 0.8[- Razoável; [0.8 e 0.9 – Bom e [1 e 0.9] – Excelente (Hill & Hill, 2002; Marôco, 2007).

Citando alguns exemplos, no estudo o valor *alfa de Cronbach* da escala Perceção do Trabalho em Equipa é de 0.93, apresentando uma consistência excelente (Grilo & Mendes, 2011).

- **Estabilidade:** consiste no grau de concordância entre duas medidas obtidas através do teste/reteste, nas mesmas pessoas, em momentos diferentes, cujo intervalo pode ir de duas a quatro semanas (Polit & Beck, 2010). Esta análise preconiza-se ser utilizada para variáveis relativamente estáveis no tempo (ex: traços de personalidade), assim como atendendo a circunstâncias em que não é previsível a alteração das variáveis em estudo (ex: humor). Considera-se que quanto menor for o intervalo de tempo entre os dois momentos, maior será o efeito de aprendizagem dos respondentes no seu preenchimento, o que poderá funcionar como desvantagem. A relação entre as duas medidas é expressa por um coeficiente de estabilidade e variam pouco se se obtiver um coeficiente de correlação (r) elevado, ou seja, progressivamente maior que 0.70.
- **Equivalência:** quando dois ou mais observadores têm uma alta correlação na coerência ou concordância de um mesmo comportamento observado, ou quando se utilizam formas alternadas de um mesmo teste e se obtém uma alta correlação (LoBiondo-Wood & Haber, 2001).

VALIDADE

A validade é o segundo elemento mais importante para avaliar a qualidade de um instrumento de medida, é o grau de precisão (garantia) com que o instrumento mede o que é suposto (Polit & Beck, 2010). A validade é uma característica importante que o instrumento deve possuir e refere-se ao grau que o mesmo mede e se supõe medir (Ribeiro *et al*, 2012).

Embora sejam elementos de avaliação distintos, a validade e fidelidade não se comportam de forma total-

Luís Manuel Mota de Sousa, Mestre. Enfermeiro Especialista em Reabilitação no Hospital Curry Cabral. Professor Assistente na Universidade Atlântica. Doutorando em Enfermagem na Universidade Católica Portuguesa; **Cristina Maria Alves Marques-Vieira**, Mestre. Enfermeira Especialista em Reabilitação. Assistente no Instituto de Ciências da Saúde da Universidade Católica Portuguesa. Doutoranda em Enfermagem na Universidade Católica Portuguesa; **Maria Leonor Ramos Carvalho**, Enfermeira Especialista em Enfermagem Médico-Cirúrgica, Professora Assistente na Universidade Atlântica. Doutoranda em Ciências de Educação na Universidade de Sevilha; **Filipa Veludo**, Mestre. Enfermeira Especialista em Enfermagem Médico-Cirúrgica. Assistente no Instituto de Ciências da Saúde da Universidade Católica Portuguesa. Doutoranda em Enfermagem na Universidade Católica Portuguesa; **Helena Maria Guerreiro José**, Doutora, Professora Auxiliar no Instituto de Ciências da Saúde da Universidade Católica Portuguesa, *Scholar of The European Academy of Nursing Sciences*.



“ A utilização de instrumentos de medida válidos conduz a uma quantificação de resultados passíveis de estudar o fenómeno em si, de definir novas linhas de investigação e mensurar ganhos em saúde para a população das intervenções implementadas ”

mente independente (Polit & Beck, 2010; Hill & Hill, 2002; Pais-Ribeiro, 2007; Polit *et al*, 2004).

Embora uma medida possa assumir boa fiabilidade e validade fraca, não evidencia coerência, no sentido em que não poderá medir validamente um atributo se não for fiável (Polit & Beck, 2010). A avaliação da validade pode centrar-se em quatro aspetos (Murphy. & Davidshofer, 1998).

VALIDADE DE CONTEÚDO

A validade do conteúdo deve ser estabelecida, sendo o pré-requisito para todos os tipos de validade, uma vez que não requer qualquer tipo de tratamento estatístico e revela o quão representativas são as perguntas do instrumento dentro do universo de todas as perguntas que poderiam ser feitas sobre uma temática específica (Polit *et al*, 2004).

Podendo ser analisada em cinco etapas (tabela 1), a validade de conteúdo assegura a abrangência e representatividade dos enunciados do um instrumento (itens) para medir o conceito em estudo, permitindo que este seja interpretado de forma apropriada (Paschoal, 2000) e centrando-se na procura da representabilidade dos itens, no domínio em estudo (Polit *et al*, 2004).

Tabela 1: Etapas na análise de validade de conteúdo (Green & Lewis, 1986)

Etapas na análise de validade de conteúdo

1. Revisão da literatura do conceito em estudo
2. Reflexão sobre a significação do conceito no contexto em estudo
3. Identificação das dimensões do conceito
4. Formulação de enunciados utilizando os indicadores específicos identificados
5. Análise das relações entre os enunciados

Pela dificuldade de análise da validade de conteúdo, de um instrumento de medida, de forma objetiva solicita-se a colaboração de peritos, de três a cinco, pretendo-se que estimem se este mede o que é suposto. Esta avaliação poderá também ser objetivada com a utilização de indicadores de validade de conteúdo, ou seja, pela percentagem de concordância entre peritos na sua aplicação, classificando cada item individualmente: 1 - Pertinente; 2 - Pouco pertinente; 3 - Pertinente e 4 - Bastante pertinente (Walzy. & Bausse, 1981).

O índice de validade de conteúdo é o cálculo efetuado, tendo por base o número de enunciados pontuados com 3 e 4, sobre o número total de enunciados ou questões, sendo aceitável um resultado superior a 0.80.

VALIDADE DE CONSTRUCTO

As leis científicas são construídas tendo como base os constructos (Pais-Ribeiro, 2007). Os constructos ocorrem com regularidade e estão ligados a fenómenos observáveis e concretos, sendo que quanto mais abstratos forem os conceitos, mais difícil será estabelecer a validade de constructo (Polit *et al*, 2004; Murphy. & Davidshofer, 1998). A definição de um constructo poderá ser operacionalizada em diferentes fases (Tabela 2).

Tabela 2: Fases para definição de um constructo (Pais-Ribeiro, 2007)

Fases para definição de um constructo

- Identificação dos comportamentos que constituem o constructo a medir
- Identificação de outros constructos, para estabelecer a relação com o que se pretende medir
- Determinação da forma como o comportamento do constructo em estudo se relaciona com outros identificados anteriormente

• A validade de constructo pode ser medida de diversas formas, tendo por base o teste das relações centrado nos conceitos teóricos, pretendendo dar resposta ao



As principais razões para a utilização deste tipo de análise são a informação e a validação dos resultados de uma avaliação e o desenvolvimento de uma teoria, tendo em consideração a natureza do constructo e a síntese de relações,...



que o instrumento realmente mensura (Polit & Beck, 2010), podendo ser realizada através de validade convergente e divergente, grupos de contraste e análise fatorial (LoBiondo-Wood & Haber, 2001).

Por validade convergente entende-se a verificação de maior correlação entre instrumentos que meçam o mesmo constructo, quando comparada com a encontrada entre instrumentos que medem constructos diferentes, nos mesmos sujeitos. A validade divergente refere-se à baixa correlação obtida, quando se utilizam instrumentos que medem constructos diferentes, face aos que medem o mesmo constructo (LoBiondo-Wood & Haber, 2001; Polit & Beck, 2010; Pais-Ribeiro, 2007). A análise de grupos contrastantes consiste na comparação de um grupo, com características específicas, com outro, no qual tais características não existem (LoBiondo-Wood & Haber, 2001). A análise fatorial permite determinar se os conceitos de um instrumento estão fortemente ligados entre si, bem como, se os enunciados se reagrupam em torno de um só fator (Nunnally & Bernstein, 1994).

A abordagem estatística para avaliar a validade de constructo passa pela aplicação de um dos dois tipos de procedimentos possíveis: a análise fatorial exploratória e a análise fatorial confirmatória. Estas duas abordagens refletem perspetivas paradigmáticas diferentes necessitando da respetiva contextualização ao fenómeno em estudo (Pais-Ribeiro, 2007). A análise fatorial exploratória enquadra-se nos métodos indutivos, onde se perspetiva a descoberta de um conjunto ótimo de variáveis latentes. Neste procedimento utiliza-se software como o SPSS e não se conhecem hipóteses prévias (Marôco, 2010a). A análise fatorial confirmatória é um método hipotético dedutivo que pretende testar hipóteses sobre relações entre variáveis, que constituem fatores comuns. O número e a interpretação dos resultados já são conhecidos, antes da realização da análise fatorial confirmatória. Nesta análise utiliza-se *software* como o AMOS (Marôco, 2010b) ou o LISREL.

Descrevem-se agora, de forma pormenorizada, os procedimentos para cada análise fatorial mencionada.

ANÁLISE FATORIAL EXPLORATÓRIA

Foi concebida para efetuar a redução dos dados iniciais, permitindo procurar as variáveis latentes (fatores) que apresentem um poder explicativo relevante. Esta análise é utilizada como ferramenta base na determinação das qualidades métricas de um instrumento de medida, possibilitando a descrição de um conjunto de variáveis (ou itens), que são elementos comuns entre os vários itens (Martinez & Ferreira, 2007). Quando se agrupam

essas variáveis obtêm-se os fatores, através de inter-relações, sendo que nesta análise, o investigador não tem uma ideia pré-definida sobre o número de fatores a extrair, daí a denominação de análise fatorial exploratória (Reyment & Jöreskog, 1996).

As principais razões para a utilização deste tipo de análise são a informação e a validação dos resultados de uma avaliação e o desenvolvimento de uma teoria, tendo em consideração a natureza do constructo e a síntese de relações, na forma de um conjunto de fatores que podem ser utilizados em análises posteriores (Thompson, 2004).

O número de sujeitos da amostra, necessário para uma análise fatorial, é no mínimo cinco por item (ex: se pretender fazer uma análise fatorial de um teste com 25 itens, devem incluir-se pelo menos 125 sujeitos). Este critério é também utilizado na regressão múltipla em que a análise fatorial deverá incluir pelo menos 100 casos (Bartlett *et al*, 2001). Constatam-se outras recomendações (10 > 20 sujeitos por item) que, no entanto, também assumem como possível o mínimo de cinco sujeitos por item (Gorsuch, 2003). Num estudo (Chaves *et al*, 2010), utilizou-se uma amostra de 52 pessoas (considerada pequena), contudo a escala apresentava 5 itens, o que permitiu ter 10 respondentes por item. Numa tentativa de uniformização apresenta-se uma forma de classificação para as amostras adequadas, à realização de análises fatoriais exploratórias: n=50 – Muito pobre; n=100 – Pobre; n=200 – Razoável; n=300 – Bom e n=500 – Muito bom (Comrey & Lee, 1992).

Em situação de amostras inferiores a 60 pode realizar-se a análise fatorial, desde que as comunalidades sejam superiores a 0.60. Em casos de comunalidades de 0.50, são necessárias amostras superiores a 100 (MacCallum *et al*, 1999).

Quando os itens, ou variáveis, apresentam correlações, com os fatores superiores a 0.80, não requerem amostras grandes, sendo suficiente 150 sujeitos (Guadagnoli & Velicer, 1988).

Na extração de fatores, por vezes, surgem dúvidas relativamente ao método a utilizar na análise fatorial exploratória (se a opção é a análise dos componentes principais ou análise fatorial). De facto, existem diferenças nos dois métodos já que a análise dos componentes principais visa retirar a máxima variância dos dados, com o mínimo de componentes ortogonais, incluindo o erro associado à variância da variável observada. Este método é recomendado no tratamento empírico dos dados, quando se pretende reduzir um número elevado de variáveis, obtendo-se uma única solução matemática.

Na análise fatorial o erro da variância é omitido, apenas se encontrando disponível, para análise, a variância que a variável observada partilha com as outras. Nem todas as soluções são únicas com este método, pois a análise fatorial parece ser a mais apropriada em estudos teóricos que não sejam “contaminados” por erros de variância (Martinez & Ferreira, 2007).

A análise fatorial procura estimar os fatores comuns e os fatores específicos dos dados, mas antes de iniciar a estimação dos seus pesos, é necessário definir quando é que as correlações entre as variáveis originais são suficientemente elevadas para que tenha utilidade na estimação dos fatores comuns. Neste contexto, existe o teste de Esfericidade de Bartlett para testar [H_0 : Matriz (π) de correlações populacionais = variância (I) vs H_1 : $\pi \neq I$]. Quando se rejeita H_0 conclui-se que as variáveis estão correlacionadas significativamente. Contudo, em amostras de grandes dimensões, poder-se-á rejeitar H_0 , mesmo quando as correlações são muito reduzidas. Este teste exige que as variáveis apresentem distribuição normal, sendo muito sensível à violação deste pressuposto (Marôco, 2010a). Neste sentido é pouco utilizado uma vez que é muito sensível ao tamanho da amostra].

O método de utilização mais geral é a “medida da adequação da amostragem de Kaiser-Meyer-Olkin - KMO”, sendo uma medida da homogeneidade das variáveis, que compara as correlações simples com as correlações parciais observadas entre as variáveis. Os valores do KMO são < 0.50 são inaceitáveis, entre 0.5 e 0.6 são maus, mas aceitáveis, entre 0.6 e 0.7 são medíocres, entre 0.7 e 0.8 são médios, entre 0.8 e 0.9 são bons e entre 0.9 e 1.0 são excelentes (Kaiser, 1970; Kaiser & Rice, 1974).

Com valores superiores a 0.5 pode utilizar-se todas as variáveis em estudo e prosseguir a análise (Marôco, 2010a).

Após a obtenção de uma solução fatorial, ou seja, depois de estimados os fatores que explicam o comportamento apenas, o número mínimo de fatores que permitam explicar convenientemente o fenómeno em estudo. As regras mais utilizadas, na prática, para decidir o número de fatores mais apropriados são (Marôco, 2010a):

- O critério Kaiser propõe a extração de fatores com *eigenvalues* (valores próprios) superiores à unidade (Guttman, 1954). Os números próprios representam a variância dos fatores. Esta tese tem na sua égide o facto da variância, de cada variável estandardizada, contribuir para a extração dos componentes principais com apenas uma unidade, por isso não se aceitam valores inferiores a um (Pais-Ribeiro, 2007; Martinez & Ferreira, 2007). Este critério tem sido indicado para situações em que o número de variáveis é menor que trinta e a média das comunalidades é maior que 0.70 ou, então, se a amostra for superior a 250 casos e a comunalidade média for maior ou igual a 0.60 (Bryman & Cramer, 1993);
- Critério do *Scree plot* consiste em representar num gráfico os valores próprios associados a cada fator (Cattell, 1966). A retenção dos fatores faz-se até ao momento em que começa a inflexão da curva (cotovelo). De acordo com este critério, é neste ponto que se obtém o número mínimo de fatores (Pais-Ribeiro, 2007; Marôco, 2010a);



- Na variância extraída por cada fator e variância extraída total, os fatores são retidos se apresentarem pelo menos 5% da variância total. Deve-se extrair um número mínimo de fatores de forma a explicar pelo menos 50% da variância total das variáveis originais, sendo este o valor mínimo de variância aceitável (Marôco, 2010a). A título de exemplo, na validação da escala de espiritualidade de Pinto e Pais-Ribeiro em pessoas com insuficiência renal crónica em hemodiálise obtiveram uma variância total de 75,1%, sendo considerada válida, pois foi superior a 50% (Chaves *et al*, 2010);
- O ajustamento significativo é aplicado quando se utiliza o método da máxima verosimilhança na extração dos fatores, em que se utiliza a estatística χ^2 para comparar modelos com diferentes números de fatores. Neste caso, faz-se a adição de fatores um a um e termina-se quando a variância da estatística do χ^2 do “modelo com mais um fator” *versus* “modelo anterior com menos um fator”, não for significativa. Quando as diferenças da estatística χ^2 dos dois modelos são superiores a 4 indicam, aproximadamente, que o modelo com mais um fator se ajusta significativamente melhor à estrutura correlacional das variáveis originais, do que o modelo com menos um fator aceitável (Marôco, 2010a).

A nível teórico existem mais dois critérios para seleccionar o número de fatores que deverá expressar a escolha do número de processos teóricos subjacentes à área que está a ser investigada (Pais-Ribeiro, 2007):

- Seleção semelhante ao critério de Kaiser, que resulta do número de variáveis a dividir por três, ou cinco. Este critério é razoável se a amostra for grande e o número de itens for inferior a 40 (Tabachnick & Fidell, 1996);
- Para cada fator têm de existir pelo menos três itens que tenham uma carga, que respeite a relação entre o valor da carga fatorial mais elevada e a menor, tendo em consideração a fórmula: $C = \sqrt{H/2}$ (H: valor do item com carga fatorial mais elevada; C: valor mínimo de carga fatorial considerado aceitável para um item) (Boyle & Harrison, 1981).

Além dos critérios que são utilizados na determinação e retenção de fatores, na análise fatorial exploratória,



A validade de critério indica a eficiência de um teste em prever o comportamento de pessoa numa determinada situação.



existem disponíveis em programas estatísticos vários métodos de extração dos fatores (Martinez & Ferreira, 2007), onde destacamos:

- O Método dos componentes principais é o mais utilizado e recomendado como primeira abordagem (Marôco, 2010a; Tabachnick & Fidell, 1996). Assume que os valores das variáveis possuem uma consistência interna perfeita e tenta reproduzir a variância de informação dos dados e não da população. Não sendo aconselhado em amostras não representativas da população, é bastante recomendado quando existe um elevado número de variáveis (itens), num número reduzido de componentes (Martinez & Ferreira, 2007);
- A análise dos eixos principais (*Principal Axis Factoring*) utiliza o coeficiente de comunalidades para substituir os valores da diagonal da matriz fatorial. Este processo efetua as interações entre as variáveis até que os valores da variável estabilizem, formando posteriormente fatores. Este método é o mais desejável, uma vez que assume uma estrutura fatorial composta por fatores comuns e específicos. Contudo, apresenta como desvantagem o facto de ser fortemente dependente da amostragem efectuada (Marôco, 2010a);
- A análise fatorial do *Alpha* (*Alpha factor Analysis*) consiste na formação de fatores com valores de consistência interna máximas (Martinez & Ferreira, 2007);
- O método de máxima verosimilhança (*Maximum Likelihood*) permite criar os fatores que reproduzem correlações ou matrizes de covariância na população ao invés da amostra. A estimação dos valores da população é obtida, tendo em consideração a elevada probabilidade de pertencerem à amostra da matriz de correlações observada (Martinez & Ferreira, 2007). Este método é pouco recomendado, pois exige que as variáveis em estudo apresentem distribuição normal multivariada, no entanto, em caso de amostras de grandes dimensões, permite calcular índices avaliativos da qualidade do modelo fatorial (Marôco, 2010a).

No processo de desenvolvimento das matrizes da análise fatorial exploratória, realiza-se a rotação dos eixos das estruturas fatoriais que obedecem a dois critérios de rotação: ortogonal e oblíqua. Na rotação ortogonal, um ângulo de 90° é mantido durante este processo, permitindo assim que os fatores se mantenham não correlacionados. Este tipo de rotação é utilizada para maximizar as elevadas correlações e minimizar as inferiores (Martinez & Ferreira, 2007). Os dois métodos mais populares são o método *varimax* e o *quartimax*.

O objetivo no método *varimax* é obter uma estrutura fatorial na qual uma e apenas uma das variáveis originais, esteja fortemente associada com um único fator e pouco associada com os restantes (Kaiser, 1958).

No método *quartimax*, o objetivo é encontrar uma estrutura fatorial onde todas as variáveis possuam pesos elevados num fator (geral) e que cada variável tenha pesos fatoriais elevados num outro fator (específico) e pesos reduzidos nos outros fatores (Marôco, 2010a). A rota de difícil interpretação.

Este método forma estruturas mais simples e interpretáveis (Martinez & Ferreira, 2007), contudo alguns autores não o recomendam, por ir contra ao fundamento da análise fatorial exploratória (Marôco, 2010a).

ANÁLISE FATORIAL CONFIRMATÓRIA

Consiste numa técnica de modelação linear geral, que tem por objetivo identificar um conjunto reduzido de variáveis latentes (fatores), que permitem explicar a estrutura correlacional observada entre um conjunto de variáveis manifestas ou itens. Esta análise é utilizada quando existe informação prévia sobre a estrutura fatorial que é necessário confirmar, ou seja, pretende verificar se determinados fatores latentes são responsáveis pelo comportamento de determinadas variáveis manifestas, específicas, tendo em conta um determinado estudo ou teoria. Por defeito, a análise fatorial confirmatória, considera que os fatores estão correlacionados. Os fatores específicos (erros ou resíduos), podem estar correlacionados, sugerindo a existência de problemas com a conceptualização das medidas das variáveis manifestas e/ou a existência de fatores comuns que não estão estabilizados no modelo (Marôco, 2010a). O modelo análise fatorial confirmatória é estabelecida pelo investigador, tendo em conta a teoria ou estudos anteriores.

A estimação do modelo consiste na obtenção de estimativas dos parâmetros deste, que melhor reproduzem os dados observados na amostra, sendo efetuada a partir das matrizes de covariância das variáveis manifestas. Esta estimação é feita através de métodos que procuram maximizar a verosimilhança das covariâncias entre as variáveis manifestas, ou através da procura de minimizar os erros de ajustamento.

Na fase de avaliação da qualidade do ajustamento do modelo, pretende-se avaliar se o modelo é capaz de reproduzir a estrutura correlacional das variáveis manifestas, observadas na amostra. A qualidade do ajustamento do modelo é realizada, usualmente, através de testes de ajustamento, de índices empíricos que baseiam-se nas funções de verosimilhança ou na matriz dos resíduos obtidos durante o ajustamento e análise dos resíduos e da significância dos parâmetros (Almeida & Freire, 2003).

O teste de χ^2 de ajustamento permite testar a significância da função de discrepância minimizada durante o ajustamento do modelo (Marôco, 2010b).

Esta análise não será desenvolvida neste artigo.

VALIDADE DE CRITÉRIO

A validade de critério indica a eficiência de um teste em prever o comportamento de pessoa numa determinada situação (Anastasi, 1990). Existem dois tipos de validade de critério: preditiva (capacidade de prever uma situação futura, características e atitudes, a partir de resultados atuais) e concorrente (quando se verifica, ao mesmo tempo, correlação entre duas medidas que medem o mesmo conceito). Um valor (r), igual ou superior a 0.70, indica que o grau de validade é aceitável (Polit *et al*,

2004). Na avaliação da validade convergente da escala de espiritualidade de Pinto e Pais-Ribeiro com a escala de qualidade de vida da Organização Mundial de Saúde, obteve-se uma correlação fraca a moderada e significativa ($r=0,33$), mas para os autores era um instrumento de referência por se encontrar validado para português (Chaves *et al*, 2010).

Finda esta sistematização das etapas para equivalência de instrumentos de medida, optámos por as plasmar na figura 1.

Figura1: Etapas da sistematização da equivalência de instrumentos de medida



Fonte: Adaptado de Pais-Ribeiro (2007)

Considerações finais

Este artigo deu continuidade ao anteriormente publicado com o título “Construção, adaptação transcultural e adequação de instrumentos de medida”. Pretendeu-se com estas publicações reforçar a ideia que a construção e a adequação, mediante a validação do instrumento de medida de autopreenchimento, são fundamentais para avaliar os constructos identificados na primeira etapa do processo de investigação. Deste modo, foi propósito central, a sistematização do processo, através de uma síntese dos principais procedimentos e critérios, passíveis de facilitar a análise dos *outputs* obtidos pelos *softwares* estatísticos.

Na realidade consideramos que esta síntese poderá ser facilitadora da tomada de decisão e do desenho rigoroso do percurso inerente ao estudo da fidelidade e validade dos instrumentos de medida, utilizados na área da saúde.

Importa realçar que, em todas as situações referidas, os autores consultados apontam para a necessidade de utilização de instrumentos padronizados, precisos e validados, de modo a garantir a fiabilidade dos resultados obtidos.



BIBLIOGRAFIA

- Almeida, L.S. & Freire, T.(2003). *Metodologia da investigação em psicologia e educação*. 3ª ed. Braga: Psiquilibrios.
- Anastasi, A. (1990). *Psychological testing*. New York: MacMillan.
- Bartlett, J.E., Kotriak, J.W. & Higgins, C.C. (2001) Organizational research: Determining appropriate sample size in survey research. *Information Technology, Learning and Performance Journal*, 19(1), 43-50.
- Boyle, E. & Harrison, B. (1981). Fator structure of the health locus of control scale. *Journal of Clinical Psychology*, 37(4), 819-824.
- Bryman, A. & Cramer, D. (1993). *Análise de dados em ciências sociais: Introdução às técnicas utilizando o SPSS*. Oeiras: Editora Celta.
- Cattell, R.B (1966). *Handbook of multivariate experimental psychology*. Chicago: Rand McNally.
- Chaves, E.C.L., Carvalho, E.C., Dantas, R.A.S, Terra, F.S., Nogueira, D.P. & Souza L. (2010). Validação da escala de espiritualidade de Pinto e Pais-Ribeiro em pacientes com insuficiência renal crônica em hemodiálise. *Revista de Enfermagem UFPE On Line*, 4(2),715-721.
- Comrey, A.L. & Lee, H.B. (1992). *A first course in fator analysis*. Hillsdale, NJ: Erlbam.
- Ferreira, P.L. & Marques, F.B.(1998). *Avaliação psicométrica e adaptação cultural e linguística de instrumentos de medição em saúde: Princípios metodológicos gerais*. Documento de trabalho 1. Centro de Estudos e Investigação em Saúde da Universidade de Coimbra.
- Gorsuch, R.L. (2003). Fator analysis. In J.A. Schinka & W. F. Velicer (Eds), *Handbook of Psychology: Research Methods in Psychology*.
- Green, L.W. & Lewis, F.M. (1986). *Measurement and Evaluation in Health Education and Health Promotion*. Palo Alto: Mayfield Public.
- Grilo, C.M.C. & Mendes, F.R.P.(2011). The healthcare professionals perception on teamwork in a hospital context. *Journal of Nursing UFPE*, 5(2), 336-343
- Guadagnoli, E. & Velicer, W.F. (1988). Relation of sample-size to the stability of coponent patterns. *Psychological Bulletin*, 103(2), 265-275.
- Guttman, L. (1954). A new approach to factor analysis: The radex. In *Mathematical Thinking in the Social Sciences*. New York: Ed. Lazarsfeld.
- Hill, M.M. & Hill, A.B. (2002). *Investigação por Questionário*. 2ª Edição. Lisboa, Edições Sílabo.
- Kaiser, H.F. & Rice, J.L. (1974). *Educational and Psychological Measurement*, 34, 111-117.
- Kaiser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.
- Kaiser, H.F. (1970). A second generation Little Jiffy. *Psychometrika*, 35, 401-415.
- LoBiondo-Wood, G. & Haber, J (2001). *Pesquisa em enfermagem: Métodos, avaliação crítica e utilização*. 4ª ed. Rio de Janeiro: Editora Guanabara Koogan S.A.
- MacCallum, R.C., Widaman, K.F., Zhang, S. & Hong, S.(1999). Sample size in factoranalysis. *Psychological Methods*, 4, 84-99.
- Marôco J (2007). *Análise estatística com utilização do SPSS*. 3ª Ed. Lisboa: Edições Sílabo.
- Marôco, J. (2010a). *Análise estatística com o PASW Statistic (ex-SPSS)*. Pero Pinheiro: ReportNumber.
- Marôco, J. (2010b). *Análise de equações estruturais: fundamentos teóricos, software e aplicações*. Pero Pinheiro: ReportNumber.
- Marôco, J. & Bispo, R (2003). *Estatística aplicada às ciências sociais e humanas*. Manuais universitários 27. Lisboa: Climepsi Editores.
- Martinez, L. & Ferreira, A. (2007). *Análise de dados com SPSS: primeiros passos*. Lisboa: Escolar Editora.
- Murphy, K. & Davidshofer, C. (1998). *Psychological testing: principles and applications*. 4th ed. New Jersey: Prentice-Hall International, Inc.
- Nunnally, J.C. & Bernstein, I.H. (1994). *Psychometric theory*. 3rd ed. New York: McGraw-Hill.
- Pais-Ribeiro, J.L.(2007). *Metodologia de investigação em psicologia e saúde*. Porto: Legis Editora.
- Paschoal, S.M.P.(2000). *Qualidade de vida do Idoso: Elaboração de um instrumento que privilegia sua opinião*. Dissertação de mestrado em Medicina na Universidade de São Paulo.
- Polit, D. & Beck, C. (2010). *Nursing Research: Appraising evidence for nursing practice*. 7td ed. Philadelphia: Lippincott Williams & Wilkins.
- Polit, D., Beck, C. & Hungler, B.(2004). *Fundamentos de pesquisa em enfermagem: métodos, avaliação e utilização*, 5ª Ed. Porto Alegre: Artmed.
- Reyment, R. & Jöreskog, K.G. (1996). *Applied fator analysis in the natural sciences*. Cambridge, UK: Cambridge University Press.
- Ribeiro, M.C.O., Pereira, C.U., Hora, E.C., Nunes, M.S., Silva, C.B. & Santos, D.S.(2012) Construção e validação de instrumento de coleta de dados para vítimas de trauma craneocefálico. *Revista de Enfermagem UFPE On Line*, 6(4), 1118-29.
- Tabachnick, B.G. & Fidell, L.S. (1996). *Using multivariate statistics*. 3rd ed. New York: Harper Collins.
- Thompson, B. (2004). *Exploratory and confirmatory fator analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association; 2004.
- Walyz, C.F. & Baussel, R.B. (1981). *Nursing research: Design, statistics and computer analysis*. Philadelphia: F.A. Davis.