

## CONSTRUÇÃO E VALIDAÇÃO DE ESCALAS DE ATITUDES

*Verônica Yumi Kataoka*

*Universidade Estadual de Santa Cruz /Universidade Bandeirante de São Paulo*

*veronicayumi@terra.com.br*

### **Resumo:**

O objetivo nesse artigo é apresentar, de forma resumida, duas técnicas estatísticas: Teoria de Resposta ao Item (TRI) e Análise Fatorial, utilizadas na validação de escalas de atitudes. A atitude, por ser um construto, não pode ser medida diretamente, sendo necessário utilizar instrumentos que possam medir variáveis secundárias que compõem esse construto. Um instrumento utilizado por muitos pesquisadores para avaliar as atitudes dos alunos em relação a uma determinada disciplina é a escala do tipo likert. No processo de construção de uma escala é importante buscar evidências de validade, para que o instrumento proposto possa atender a duas características principais: confiabilidade (capacidade da escala medir com fidelidade o constructo avaliado) e validade, e, por conseguinte, possa auxiliar no processo de avaliação do construto estudado.

**Palavras-chave:** Evidências de validade; escalas tipo likert; análise fatorial; Teoria de Resposta ao Item.

### **1. Introdução**

A atitude é um construto ou traço latente, sendo assim uma característica de um indivíduo que não pode ser observada diretamente e que deve ser avaliada a partir de observações de variáveis secundárias que estejam relacionadas a ela (ANDRADE; TAVARES; VALE, 2000).

Segundo Novaes (2007), uma das formas de medir as atitudes em relação a uma determinada disciplina é por meio de escalas do tipo Likert. Estas escalas são compostas de itens que podem contemplar uma ou mais dimensões, que podem ser, por exemplo, afetiva, cognitiva, valor frente a disciplina, utilidade, entre outras. A partir das dimensões é possível indentificar se as atitudes dos alunos são mais positivas ou mais negativas, conhecendo-se assim o comportamento dos mesmos em relação à uma disciplina.

Brito (1998) justifica a utilização da escala de atitudes em relação à Matemática, para que o professor verifique as atitudes de seus alunos no início do período letivo e, reaplique o instrumento após a intervenção, analisando dessa forma se ocorreu mudança

nas atitudes, principalmente se forem detectadas atitudes negativas, a fim de estabelecer estratégias apropriadas de ensino.

Diversos pesquisadores propuseram escalas de atitudes em relação à Matemática, como por exemplo, a escala unidimensional (dimensão afetiva) do tipo likert de Aiken e Dreger (1961), composta por 20 itens, e que foi aplicada pela primeira vez a 127 alunos matriculados na disciplina de Matemática no primeiro ano de graduação. Nesse estudo os pesquisadores verificaram uma relação positiva entre as atitudes e a nota na disciplina.

A partir da escala de Aiken e Dreger (1961), Brito (1996) propôs e validou, em português, uma escala de atitudes em relação à Matemática, composta por 20 itens, sendo 10 itens em relação às atitudes positivas e 10 em relação às atitudes negativas, com respostas do tipo Likert que variam entre discordo fortemente (DF), discordo (D), concordo (C) e concordo fortemente (CF). Ressalta-se que essa mesma autora buscou evidências de validade na sua escala utilizando a técnica estatística da Análise Fatorial.

Nesse contexto, o objetivo nesse artigo é apresentar, de forma resumida, duas técnicas estatísticas: Teoria de Resposta ao Item (TRI) e Análise Fatorial, utilizadas na validação de escalas de atitudes do tipo likert.

## **2. Evidências de validade**

A validade pode ser compreendida como uma verificação direta da possibilidade do instrumento satisfazer o seu objetivo (PASQUALI, 2003). São diversas as técnicas empregadas para a determinação do índice de validade de um instrumento, porém, neste artigo teremos como base a evidência de validade de construto. Segundo Dias e Vendramini (2008), a validade de construto de um instrumento é a extensão em que se pode dizer que o instrumento mede de forma precisa um construto. E nesse sentido, o uso de escalas previamente validadas permite a comparação de resultados em populações diferentes.

Para avaliar se existem evidências de validade de um instrumento, podem ser utilizadas duas técnicas de estatística multivariada que avaliam as propriedades psicométricas do mesmo: a Análise Fatorial (AF) e a Teoria de Resposta ao Item (TRI). Nesse contexto, entende-se por propriedades psicométricas neste contexto, o estudo de duas características principais para validar uma escala: confiabilidade (capacidade da escala medir com fidelidade o constructo avaliado) e validade (PASQUALI, 2003).

Salienta-se que antes de proceder às análises estatísticas, é necessário, no caso das escalas tipo likert, atribuir uma pontuação para cada possibilidade de resposta. No caso da escala de atitude proposta por Brito (1996), para cada possibilidade de resposta DF, D, C e CF, são atribuídas, respectivamente, pontuações de 1 até 4 para as afirmativas positivas e de 4 até 1 para as afirmativas negativas. Por exemplo, para o aluno que responder a 5 itens (3 afirmativas positivas e 2 negativas) dessa escala conforme apresentado na Figura 1, será associada uma pontuação total de 11 pontos .

Item	Afirmção	DF	D	C	CF	Pontos
1	Eu fico sempre sob uma terrível tensão na aula de Matemática.		X			3
2	Eu não gosto de Matemática e me assusta ter que fazer essa matéria.			X		2
3	Eu acho a Matemática muito interessante e gosto das aulas de Matemática.	X				1
4	A Matemática é fascinante e divertida.			X		3
5	A Matemática me faz sentir seguro(a) e é, ao mesmo tempo, estimulante.		X			2
<b>Pontuação Total</b>						<b>11</b>

Figura 1. Exemplo da pontuação atribuída a 5 itens da escala de atitudes em relação à Matemática (BRITO, 1996).

## 2.1 Análise Fatorial

Reis (1997) afirma que a Análise Fatorial (AF) é uma técnica estatística que tem por objetivo representar um número de variáveis iniciais a partir de um número menor de variáveis hipotéticas. No contexto de uma escala, a AF seria utilizada para descrever a estrutura de dependência dos itens, ou seja, identificar as dimensões de uma escala a partir do agrupamento de itens comuns.

Na AF para avaliar a consistência interna de uma escala pode ser utilizado o coeficiente Alfa de Cronbach (1951), que segundo Streiner (2003) é a medida das correlações entre os itens de um instrumento. Pode-se entender também que este coeficiente representa a medida pelo qual o construto está presente em cada item.

Para estudar a viabilidade de uma escala utiliza-se o teste de Kaiser- Meyer-Olkin - KMO (KAISER, 1970), além do *measure adequacy of sampling* – MSA (KAISER, 1970). De acordo com Barroso e Artes (2003) esses dois índices possibilitam verificar a existência de uma estrutura fatorial nos dados, esses mesmos autores afirmam que “valores baixos de MSA, são indícios que a respectiva variável pode ser retirada da análise sem maiores prejuízos” (p. 96) e que valores de KMO “entre 0,80 a 1,00 pode ser considerado

excelente, de 0,70 a 0,80 ótimo, de 0,60 a 0,70 bom, 0,50 a 0,60 regular e de 0,00 a 0,50 insuficiente” (p. 69).

Como dito, o estudo de Brito (1996), com 2007 estudantes do ensino fundamental e do médio, é um exemplo de utilização da AF para buscar evidências de validade de uma escala de atitudes em relação à Matemática. Dentre outros resultados, pode-se destacar que o valor de alfa de Conbrach determinado foi de 0,9494, o que de acordo com a mesma autora sugerem que os itens estão fortemente relacionados; que o valor do KMO foi de 0,9720, que pode ser considerado excelente, e que foram identificados dois fatores, um que agrupou todas as 10 afirmativas positivas e outro fator que reuniu as 10 afirmativas negativas. Com base nos resultados obtidos, essa mesma autora considerou que foram encontradas evidências de validade na escala proposta.

Outro exemplo do uso da AF é o estudo de Silveira (2011), que teve como objetivo propor e buscar evidências de validade de uma escala de atitudes em relação à Estatística, aplicada a 175 alunos do terceiro do ensino médio de sete escolas públicas do estado de São Paulo. Essa escala inicialmente proposta com 30 itens, após a AF ficou composta por 20 itens, sendo 9 afirmativas positivas e 11 negativas. De acordo com os resultados da AF, o valor de alfa de Conbrach foi de 0,844, do KMO foi de 0,813, e foram identificadas quatro dimensões: afetiva positiva, afetiva negativa, dificuldade positiva e dificuldade negativa. Com os resultados obtidos pela AF esse autor teve os primeiros indícios de que parecia haver evidências de validade na escala proposta, mas para reforçar essas evidências foi utilizada também nesse estudo a técnica da Teoria de Resposta ao Item (TRI).

## **2.2 Teoria de Resposta ao Item - TRI**

A TRI está baseada em modelos matemáticos nos quais as estimativas dos traços dependem das respostas dos sujeitos e das propriedades dos itens avaliados em uma mesma escala métrica (EMBRETSON; STEVEN, 2000). Um dos modelos da TRI é o de créditos parciais de Rasch (RASCH, 1980; MASTERS, 1982), que considera o escalonamento hierárquico das categorias de resposta dos itens e a interação entre pessoas e itens para estimar a probabilidade de cada pessoa responder a cada item.

Com o uso da TRI é possível analisar as propriedades psicométricas dos itens de uma escala de atitudes, o nível do construto que é medido pelos itens (parâmetro de

locação do item na escala de atitudes) e quanto cada item está relacionado ao construto subjacente medido pela escala (coeficiente de determinação do item).

Para avaliar a confiabilidade (consistência interna) de uma escala de atitude no uso da TRI, analisa-se, inicialmente, a qualidade dos itens para saber se todos devem fazer parte da escala ou se alguns devem ser excluídos por ter propriedades psicométricas ruins, isto é, que não se ajustem ao modelo de créditos parciais de Rasch. Para essa análise podem ser utilizados os índices de aderência, o coeficiente de alfa de Cronbach e medidas de ajuste ao modelo: *Infit* e *Outfit*<sup>1</sup> e a correlação entre o item e a medida de Rasch<sup>2</sup>.

Para testar a validade do construto utiliza-se a análise dos componentes principais, baseada no resíduo do modelo de Rasch. Para um bom ajuste do modelo de Rasch, espera-se encontrar uma variância explicada pelo modelo acima de 60% e variâncias, não explicadas pelos contrastes, inferiores a 5%. Nessa análise de resíduos, o que se deseja é encontrar o número mínimo de contrastes que expliquem o máximo de variância possível, o que é um pressuposto para considerar que o instrumento é unidimensional (mede apenas um construto).

Como dito, Silveira (2011) utilizou a TRI para buscar evidências de validade na escala de atitudes em relação à Estatística proposta em seu estudo. Esse autor observou que os índices de locação dos 20 itens da escala variaram de -0,81 a 0,81, apresentaram índices de ajuste (*infit*, *outfit*, correlação) dos itens ao modelo de Rasch considerados aceitáveis e valor de alfa de Cronbach de 0,84. Dessa forma, os resultados globais da TRI indicaram que essa escala era unidimensional, e apresentava evidências de validade, e, por conseguinte, que os seus itens estavam medindo o construto atitude. Mas, como o valor da variância explicada ficou abaixo de 60% (37,2%), esse mesmo autor concluiu que após a revisão dos itens, novos estudos deveriam ser realizados para detectar outras evidências de validade.

Outro exemplo, fora do contexto da escala de atitudes, é a pesquisa de Oliveira, Silva, Kataoka e Vendramini (2010), que buscou evidências de validade de uma escala que

---

<sup>1</sup> Os valores de *infit* e *outfit* indicam se a relação entre o índice de aderência de uma pessoa e a qualidade do item atende aos pressupostos do modelo proposto, isto é, valores adequados desses índices revelam que uma pessoa que responde positivamente a um item que revela mais do construto, tenderá a responder positivamente a um item que forneça menos informação desse construto. Segundo Linacre (2002), os valores desejáveis de *outfit* e *infit* devem variar entre 0,5 a 1,5; abaixo ou acima dessa faixa, recomenda-se que o item seja eliminado.

<sup>2</sup> De acordo com Linacre (2002) a correlação esperada são valores maiores do que 0,2, abaixo desse índice é recomendável a exclusão do item, por indicar a ineficácia do mesmo para obtenção de informação sobre o construto que está sendo medido.

avalia o uso intencional de estratégias de atenção e interação no processo de autorregulação da aprendizagem de Estatística de 236 alunos de cursos tecnológicos de graduação. Essas autoras observaram que os índices de locação dos 16 itens variaram de -1,00 a 1,12, com exceção de um item com valor igual a -2,15, que apresentou também o menor coeficiente de correlação item-medida de Rasch ( $r = 0,13$ ), indicando assim a sua ineficácia para obtenção de informação sobre autorregulação. Com a retirada desse item a consistência interna da escala aumentou, apresentando um valor de alfa de Cronbach de 0,85. Todos os índices de ajuste (*infit*, *outfit*, correlação) dos itens ao modelo de Rasch apresentaram valores considerados aceitáveis. No que se referem a avaliação da validade, a variância explicada ficou abaixo de 60% (43,6 %), indicando que os itens dessa escala devem ser revisados, inclusive, porque alguns itens apresentaram índices de locação semelhantes na escala de autorregulação, sugerindo que um ou mais itens possam ser removidos numa nova formatação da escala. Sendo assim, como Silveira (2011), as autoras concluíram que novos estudos deveriam ser realizados.

### 3. Considerações finais

No contexto da medição de construtos é fundamental a utilização de instrumentos psicométricos que possam ser considerados válidos e fidedignos. Sendo assim, os pesquisadores no processo de construção ou no aperfeiçoamento desses instrumentos devem buscar evidências de validade dos mesmos utilizando-se de técnicas estatísticas, como por exemplo, a Análise Fatorial e/ou a TRI, o que por conseguinte, poderá auxiliar no processo de avaliação do construto estudado.

### 4. Referências

AIKEN, L. R.; DREGER, R. M. The effect of attitudes on performance in Mathematics. **Journal of Educational Psychology**, v.52, n.1, p. 19-24, 1961.

ANDRADE, D.F. ; TAVARES, H.R.; VALLE R. da C. Teoria de Resposta ao Item: Conceitos e Aplicações. In: 14º Simpósio Nacional de Probabilidade e Estatística, 2000, São Paulo. **Anais...** São Paulo: Associação Brasileira de Estatística, 2000. p. 149.

BARROSO, L. P.; ARTES, R. Análise multivariada. In: Reunião anual da RBRAS, 48, SEAGRO,10, 2003, Londrina. **Anais...** Lavras: UFLA, 2003. p. 151.

BRITO, M. R. F. **Um Estudo sobre as Atitudes em relação a Matemática em estudantes de 1º e 2º graus**. Tese de Livre Docência, FE-UNICAMP. Campinas-SP, 1996.

\_\_\_\_\_, M. R. F. Adaptação e validação de uma escala de atitudes em relação à Matemática. **Journal of Educational Psychology**, v. 52, n.1, p. 19-24, 1998.

CRONBACH, L. J. **Coefficient alpha and the internal structure of tests**. **Psychometrika**. Williamsburg, v. 16, n. 3, p. 297-334, 1951.

DIAS, A. S. ; VENDRAMINI, C. M. M. Análise fatorial com informação completa de uma prova de Compreensão em Leitura em Estatística. **Psicologia Escolar e Educacional**, v. 12, p. 357-367, 2008.

EMBRETSON, S. E.; STEVEN P. R. **Item Response Theory for Psychologists**. New Jersey: Lawrence Erlbaum Associates, 2000.

KAISER, H. F. A Second Generation Little Jiffy. **Psychometrika**, Williamsburg, v. 35, n.4, p. 401–415, 1970.

LINACRE, J. M. **A user's guide to WINSTEPS® Rasch-Model computer program: Program Manual 3.68**, Beaverton, Oregon, 2009. Disponível em: <http://www.winsteps.com/a/winsteps.pdf>. Acesso em: 22 jun. 2010.

MASTERS, G. N. A Rasch model for partial credit scoring. **Psychometrika**, Williamsburg, v. 47, n.2, p. 149-174, 1982.

NOVAES, M. B. C.; LANÇA, E. C. R.; OLIVEIRA, A. L. P.; MIGUEL, L. F. ; MAZZALI, L. A construção de escalas de atitudes em ciências sócias. In: Anais do XI SEMEAD Seminário em Administração FEA USP, São Paulo, 2007. **Anais...** São Paulo, 2007.

OLIVEIRA, M.H.P.; SILVA, C.B.; KATAOKA, V.Y.; VENDRAINI, C.M.M. Validação de uma escala de autorregulação da aprendizagem de Estatística: um estudo com universitários de cursos tecnológicos de São Paulo. In: Acta Latinoamericana de Matemática Educativa - ALME 23, 23 ed, México, 2010. **Anais...** Mexico: Comité Latinoamericano de Matemática Educativa, 2010, v. 23, p. 01-09.

PASQUALI, L. **Psicometria: teoria dos testes na Psicologia e na educação**. Petrópolis: Vozes, 2003.

RASCH, G. **Probabilistic models for some intelligence and attainment tests**. Chicago: University of Chicago Press, 1980.

REIS, E. **Estatística Multivariada Aplicada**. Lisboa: Sílabo. 1997

SILVEIRA, E. M. **Elaboração e validação de uma escala de atitudes em relação à Estatística para o ensino médio**. Dissertação de Mestrado Acadêmico, Universidade Bandeirante de São Paulo, São Paulo, 2011.

STREINER, D. L. Being inconsistent about consistency: When coefficient alpha does and doesn't matter. **Journal of Personality Assessment**, New York, v. 80, p. 217-222, 2003.