

FIDEDIGNIDADE DE ESCALAS

LUCILA MARIA COSTI SANTAROSA
Profa. Assistente, Dep. de Est.
Espec., Fac. de Educ., UFRGS
Mestre em Educação

Resumo

Descrição de métodos de fidedignidade aplicáveis às escalas de mensuração. Define-se fidedignidade em termos de estabilidade precisão e erro de mensuração. Descreve-se a teoria da fidedignidade, delineando-se o modelo de segregação da variância total em variância verdadeira e variância do erro ($\sigma x^2 = \sigma_{\infty}^2 + \sigma e^2$). Relatam-se os quatro métodos clássicos para a estimação da fidedignidade: 1. aplicações repetidas da mesma escala; 2. aplicações de formas paralelas da escala; 3. divisão da escala em duas metades; 4. método de Kuder-Richardson, apontam-se as vantagens e limitações de cada um concernente à situação específica de escalas. Apresentam-se, ainda, outros métodos: 5. de Horst-Richardson; 6. de Rulon; 7. coeficiente de Hoyt; 8. coeficiente α de Cromback, apontados como mais adequados para estimar a fidedignidade de escalas. Conclui-se exemplificando a utilização do coeficiente de Cromback estimando a fidedignidade de uma escala de mensuração de atitudes tipo Likert.

1 INTRODUÇÃO

Fidedignidade e validade são apontadas como elementos essenciais para avaliar a qualidade de qualquer instrumento de mensuração. Há uma íntima conexão entre os dois conceitos. Ambos se referem a diferentes aspectos do que é essencialmente a mesma coisa, ou seja, a eficiência do instrumento de medida. Para que um instrumento seja válido ele deve ser fidedigno. Se essas duas condições não são satisfeitas, os resultados obtidos também não serão dignos de confiança. Desconhecendo-se o que está sendo medido e com que precisão está sendo mensurado determinada variável, não se terá segurança em apresentar qualquer conclusão a respeito dos resultados alcançados. Validade e fidedignidade são, portanto, componentes fundamentais a serem considerados, em qualquer instrumento de medida, por um pesquisador que deseje desenvolver um trabalho científico.

Nas ciências sociais o problema da precisão da mensuração parece ser mais

sério do que em outras ciências. Na área da educação tem-se observado uma preocupação muito grande neste sentido. Verifica-se que existe maior atenção com referência à tecnologia de instrumentos nas dimensões afetiva e cognitiva, embora nessa última área apareça com maior abundância na literatura pertinente. Consta-se, por exemplo, que os métodos para estimar a fidedignidade se direcionam quase que exclusivamente para testes que envolvem aspectos cognitivos. Na área afetiva tem-se observado maior desenvolvimento relacionado às escalas de mensuração de atitudes, das quais grande parte possui métodos de validação peculiares. Este estudo tem como propósito focalizar o problema da fidedignidade de escalas de medida. Embora a eficiência de uma escala esteja ligada aos dois componentes, validade e fidedignidade, este trabalho se centraliza no último aspecto, tendo presente que a fidedignidade é condição necessária para que um instrumento de mensuração seja válido.

O objetivo deste estudo é o de apresentar métodos de fidedignidade aplicáveis às escalas, apontando a adequacidade dos mesmos face às vantagens e desvantagens que cada um apresenta.

Inicialmente serão apresentados conceitos e teorias sobre fidedignidade. Posteriormente destacar-se-ão métodos para estimar a fidedignidade de escalas, com exemplificações ou direcionamentos que facilitarão seu uso.

2 CONCEITO DE FIDEDIGNIDADE

Segundo Kerlinger (1973) é possível definir fidedignidade de três maneiras:

- a) A primeira definição é focalizada em termos de *estabilidade*. Nessa abordagem, a fidedignidade está ligada à concepção de que sucessivas aplicações do instrumento produzem os mesmos resultados. Bugeda (1974, p. 183), em uma conceituação similar, descreve que “*uma escala é fidedigna na medida em que aplicada várias vezes a uma mesma situação proporciona sempre resultados iguais*”. Vianna (1973) parece realçar também essa dimensão quando coloca que a fidedignidade se refere à estabilidade dos resultados de um instrumento ou ao grau de consistência dos escores.
- b) A segunda abordagem enfoca a *precisão* da mensuração, ou seja, verifica se os resultados obtidos em um instrumento correspondem às medidas verdadeiras da propriedade que está sendo mensurada. O autor acrescenta que a definição de precisão implica na definição de estabilidade. Nunally (1972), também, salienta o aspecto da precisão da mensuração quando define fidedignidade.
- c) A terceira definição refere-se ao *erro de mensuração*. Esse pode ser de dois tipos: sistemáticos e aleatórios. Isto implica na existência de uma variância sistemática e uma variância do erro. Os erros de mensu-

ração são os aleatórios ou randômicos, provenientes de várias fontes.

Nessa dimensão, a fidedignidade pode ser definida como a ausência relativa de erros de mensuração em um instrumento de medida. Guilford (1965) coloca que fidedignidade tem sido definida como a porção da variância total que é a variância verdadeira. Esse autor apresenta um esquema semelhante ao de Kerlinger.

Para melhor compreensão dessa última abordagem necessário se faz descrever os aspectos teóricos sobre fidedignidade.

3 TEORIA DA FIDEDIGNIDADE

A teoria da fidedignidade não representa apenas uma concepção do ponto-de-vista especulativo, mas também do ponto-de-vista prático. Os aspectos que serão abordados auxiliam na compreensão dos vários métodos que posteriormente aparecerão relatados.

Qualquer conjunto de medidas possui uma variância máxima denominada variância total, σx^2 . Essa variância pode ser decomposta em vários elementos. Essa idéia não é nova e pode ser encontrada na maioria dos livros de Estatística Inferencial. Pretende-se delinear aqui o modelo da segregação da variância total em dois tipos de variâncias: a variância verdadeira, σ_{∞}^2 , e a variância do erro, σe^2 .

Dessa forma:

$$\sigma x^2 = \sigma_{\infty}^2 + \sigma e^2 \quad (1)$$

Este raciocínio leva a uma equação simples e básica da teoria:

$$\begin{array}{l} X = X_{\infty} + X_e \\ \text{onde:} \\ X = \text{escore obtido} \\ X_{\infty} = \text{escore verdadeiro} \\ X_e = \text{componente do erro} \end{array} \quad (2)$$

Essa equação expressa que qualquer escore obtido possui dois elementos: o escore verdadeiro e o componente do erro. Dessa forma, a eliminação de qualquer fonte de erro produzida por instrumentos perfeitos e em condições ideais, levará à obtenção do escore verdadeiro. Infelizmente essa perfeição não é atingida e sempre se terá presente o componente do erro. Ocorre ainda que nunca se conhece o verda-

deiro escore e nem o componente do erro. Contudo, é possível estimar a variância do erro. Assim, subtraindo essa variância total se obterá a variância verdadeira. Decorre daí que a fidedignidade vem a ser a proporção da variância verdadeira em relação à variância total dos escores obtidos pelo instrumento de medida.

Para satisfazer essa definição, faz-se mister encontrar a proporção da variância total. Retomando a fórmula (1) tem-se:

$$\frac{\sigma_x^2}{\sigma_x^2} = \frac{\sigma_\infty^2}{\sigma_x^2} + \frac{\sigma_e^2}{\sigma_x^2} = 1,00$$

A fidedignidade é dada pela razão $\sigma_\infty^2 / \sigma_x^2$ ou, em outra forma, por $1 - \sigma_e^2 / \sigma_x^2$.

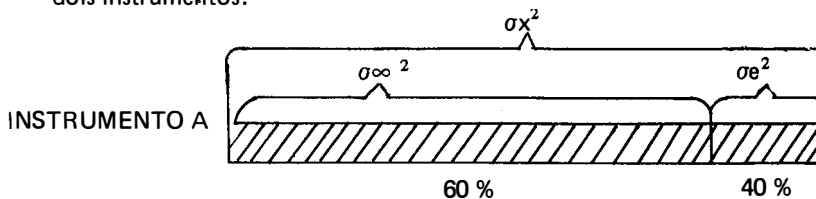
Usando o r_t como representação simbólica do coeficiente de fidedignidade, as duas equações alternativas para expressá-lo, segundo as colocações acima, são:

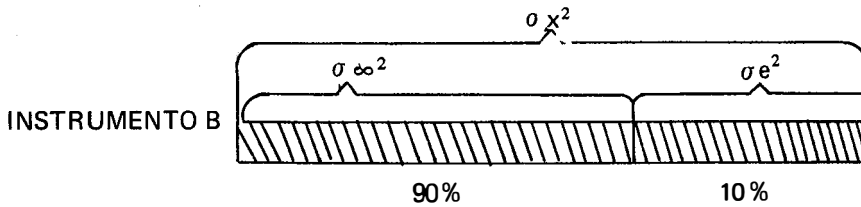
$$\begin{aligned} r_t &= \frac{\sigma_\infty^2}{\sigma_x^2} \\ r_t &= 1 - \frac{\sigma_e^2}{\sigma_x^2} \end{aligned} \quad (3)$$

A primeira fórmula define a fidedignidade como a proporção da variância verdadeira em relação à variância total.

A segunda fórmula permite definir fidedignidade como a proporção da variância do erro em relação a variância total subtraída de 1,00, que indica um coeficiente de fidedignidade perfeito.

Kerlinger (1973) apresenta graficamente exemplos de coeficientes, obtidos de dois instrumentos:





Pelos gráficos acima pode-se detectar que quanto maior a σ_e^2 menor o r_t e vice-versa.

4 MÉTODOS PARA ESTIMAR A FIDEDIGNIDADE

Existe literatura abundante descrevendo os métodos para estimar a fidedignidade, direcionando-se, principalmente, para instrumentos como testes. A maioria dos autores apresenta os quatro métodos mais usados, apontando as vantagens e restrições quanto à utilização de cada um. Embora sejam empregados em diferentes tipos de instrumentos de medida, a ênfase neste trabalho, como já foi referido, será dada às escalas.

Os quatro métodos apontados são:

- a) Aplicações repetidas da mesma escala;
- b) Aplicações de formas paralelas da escala;
- c) Divisão da escala em duas metades;
- d) Método de Kuder-Richardson.

4.1 Aplicações Repetidas da Mesma Escala

Consiste em *construir uma única escala e aplicá-la duas vezes* ao mesmo grupo. O coeficiente de fidedignidade é obtido pela correlação entre os resultados das duas aplicações. Esse método produz, na realidade, um *coeficiente de estabilidade* entre os resultados das duas medidas.

As técnicas de correlação^a mais utilizadas têm sido a linear de Pearson, $\rho_{x y}$, e a ordinal de Spearman, $R_{x y}$, simbolicamente assim representadas:

$$\rho_{x y} \text{ ou } r_t = \sqrt{\frac{n \sum x y - (\sum x)(\sum y)}{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad (4)$$

$$R_{x y} \text{ ou } r_t \equiv \frac{6 \sum d^2}{n(n^2 - 1)} \quad (5)$$

^a Para desenvolvimento e aplicação, ver Santarosa (1975, p. 183-98).

O método da repetição tem sido pouco utilizado devido às limitações que apresenta. Os efeitos da memória entre as respostas dadas em uma e outra situação afetarão o coeficiente de correlação. Se o espaço de tempo transcorrido, entre as duas aplicações for pequeno, possivelmente as respostas dadas na primeira ocasião serão lembradas e repetidas na segunda situação. Isso tenderá a inflacionar o coeficiente de correlação, produzindo uma superestimação. Os efeitos da memória, ainda, são facilitados se a escala apresentar baixo número de itens, o que freqüentemente ocorre com alguns tipos específicos de escala de atitudes (Summers, 1976). Se, por outro lado, a escala contiver muitos itens, a reaplicação pode ser tediosa para o indivíduo. Neste caso, problemas de interesse, motivação e cansaço possivelmente influenciarão as respostas dadas e, conseqüentemente, o coeficiente calculado.

Se o espaço de tempo transcorrido, entre a aplicação e a reaplicação da escala, for longo, provavelmente ocorrerão mudanças que tenderão a subestimar o coeficiente de correlação. Em uma escala de opinião, por exemplo, poderão ocorrer mudanças entre as duas aplicações, o que possivelmente produzirá diferenças significativas entre os resultados. A principal dificuldade nesse método é distinguir os efeitos da memória daqueles devidos à mudança efetiva.

Além das limitações apontadas, é extremamente difícil, senão impossível, manter as mesmas condições nas duas aplicações da escala. Essa estabilidade torna-se mais difícil quando se trata das condições internas do indivíduo.

Em decorrência dessas colocações, a utilização desse método não assegurará a obtenção do coeficiente de fidedignidade de uma escala.

4.2 Aplicações de Formas Paralelas da Escala

Consiste em construir duas escalas similares e aplicá-las ao mesmo grupo de sujeitos. O coeficiente de fidedignidade é obtido pela correlação entre os resultados das duas escalas. Esse método produz o *coeficiente de equivalência* entre as duas formas da escala. As técnicas de correlação mais utilizadas são as já referidas no método anterior, cujas fórmulas (4) e (5) foram mostradas. As vantagens desse método em relação ao anterior referem-se, basicamente, aos efeitos da memória e estabilidade de condições de aplicação, uma vez que ambas as formas podem ser aplicadas no mesmo momento. O problema reside em construir duas formas iguais de escalas. Bugeda (1974), coloca que é relativamente fácil construir formas paralelas de um teste, onde se pode controlar item por item e fazer as substituições necessárias. Todavia, uma forma paralela de uma escala é sempre outra escala e essa, provavelmente, terá um significado diferente e medirá dimensões diversas com outra unidade de medida.

Não obstante, o problema maior parece residir no coeficiente de correlação obtido, que indica serem as escalas comparáveis entre si ou equivalentes. Vianna (1973), esclarece que este grau de relação não reflete, necessariamente, que os ins-

trumentos de medida sejam fidedignos, uma vez que dois instrumentos, quando comparados, podem ter correlação baixa e possuírem fidedignidades elevadas.

Não há pois aplicabilidade prática desse método com referência às escalas.

4.3 Divisão da Escala em Duas Metades

Consiste em *construir uma única escala e aplicá-la somente uma vez ao grupo*. Divide-se a escala em duas metades, correlacionando-se os resultados das duas partes para estimar o coeficiente de fidedignidade. As técnicas de correlação mais empregadas foram apresentadas nos métodos anteriores, correspondendo às fórmulas (4) e (5). O coeficiente obtido traduz a correlação de uma escala que foi reduzida à metade de sua extensão original. Em conseqüência, para estimar a fidedignidade da escala inteira deve-se aplicar a fórmula de predição de Spearman-Brow. Essa fórmula permite calcular nova fidedignidade de um instrumento quando ele é aumentado (ou diminuído) certo número de vezes. Sua expressão geral é:

$$r_{nn} = \frac{n r_t}{1 + (n-1) r_t}$$

onde:

n = ao número de vezes que o instrumento foi aumentado

r_t = é a fidedignidade do instrumento

r_{nn} = é a fidedignidade para o instrumento aumentado n vezes.

(6)

Esse método parte do pressuposto de que ao construir uma escala, por exemplo, de atitudes, cada item do instrumento mede a mesma atitude. Em outras palavras, os itens, individualmente, podem ser considerados como uma medida da mesma atitude. As estimativas da fidedignidade que medem a igualdade de cada item são chamadas, com propriedade, medidas de equivalência. Dessa forma, obtém-se um *coeficiente de equivalência* entre as duas metades da escala.

As vantagens desse método em relação ao primeiro, referem-se aos efeitos da memória e estabilidade das condições, porquanto envolve uma única aplicação do instrumento. Em relação ao segundo método citado, oferece vantagens traduzidas pela facilidade e economia de tempo na construção e aplicação de uma única escala.

Por vários motivos, esse método tem caído em desuso. Uma das principais razões está presente na seleção das duas metades da escala. Alguns pesquisadores correlacionam os itens pares com os ímpares, outros tomam a primeira metade dos itens e correlacionam com a segunda metade, outros ainda, utilizam-se da seleção aleatória dos itens. Cada uma dessas divisões tende a apresentar coeficientes de fidedignidade diferentes. Uma escala dividida em duas metades com n itens terá $(2n)! / 2(n!)^2$ combinações possíveis. Dependendo do procedimento utilizado, as várias combinações resultarão em n coeficientes, alguns superiores e outros inferiores com referência à verdadeira fidedignidade da escala. Conseqüentemente, a divisão das duas metades pode estar longe de ser equivalente.

A participação por itens pares e ímpares tem sido mais empregada e, até mesmo, recomendada. Contudo, teoricamente, qualquer combinação é aceitável.

Bugeda (1974), ainda, acrescenta que o método da partição em duas metades produz sempre coeficientes *conservadores*, ou seja, menores do que realmente deveriam apresentar.

4.4 Método de Kuder-Richardson

As fórmulas de Kuder e Richardson, particularmente as simbolizadas por KR_{20} e KR_{21} têm sido apontadas como extremamente úteis para estimar a fidedignidade. O coeficiente resulta de uma única aplicação da escala. Simbolicamente as fórmulas são assim representadas:

$$KR_{20} \text{ ou } r_t = \frac{K}{K-1} \left(1 - \frac{\sum pq}{\sigma x^2} \right)$$

onde:

X = nº de itens do instrumento

P = proporção dos que concordaram com o item

q = proporção dos que discordaram com o item

$\sum pq$ = somatório das variâncias dos itens

σx^2 = variância dos resultados obtidos pelos indivíduos no instrumento

(7)

$$KR_{21} \text{ ou } r_t = \frac{K}{K-1} \left[1 - \frac{M_x \left(1 - \frac{M_x}{K} \right)}{\sigma_x^2} \right] \quad (8)$$

onde:

M_x = é a média dos resultados obtidos pelos indivíduos no instrumento

Ambas as fórmulas pressupõem que todos os itens sejam paralelos. A KR_{21} pressupõe, ainda, que as questões apresentam o mesmo índice de dificuldade.

A estimação da fidedignidade pelas fórmulas KR é superior à obtida pelos métodos anteriores. Ademais, é uma estimação mais racional de vez que traduz a consistência interna do instrumento. Essas fórmulas foram desenvolvidas para itens dicotômicos, o que limita sua aplicabilidade em instrumentos como escalas, os quais, geralmente, se apresentam com mais de duas dimensões.

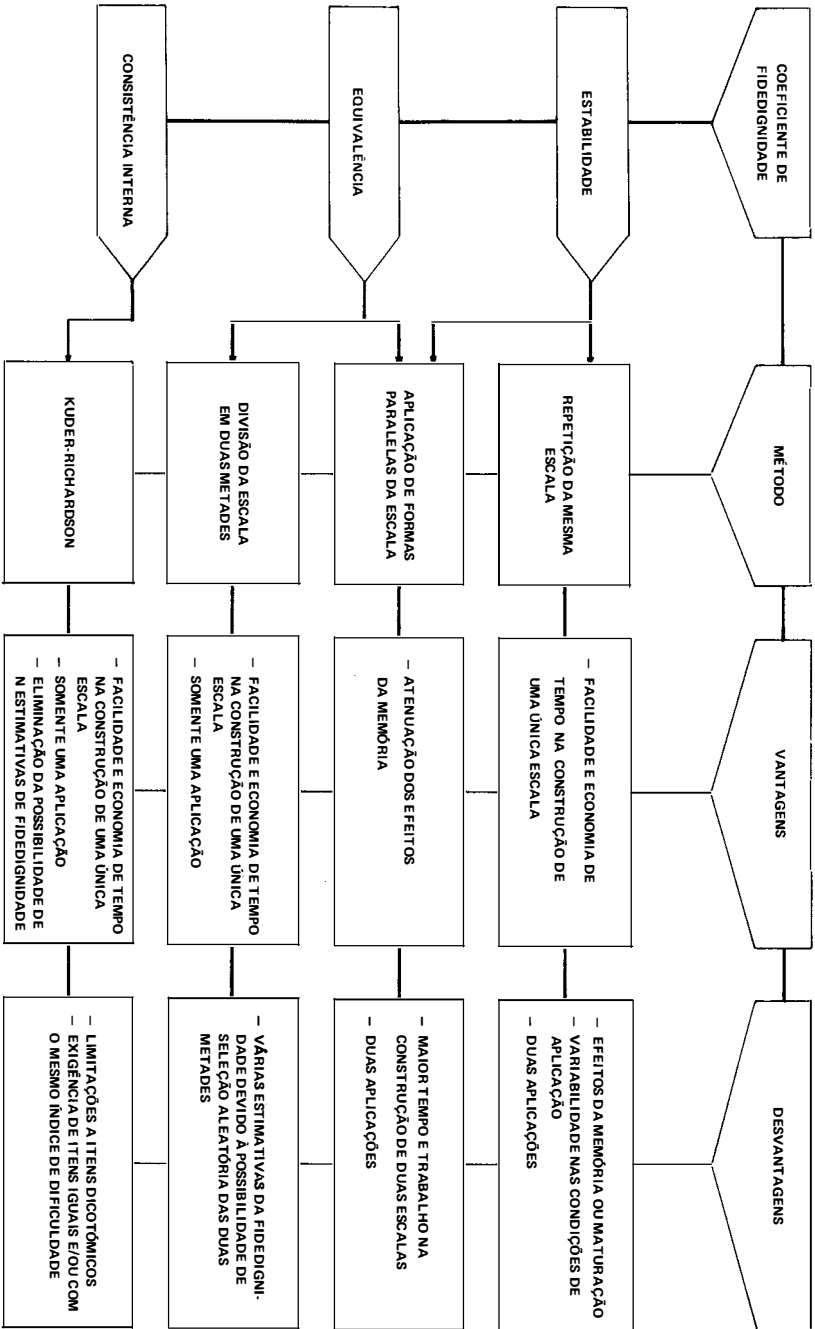
Artifícios têm sido utilizados, com a finalidade de aplicar, principalmente, a KR_{20} quando os itens da escala não são dicotômicos. Uma das formas empregadas, por exemplo, em escalas de mensuração de atitudes, tem sido a de aglutinar as posições extremas de concordância, a partir do ponto neutro, em uma categoria e, as dimensões restantes em outra categoria. Esse procedimento dicotomiza os itens e torna possível, a utilização da KR_{20} . Contudo, não é um processo recomendável.

Os quatro métodos descritos, até este momento, parecem atender mais à instrumentação referente a testes do que à relacionada as escalas. As vantagens de um procedimento em relação ao outro parecem justificar seu emprego, tendo presente as limitações de cada um.

O ESQUEMA I sintetiza os métodos apresentados, dando uma rápida visão do que foi levantado quanto às principais vantagens e desvantagens de cada um deles.

Em detrimento às desvantagens ou à busca de maior utilização, novas técnicas têm sido propostas e se constituem em derivações ou generalizações dos métodos apresentados, especialmente com referência ao de Kuder-Richardson.

ESQUEMA 1 — Vantagens e desvantagens dos métodos de fidedignidade



Para estimar a fidedignidade de escalas serão, ainda, apresentados:

- a) Método de Horst-Richardson;
- b) Método de Rulon;
- c) Coeficiente de Hoyt;
- d) Coeficiente α de Cromback.

4.5 Método de Horst - Richardson

Horst^b (apud Guilford, 1965; Bugada, 1974; Wainerman, 1976) sugere uma modificação da KR_{20} que possibilita a variação dos itens de dificuldade. Essa modificação se baseia no cálculo da variância máxima que é definida como:

$$\sigma_m^2 = 2 \sum_{i=1}^K R_i P_i - M_x (1 + M_x)$$

onde:

σ_m^2 = variância máxima

R_i = é a posição do item i na escala
Os itens são ordenados por dificuldade
O item mais fácil terá ordenação igual a 1

P_i = é a proporção de aceitação do item

M_x = é a média total da escala

(9)

Com essa estimativa da variância máxima, a fórmula denominada Horst-Richardson, será:

$$r_t = \left(\frac{\sigma_x^2 - \sum pq}{\sigma_m^2 - \sum pq} \right) \left(\frac{\sigma_m^2}{\sigma_x^2} \right)$$
(10)

^b Para desenvolvimento e aplicação dessa técnica, ver Guilford (1965, p. 461)

4.6 Método de Rulon

Rulon^C (apud Guilford, 1965) desenvolveu um método para calcular o erro padrão de medida, σe^∞ , em duas metades de um instrumento. Devido à relação entre r_t e σ^∞ , essa abordagem leva a outro tipo de estimativa da fidedignidade. Usualmente a técnica citada é aplicada ao método das metades de um instrumento, mas pode ser expandida ao método das formas paralelas.

Sendo $\sigma e^{2\infty}$ uma medida da variância do erro, outra estimativa da fidedignidade pode ser dada pela fórmula:

$$r_t = 1 - \frac{\sigma e^{2\infty}}{\sigma x^2} \quad (11)$$

onde:

$$\sigma e^{2\infty} = \Sigma d^2 / n$$

Observa-se a similaridade dessa expressão com a fórmula (3) anteriormente apresentada.

O coeficiente obtido pela fórmula acima, conduz a um resultado similar ao obtido pelas técnicas utilizadas no método das duas metades. Entretanto este coeficiente é bem menor do que o obtido pela técnica anteriormente referida. Além disso, possui as mesmas restrições apontadas na descrição desse método. Observa-se, também, que o coeficiente de fidedignidade é dado pelo total dos escores do instrumento e não pelas duas metades. Conseqüentemente a fórmula de profecia de Spearman-Brow não deveria ser aplicada.

Se a fórmula de Rulon for utilizada em formas paralelas de escalas, deve-se realizar a correção pela fórmula (6), substituindo o n por $0,5n$.

^C Para desenvolvimento e aplicação desta técnica, ver Guilford (1965, p. 462).

4.7 Coeficiente de Hoyt

Hoyt^d (apud Wainerman, 1976) faz a estimação do coeficiente de fidedignidade, utilizando a técnica de análise de variância (ANOVA). O autor parte do pressuposto de que o resultado obtido por um indivíduo pode ser dividido em quatro componentes independentes: um associado ao item do instrumento, outro associado ao indivíduo, um terceiro comum a todos os indivíduos e a todos os itens e o último, devido ao erro.

E, também, se utiliza dos pressupostos que embasam o uso do modelo de análise de variância sintetizadas em Santarosa (1976). Cumpridas estas condições, Hoyt estima a fidedignidade do instrumento de medida pela seguinte fórmula:

$$r_t = 1 - \frac{\text{variância devida ao erro}}{\text{variância entre sujeitos}} \quad (12)$$

Exemplifica-se através da matriz abaixo como devem ser dispostos os dados. Em uma das dimensões da matriz encontram-se os sujeitos e em outra os itens do instrumento:

Sujeitos \ Itens	Itens				Σ
	1	2	3	... k	
1					b_1
2					b_2
3					.
.					.
.					.
.					.
.					.
n					b_n
Σ	a1	a2	a3	... ak	$\Sigma a_i = \Sigma b_i$

O **b** representa o escore individual de cada sujeito e o **a** representa a quantidade de respostas corretas em cada um dos itens do instrumento. A soma dos quadrados *entre sujeitos* (SQb), dividida por $n-1$, produzirá a variância entre sujeitos. Da soma dos quadrados *entre itens* (SQa) dividida por $K-1$, obter-se-á a variância

^d Para desenvolvimento e aplicação desta técnica, ver Kerlinger (1973, p. 448).

entre itens. A soma dos quadrados residual ou do erro (SQe), obtida através da soma total dos quadrados menos a soma atribuível aos fatores sistemáticos do indivíduo e do item, dividida por $(n-1)(K-1)$, produzirá a variância devida ao erro.

A variância total (SQt) corresponde ao produto entre respostas corretas e incorretas, dividido pelo total de respostas.

A tabela que segue resume esses procedimentos:

TABELA DA ANOVA

FONTES DE VARIAÇÃO	SQ	g.1	VARIÂNCIA OU QM
Entre itens	SQa	K-1	Q M a
Entre sujeitos	SQb	n-1	Q M b
Devido ao erro	SQe	$(K-1)(n-1)$	Q M e
TOTAL	SQt	n K	Q M t

Neste caso:

$$r_t = 1 - \frac{SM_e}{SM_b} \quad (13)$$

Os resultados obtidos pelo procedimento Hoyt são idênticos aos obtidos pela KR_{20} . Contudo, o enfoque da análise de variância parece ser mais útil para obter estimações da fidedignidade de escalas a partir de itens avaliadas em mais de duas dimensões. Em outras palavras não se restringe unicamente àqueles que apresentam situações dicotômicas com categorias de *concorda* e *discorda*. Selvage (1976), ao descrever a ANOVA como uma estratégia para o cálculo do coeficiente de fidedignidade intraclassas, salienta o problema do atendimento aos pressupostos para utilização de tal modelo. Relata que a ANOVA falha em computar o coeficiente de fidedignidade (que reflete a consistência de ordenações), quando os pressupostos são grosseiramente violados como no caso de ordenações virtualmente iguais. Alerta, também, para os pressupostos referentes à homocedasticidade das variâncias e à normalidade da distribuição. Apresenta uma análise detalhada de coeficientes possíveis de serem obtidos em escalas que se resumem em:

- a) ANOVA obtida com um conjunto de dados originais;
- b) ANOVA com normalidade induzida;
- c) Coeficiente-Finn com pressupostos da distribuição retangular principalmente em escalas de 5 pontos;
- d) Coeficiente Finn com pressupostos da distribuição retangular contínua.

Selvage (1976) discute essas quatro técnicas mais com o objetivo de apresentar alternativas do que indicar preferências. O importante é definir os pressupostos antecipadamente e depois calcular o coeficiente. As quatro possibilidades apontadas não devem ser julgadas para se obter o mais alto coeficiente.

As alternativas e os comentários são feitos com a finalidade de melhorar o cálculo do coeficiente de fidedignidade intraclasses.

O autor faz referência, ainda ao coeficiente oferecido por Olkin & Pratt (apud Selvage, 1976), cuja estimativa da fidedignidade é levemente superior a de Hoyt em relação a pequenas amostras, mas sem efeito em amostras maiores do que trinta ($n > 30$).

4.8 Coeficiente α de Cromback

Uma generalização da $KR_{2.0}$ que obteve maior popularidade e que é mais utilizada para estimar a fidedignidade de escalas, referente à consistência interna, é o coeficiente α de Cromback (Summers, 1976; Nunnally, 1972). O coeficiente α , bem como o coeficiente Hoyt são aplicados a escalas com qualquer tipo de pontuação nos itens, incluindo-se as dicotômicas. Em muitas situações é necessário estimar a fidedignidade de instrumentos, principalmente em escalas de atitudes, cujos itens não apresentam unicamente duas dimensões. Nesses casos, a consistência interna do instrumento não deve ser estimada através da $KR_{2.0}$, que se constitui em um caso especial do coeficiente α , a ser utilizado particularmente para itens dicotômicos. (Stanley, 1971; Nunnally, 1972).

O termo genérico do coeficiente α pode ser assim representado :

$$\alpha = \frac{K}{K-1} \left(\frac{\sum_{i=1}^K \sum_{j=1}^K \sigma_{ij}}{\sigma^2 x^2} \right)$$

onde:

$K = n^{\circ}$ de itens na escala

$\sum_{i=1}^K \sum_{j=1}^K \sigma_{ij} =$ somatório das covariâncias entre todos os itens da escala

$\sigma^2 x^2 =$ variância total da escala

O cálculo desse coeficiente é demorado, por envolver $K(k-1)/2$ possibilidades de diferentes covariâncias entre os itens. Para exemplificar, em uma escala

composta de 20 itens, obter-se-á $20(19)/2 = 190$ covariâncias. O uso de um computador é recomendável para a obtenção da matriz de variâncias dos itens da escala.

Fórmulas mais simplificadas têm sido apresentadas, deduzidas da anterior. Contudo, não se dispensa o uso do computador para facilitar os cálculos.

Assim sendo:

$$\sigma_x^2 = \sum_{i=1}^K \sigma_i^2 + \sum_{i=1}^K \sum_{j=1}^K \sigma_{ij}$$

conseqüentemente:

$$\sum_{i=1}^K \sum_{j=1}^K \sigma_{ij} = \sigma_x^2 - \sum_{i=1}^K \sigma_i^2$$

Aplicando essa equivalência à fórmula (14) tem-se:

$$\alpha = \frac{K}{K-1} \left(\frac{\sigma_x^2 - \sum_{i=1}^K \sigma_i^2}{\sigma_x^2} \text{ ou } \right)$$

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K \sigma_i^2}{\sigma_x^2} \right)$$

onde:

$$K = \text{n}^\circ \text{ de itens da escala} \quad (15)$$

$$\sum_{i=1}^K \sigma_i^2 = \text{somatório das variâncias dos itens da escala}$$

$$\sigma_x^2 = \text{variância total da escala}$$

Stanley (1971) ao referir-se à fórmula acima esclarece que com a substituição das variâncias σ^2 , pelas suas estimativas S^2 , poder-se-ia verificar a consistência interna de um instrumento composto de I itens e aplicado a P pessoas. Nessa forma de estimativa, ou seja, partindo de uma matriz de itens versos indivíduos, a fórmula seria um caso especial do coeficiente α e produziria exatamente o mesmo resultado do

coeficiente Hoyt. O autor acrescenta, ainda, que os procedimentos de Hoyt são talvez mais simples de usar porque empregam as técnicas ANOVA e, conseqüentemente a terminologia é mais familiar.

Pode-se observar também, na fórmula (15), que $\sum_{i=1}^K \sigma_i^2 = \sum_{i=1}^K \frac{p_i \cdot q_i}{KR_{20}}$, para o caso de escalas com itens dicotômicos, o que vem a ser a fórmula KR₂₀, caso especial do coeficiente α , como já foi referido.

Enquanto a técnica de Hoyt parte de uma matriz de *itens x indivíduos*, o coeficiente α parte de uma matriz de *itens x itens*.

Observando a fórmula (15) e tendo-se calculada a matriz de variâncias e covariâncias entre os itens da escala, procede-se da seguinte forma:

- Somam-se os elementos da diagonal da matriz e obtêm-se $\sum_{i=1}^K \sigma_i^2$
- Somam-se os elementos fora da diagonal da matriz com os elementos da diagonal, ou seja, $\sum_{i=1}^K \sigma_i^2$ e obtém-se σ_{x^2} , que é a variância total.
- Substituem-se estes elementos na fórmula (15) e obtêm-se o α como estimativa do coeficiente de fidedignidade da escala.

Para facilitar a compreensão desses procedimentos, será desenvolvido passo a passo o cálculo do coeficiente α .

4.8.1 Exemplo do Cálculo do Coeficiente α

Foi aplicada em 30 (trinta) alunos da disciplina de Medidas Educacionais,^e uma escala tipo Likert com o propósito de medir a atitude dos estudantes em relação ao material audiovisual, utilizado como recurso de aprendizagem em uma das unidades desenvolvidas. Foram selecionados dessa escala, 8 (oito) itens, cuja matriz de variâncias e covariâncias entre os itens vem a seguir.

Matriz de Variâncias e covariâncias^f dos itens da escala.

	1	2	3	4	5	6	7	8
1	0,2023							
2	0,1310	0,7138						
3	0,0368	0,3310	1,1540					
4	0,1356	0,1862	0,2253	0,8092				
5	0,0483	0,4172	0,1310	0,2828	0,1138			
6	0,1195	0,2483	0,3195	0,0943	0,2621	0,6851		
7	0,1287	0,3310	0,0046	0,0299	0,8552	0,2391	1,7057	
8	0,4115	0,3069	0,0529	0,2575	0,4379	0,4391	0,6161	1,1426

^e Curso de Pedagogia da Faculdade de Educação da UFRGS.

^f Obtida através de computador, utilizando-se o pacote SPSSV6

A partir dessa matriz pode-se obter facilmente os elementos que compõem a fórmula (12) do coeficiente α .

Necessita-se computar $\sum_{i=1}^K \sigma_i^2$, que corresponde à soma das variâncias

dos itens da escala, ou seja, os elementos que estão na diagonal da matriz.

$$\sum_{i=1}^K \sigma_i^2 = 0,2023 + 0,7138 + 1,1540 + 0,8092 + 1,1138 + 0,6851 + 1,7057 + 1,4126$$

$$\sum_{i=1}^K \sigma_i^2 = 7,7965$$

É necessário calcular, ainda, a variância total da escala, σ_x^2 , que se obtém somando todos os elementos da diagonal da matriz e os elementos fora da diagonal da matriz. A primeira soma, ou seja, a dos elementos da diagonal, já foi obtida acima, no cômputo de $\sum_{i=1}^K \sigma_i^2$.

Observa-se que a matriz é simétrica, portanto a soma dos elementos fora da diagonal deve ser multiplicada por 2.

Logo:

$$\sigma_x^2 = \sum_{i=1}^K \sigma_i^2 + 2(0,1310 + 0,0368 + 0,1356 + 0,6161)$$

$$\sigma_x^2 = 7,7965 + 2(7,0793)$$

$$\sigma_x^2 = 7,7965 + 14,1586$$

$$\sigma_x^2 = 21,9551$$

Substituindo os elementos da fórmula (12) obtém-se

$$\sigma\alpha = \frac{8}{8-1} \left(1 - \frac{7,7965}{21,9551} \right)$$

$$\sigma\alpha = \frac{8}{7} (1 - 0,36)$$

$$\sigma\alpha = 1,14 \times 0,64$$

$$\sigma\alpha = 0,73$$

O valor de α representa a estimativa do coeficiente de fidedignidade da escala tipo Likert, composta de 8 itens.

Entre essas derivações ou generalizações apresentadas, o coeficiente α parece ser o mais indicado quando se trata de obter estimativas da fidedignidade de instrumentos tipo escala. Contudo, as demais alternativas descritas são aplicáveis desde que se atendam aos pressupostos referentes à sua utilização.

5 REFERÊNCIAS BIBLIOGRÁFICAS

1. BUGEDA, José. **Manual de técnicas de investigación social**. 2. ed. Madrid, Instituto de Estudios Politicos, 1974.
2. GUILFORD, J. P. **Fundamental statistics in psychology and education**. 4. ed. New York, McGraw-Hill, 1965.
3. KERLINGER, Fred N. **Foundations of behavioral research**. New York, Holt, Rinehart and Winston, 1973.
4. NUNNALLY, Jum C. **Educational measurement and evaluation**. 2. ed. New York, McGraw-Hill, 1972.
5. OLKIN, I. & PRATT, J. W. Unbiased estimation of certain correlation coefficients. **Annals of Mathematical Statistics**, **29**:201-11, 1958. Apud SELVAGE, Rob. Comments on the analysis of variance strategy for the computation of intraclass reliability. **Educational and Psychological Measurement**, Durham, **36**(3):605-9, Autumn 1976.
6. SANTAROSA, Lucila Maria Costi. Análise de variância fatorial: um roteiro de aplicação para dois fatores. **Educação**, Brasília, **6**(21) :10-4, jul./set. 1976.
- 7.——. **Módulos de estatística**. Porto Alegre, Ed. A4, 1975.
8. SELVAGE, Rob. Comments on the analysis of variance strategy for the computation of intraclass reliability. **Educational and Psychological Measurement**, Durham, **36**(3) :605-9, Autumn 1976.
9. STANLEY, Julian C. Reliability. In: THORNDIKE, Robert, ed. **Educational measurement**. Washington, American Council on Education, 1971. p.357-442.
10. SUMMERS, Gene. **Medición de actitudes**. México, Trillas, 1976.

11. VIANNA, Heraldo Marelím. *Testes em educação*. São Paulo, Ibrasa, 1973.
12. WAINERMAN, C. H. et alii. *Escalas de medición en ciencias sociales*. Buenos Aires, Nueva Vision, 1976.

Abstract

Descriptions of reliability methods applied to measurement scales. Reliability is defined as stability, precision and error of measurement. Reliability theory is described and a model is proposed that divides the total variance in two parts: the variance due to error and the real variance. The four classical methods to estimate reliability are presented: 1. repeated measurement of the same scale; 2. application of parallel forms of the scale; 3. split-half division of the scale; 4. Kuder-Richardson method. For each one of them the advantages and limitations are appointed regarding the specific use of the scale. Other methods are also presented: 5. Horst-Richardson's; 6. Rulon's; 7. Hoyt's coefficient 8. α coefficient of Cromback. These four methods are appointed as the most adequate to estimate the reliability of a scale. An example of the use of Cromback's coefficient is given at the end of the article, estimating the reliability of a scale that measures attitudes as the one proposed by Likert.