

FIDEDIGNIDADE

Sérgio Fukusima

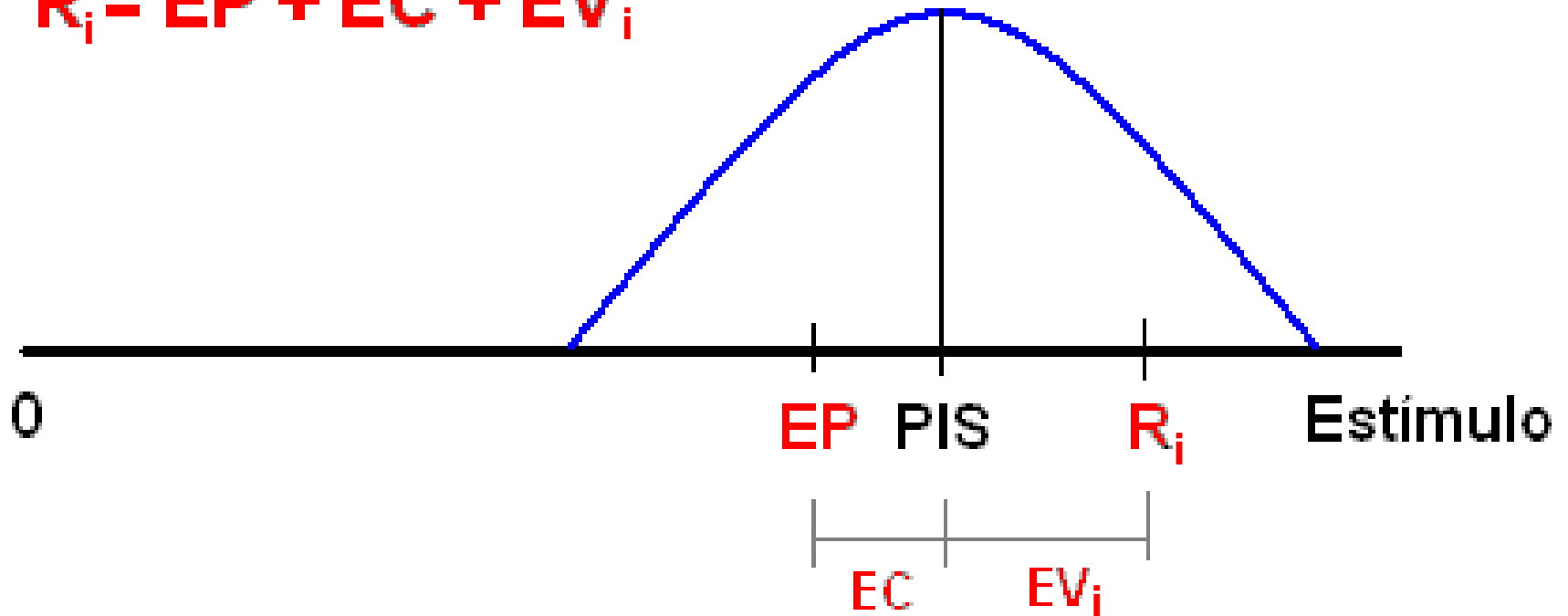
Depto Psicologia / FFCLRP

Ilusão de Müller-Lyer



Acurácia x Precisão

$$R_i = EP + EC + EV_i$$



R_i = Resposta

EP = Estímulo Padrão

PIS = Ponto de Igualdade Subjetiva (média das R_i)

EC = Erro Constante (PIS-EP)

EV_i = Erro Variável -> medida de dispersão -> LD

Erro Médio = média dos valores absolutos de EV_i

Desvio Padrão

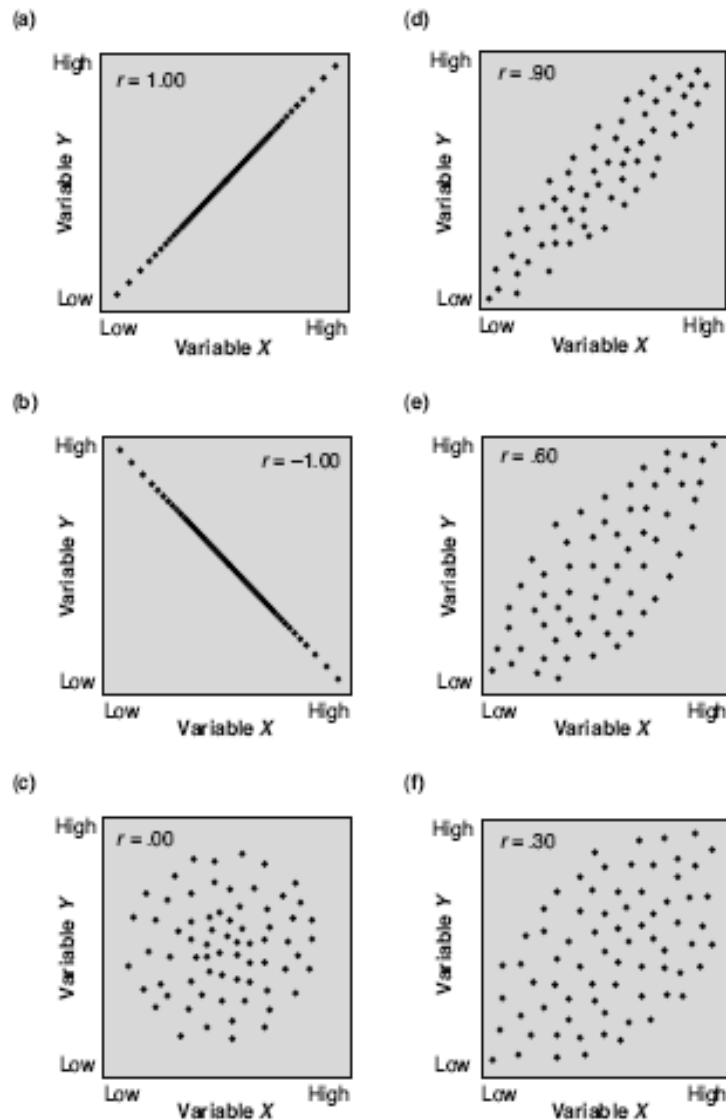


FIGURE 9 Scatterplot of Different Correlation Coefficients.

Source: Hopkins, Kenneth, *Educational and Psychological Measurement and Evaluation*, 8th ©1998. Printed and Electronically reproduced by permission of Pearson Education, Inc., Upper Saddle River, New Jersey.

TABLE 6 Calculating a Pearson Correlation Coefficient

There are different formulas for calculating a Pearson correlation coefficient and we will illustrate one of the simpler ones. For this illustration we will use the test scores we have used before as the X variable, and another set of 20 hypothetical scores as the Y variable. The formula is:

$$r_{xy} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{N\sum X^2 - (\sum X)^2} \sqrt{N\sum Y^2 - (\sum Y)^2}}$$

XY = sum of the XY products

X = sum of X scores

Y = sum of Y scores

X² = sum of squared X scores

Y² = sum of squared Y scores

Test 1 (X)	X ²	Test 2 (Y)	Y ²	(X)(Y)
7	49	8	64	56
8	64	7	49	56
9	81	10	100	90
6	36	5	25	30
7	49	7	49	49
6	36	6	36	36
10	100	9	81	90
8	64	8	64	64
5	25	5	25	25
9	81	9	81	81
9	81	8	64	72
9	81	7	49	63
8	64	7	49	56
4	16	4	16	16
5	25	6	36	30
6	36	7	49	42
7	49	7	49	49
8	64	9	81	72
8	64	8	64	64
7	49	6	36	42
X = 146	X ² = 1,114	Y = 143	Y ² = 1,067	XY = 1,083

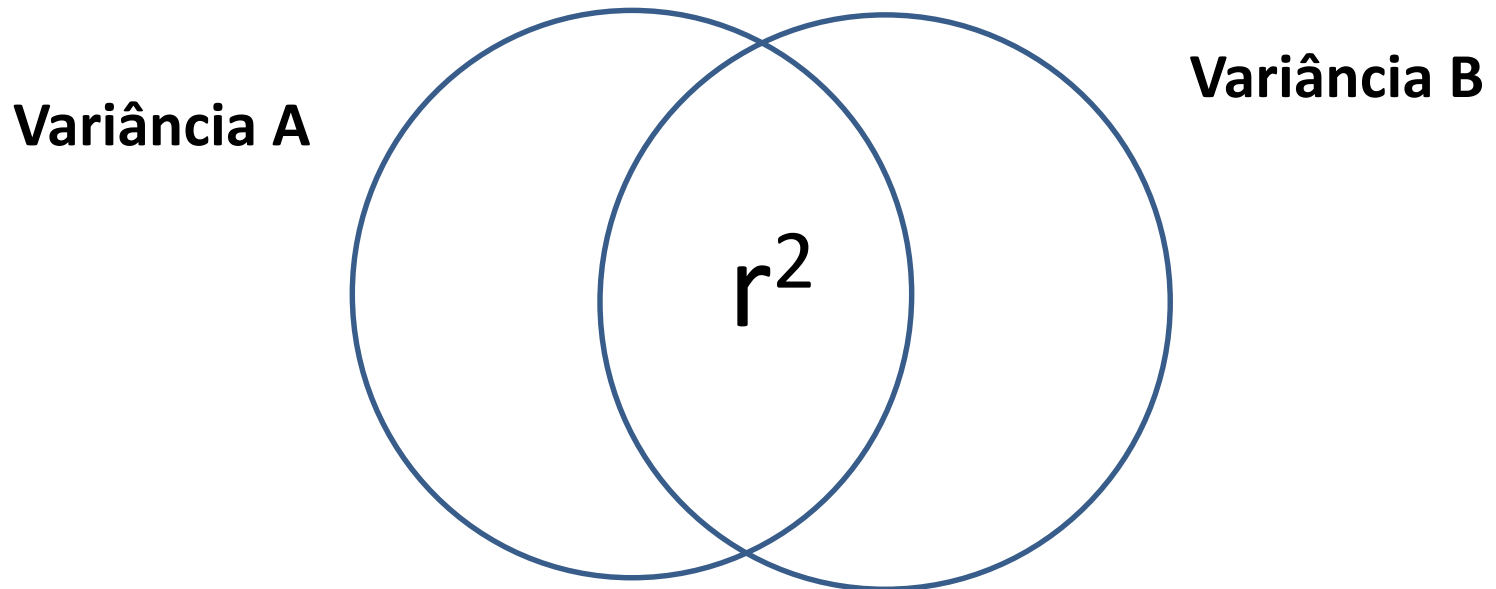
$$r_{xy} = \frac{20(1,083) - (146)(143)}{\sqrt{20(1,114) - (146)^2} \sqrt{20(1,067) - (143)^2}}$$

$$= \frac{21,660 - 20,878}{\sqrt{22,280 - 21,316} \sqrt{21,340 - 20,449}} = \frac{782}{\sqrt{964} \sqrt{891}}$$

$$\frac{782}{(31.048)(29.849)} = 0.843$$

$r^2 =$ coeficiente de determinação

Corresponde a taxa de variância comum entre as variâncias de duas variáveis correlacionadas.

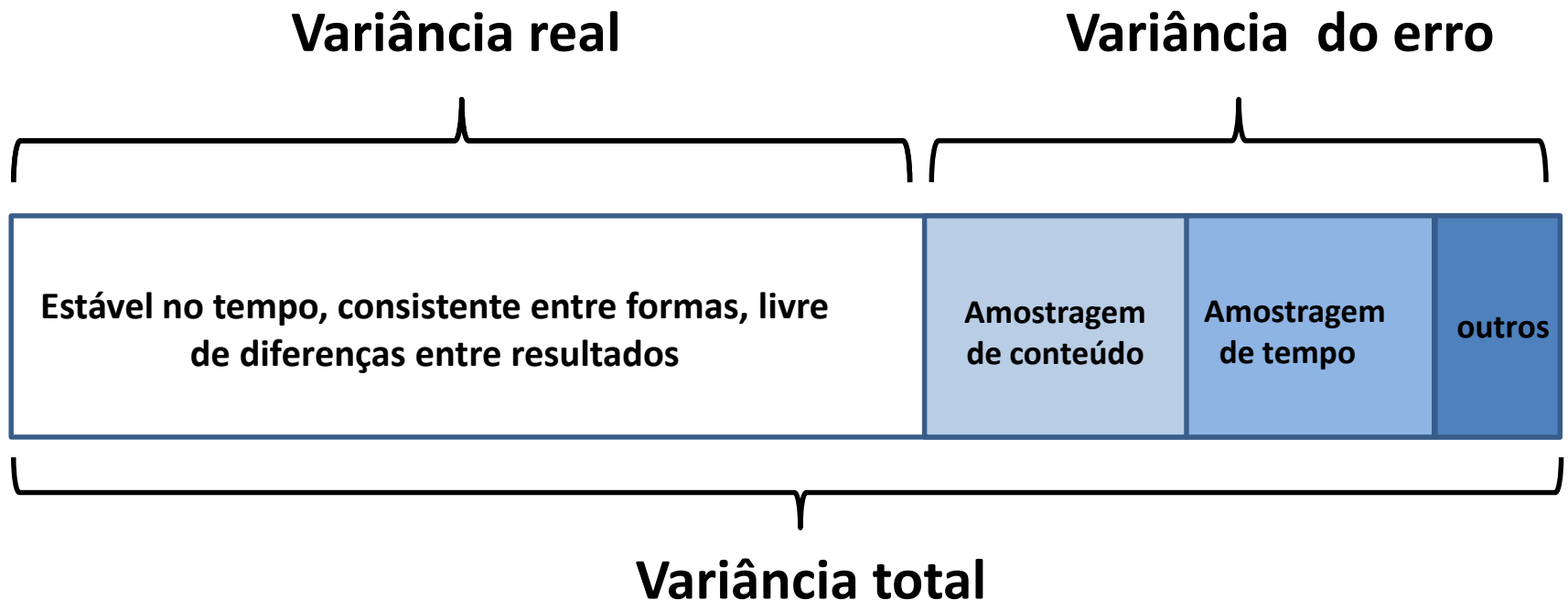


Fidedignidade

(Confiabilidade, Precisão, Consistência)

Psicometria Clássica

$$X_{\text{observado}} = X_{\text{verdadeiro}} + \text{erro}$$



Fontes de Erros

CONSULTA RÁPIDA 4.2

Fontes de erro de mensuração com os coeficientes de fidedignidade típicos usados para estimá-las

Fonte de erro	Tipo de teste propenso a cada fonte de erro	Medidas apropriadas para estimar erros
Diferenças entre avaliadores	Testes avaliados com algum grau de subjetividade	Fidedignidade do avaliador
Erro de amostragem de tempo	Testes de traços ou comportamentos relativamente estáveis	Fidedignidade de teste-reteste (r_{tt}), ou coeficiente de estabilidade
Erro de amostragem de conteúdo	Testes para os quais a consistência de resultados é desejada como um todo	Fidedignidade de forma alternativa (r_{fa}) ou fidedignidade pelo método das metades (split-half)
Inconsistência entre itens	Testes que requerem consistência entre os itens	Fidedignidade pelo método das metades ou medidas mais rígidas de consistência interna, como a fidedignidade de Kuder-Richardson 20 (K-R 20) ou o coeficiente alfa (α)
Inconsistência entre itens e heterogeneidade de conteúdo combinadas	Testes que requerem consistência e homogeneidade entre os itens	Medidas de consistência interna e evidências adicionais de homogeneidade
Erros de amostragem de tempo e conteúdo combinados	Testes que requerem estabilidade e consistência dos resultados como um todo	Fidedignidade de forma alternativa com intervalo

Estimação da Fidedignidade

- **Metódo Teste-Reteste**
(coeficiente de correlação entre os resultados da 1.a e 2.a sessão)
- **Método das Formas Paralelas (alternativas)**
(coeficiente de correlação entre os resultados da 1.a e 2.a forma)
- **Método da Divisão em Metades**
 - fórmula SB
 - fórmula K-R 20
 - alpha de Crombach

Método da Divisão em Metades

Fórmula Spearman-Brown

$$r_{SB} = \frac{nr_{xx}}{1 + (n-1)r_{xx}} \quad (4.5)$$

em que

- r_{SB} = a estimativa de um coeficiente de fidedignidade de Spearman-Brown,
- n = o multiplicador pelo qual a extensão do teste deve aumentar ou diminuir, e
- r_{xx} = o coeficiente de fidedignidade obtido com a extensão original do teste.

Fórmula Kuder-Richardson 20 (K-R 20)

Fórmulas para cálculo da consistência interna

Fórmula Kuder-Richardson 20 (K-R 20)

$$r_{K-R20} = \left(\frac{n}{n-1} \right) \frac{S^2_T - \sum pq}{S^2_T}$$

em que

- n = número de itens do teste
- S^2_T = variância dos escores totais
- $\sum pq$ = soma de p vezes q para cada item do teste
- P = proporção de pessoas que passam em cada item ou o respondem em uma direção específica
- Q = proporção de pessoas que são reprovadas em cada item ou o respondem na direção oposta

A fórmula r_{K-R20} se aplica a testes cujos itens são avaliados como certos ou errados, ou de qualquer outra forma dicotômica, como verdadeiro ou falso, se todos os itens forem formulados de tal modo que o sentido de cada alternativa é uniforme ao longo de todo o teste.

Coefficiente alfa (α) ou alfa de Cronbach

$$\alpha = \left(\frac{n}{n-1} \right) \frac{S^2_T - \sum (S^2_i)}{S^2_T}$$

em que

- n = número de itens do teste
- S^2_T = variância dos escores totais
- $\sum (S^2_i)$ = soma das variâncias dos escores de itens

Esta fórmula do coeficiente alfa é uma variação de item padronizado conhecida como alfa, que usa a correlação média entre itens em vez das variâncias de escores de itens e do escore total, são usadas para testes cujos itens têm múltiplas respostas possíveis (p. ex., concordo totalmente, concordo, discordo e discordo totalmente). Corina (1993) faz uma extensa discussão do significado das fórmulas do coeficiente alfa e os vários fatores que podem afetar seus resultados.

Análise de múltiplas fontes de variância de erro em escores de um único teste

O subteste de Vocabulário da Escala de Inteligência Wechsler para Adultos – Terceira Edição (WAIS-III) consiste em uma série de palavras de dificuldade crescente que são lidas para o testando pelo examinador e simultaneamente apresentadas visualmente em um livreto de estímulos. As definições do testando, fornecidas oralmente, são registradas literalmente e imediatamente pontuadas pelo examinador, usando uma escala de 2, 1 ou 0 pontos, dependendo da qualidade das respostas. Ao avaliarem as respostas, os examinadores são guiados por uma ampla familiaridade com as amostras de respostas fornecidas no manual para cada uma das palavras – nas três níveis de pontuação – bem como pelas definições dicionarizadas de cada palavra (Psychological Corporation, 1997).

O escore total para o subteste de Vocabulário é a soma dos pontos obtidos pelo examinando em todos os itens (palavras). Um escore deste tipo está sujeito a erros de amostragem de tempo e conteúdo, bem como à possibilidade de diferenças entre avaliadores. As estimativas médias de fidedignidade oferecidas no manual do WAIS-III (que pode ser consultado por aqueles que desejarem informações mais detalhadas) para o subteste de Vocabulário são as seguintes:

Fonte de erro/ Tipo de fidedignidade	Coefficiente médio	Proporção e porcentagem(%) da variância de erro
Amostragem de tempo/ estabilidade (teste-reteste)	0,91	$1 - 0,91 = 0,09$ (9%)
Amostragem de conteúdo/ consistência interna	0,93	$1 - 0,93 = 0,07$ (7%)
Diferenças entre avaliadores/ pontuadores	0,95	$1 - 0,95 = 0,05$ (5%)
Variância de erro total medida		$0,9 + 0,7 + 0,5 = 0,21$ (21%)
Variância verdadeira estimada		$1 - 0,21 = 0,79$ (79%)

A partir dos cálculos acima, deve ficar evidente que utilizar uma estimativa de fidedignidade de fonte única para um teste do tipo exemplificado pelo subteste de Vocabulário da WAIS-III produziria uma impressão altamente enganosa da possível quantidade de erro em seus escores. Além disso, este exemplo aponta que, para que escores que estão sujeitos a múltiplas fontes de erro sejam suficientemente confiáveis, as estimativas de fidedignidade para cada fonte, de maneira isolada, precisam ser bastante altas, na faixa de 0,90 ou mais.

Ver Exemplo 1: Aplicando o EMP no texto para uma aplicação específica desta análise de múltiplas fontes de erro e seus efeitos na fidedignidade de um escore do subteste de Vocabulário da WAIS-III.

Erro Padrão da Medida (Mensuração)

$$r_{tt} = 1 - (s_e^2 / s_t^2), \text{ portanto: } s_e^2 = s_t^2 - r_{tt}$$

$$s_e = s_t \sqrt{1 - r_{tt}}$$

Onde:

s_e = erro padrão da medida

s_t = desvio padrão da distribuição

r_{tt} = precisão do teste