

## Sôbre uma questão de nomenclatura estatística

OTÁVIO A. L. MARTINS

Técnico de educação

No mesmo número da *Revista do Serviço Público* em que foi publicado um estudo (ref. 7) no qual assinala a "falta entre nós de uma terminologia estável referente à estatística educacional", tive o prazer de ler um vocabulário de expressões relativas às funções de ligação dos fenômenos, organizado pelo prof. F. Rodrigues da Silveira, do Instituto de Educação (ref. 10).

A publicação simultânea, numa revista não especialmente dedicada ao assunto, de dois estudos, num dos quais se menciona a deficiência de nossa nomenclatura estatística e no outro são apresentados subsídios para a uniformização dessa nomenclatura, vem provar que essas deficiências já são acentuadamente sensíveis em relação ao desenvolvimento entre nós dos estudos desse assunto. Considero isto sintoma de que já é tempo de se tratar sistematicamente dessa uniformização, trabalho cuja responsabilidade, a meu ver, deverá ser exercida por um órgão coletivo, pois as iniciativas individuais desse gênero, sobretudo em nosso meio onde as questões puramente pessoais prevalecem sobre interesses de ordem científica, são sempre sujeitas a controvérsias e a discussões estereis.

Analisando o vocabulário citado e os termos usados em meu artigo, verifiquei, como principal divergência, o emprêgo de *fidedignidade* em vez de *precisão* para traduzir o inglês *reliability*.

Diz o prof. Silveira:

"FIDEDIGNIDADE — Assim chamada a "capacidade de ser medida a *coerência interna* do teste. Este nome, traduzindo o *reliability* usado por Spearman, foi proposto pelo Prof Murilo Braga, que também

"apresentou a expressão *coerência interna* "para significar o grau de constância entre "as repetições das mesmas medidas nos "mesmos indivíduos. Adotamos a denominação nos cursos por nós professados, porquanto é perfeitamente bem adequada ao papel que tais medidas representam. Na verdade, este instrumento merece crédito ou fé ao pesquisador, quer dizer, é *fidedigno* exatamente porque, entre outras vantagens, "é *fiel*. O caso contrário seria absurdo, isto é, "apresentar fidelidade porque nos merece "crédito. Não têm razão, também, alguns "pesquisadores menos avisados, quando "querem emprestar a esse instrumento o "nome de coeficiente de precisão. Aí, além "de absurdo, porque outros coeficientes medem, da mesma forma, o grau de constância, viria trazer confusão com a palavra "precisão, usada e consagrada para significar "uma relação por demais conhecida, no estudo das distribuições de frequência."

Sôbre o assunto, devo dizer que, na minha opinião, é até certo ponto indiferente o uso desta ou daquela palavra para designar determinado conceito científico, desde que a palavra seja aceita por todos e não haja possibilidade de equívoco em seu emprêgo. Disso é exemplo, em estatística, a palavra *regressão*, usada originalmente em estudo sôbre hereditariedade e hoje consagrada no campo da estatística, embora seu sentido corrente só caiba com propriedade quando aplicado ao estudo particular que lhe deu origem (cf. Walker, *Studies in the history of statistical method*, pg. 187, e Yule & Kendall, *An introduction to the theory*

of statistics (11.<sup>a</sup> edição, pg. 207). Fora da especialidade, a expressão *momento de inércia*, por exemplo, tem significação científica completamente diversa da que lhe seria atribuída na linguagem comum, o que não apresenta inconveniente, pois quem usa a expressão sabe exatamente a que se refere.

Entretanto, quando se pretende fixar a nomenclatura de um sector científico, haverá toda vantagem em que os termos escolhidos sejam designados por vocábulos idênticos ou vizinhos.

Analisando-se à luz desses princípios as vantagens e desvantagens de *fidedignidade* e de *precisão* para traduzir *reliability*, parece-me poder-se chegar às seguintes conclusões:

#### FIDEDIGNIDADE:

##### Vantagens

1. Traduz bem a palavra *reliability* tal como é usada em inglês em linguagem corrente.
2. Não é passível de causar certos equívocos, uma vez que não é empregada com outra significação em estatística.

##### Desvantagens

1. É longa e de aspecto rebarbativo.
2. Sua significação não traduz o conceito de *reliability* tal como existe em estatística em geral e na teoria dos testes em particular.
3. O sentido de *fidedignidade*, tal como se usa em linguagem comum, caberia melhor à noção de *validade*, donde a possibilidade de confusão entre conceitos que devem ser nitidamente diferenciados.

#### PRECISÃO

##### Vantagens

1. A significação da palavra, em linguagem corrente, traduz com exatidão a noção de *reliability* tal como existe em estatística em geral e na teoria dos testes em particular.
2. É a mesma usada na linguagem científica brasileira para representar conceito análogo ao de *reliability* quando aplicado a outros instrumentos de medida.
3. É palavra curta, não apresentando inconvenientes quanto à forma.

##### Desvantagens

1. Já é usada, na expressão *índice de precisão*, para designar certo parâmetro da curva normal de probabilidades.

Quando escolhi a nomenclatura adotada em meu artigo, não foi sem madura reflexão que dei preferência a *precisão* sobre *fidedignidade*. Pensadas as vantagens e desvantagens de ambas, a única desvantagem que vi na palavra *precisão* não me pareceu suficiente para eliminá-la. Realmente, quando se falar no *índice de precisão de um teste*, não haverá probabilidade de ser confundido com a noção de *índice de precisão de uma distribuição de frequência*. São aliás numerosos os exemplos em ciência de ter a mesma palavra sentidos diferentes quando relacionadas a diferentes elementos. Isto se dá em estatística com a própria palavra *reliability*. Como exemplo, na frase: "The reliability of estimated test reliabilities", usada por Guilford (*Psychometric methods*, pg. 420), há sentidos diversos, embora vizinhos, atribuídos à mesma palavra. A frase poderia ser traduzida: "A *precisão* (numérica) dos elementos que medem a *precisão* (ou *fidedignidade*) dos testes."

Quanto aos argumentos que figuram sob o verbete *fidedignidade* do vocabulário de autoria do ilustrado professor do Instituto de Educação, não me parecem suficientemente convincentes, pelos motivos que exponho a seguir.

Si *fidedignidade* é a *qualidade daquilo que merece fé ou confiança*, a palavra, quando referente a testes, se aplicaria melhor à noção de validade que à de *reliability*. Um teste poderá ter, com efeito, alta *reliability* e entretanto sua *validade* ser muito baixa; neste caso, a aceitar-se a nomenclatura advogada no vocabulário, o teste seria *fidedigno* mas seus resultados *não seriam merecedores de fé ou confiança*, pois em vista da baixa validade do teste não estaria ele medindo com exatidão aquilo que se pretendia medir, o que seria um absurdo dentro da acepção vulgar dos vocábulos usados. Este argumento me parece suficiente para neutralizar todos os que foram adiantados para justificar o uso da palavra, baseados em sua significação em linguagem corrente.

Contrariamente a isto, julgo que a significação vulgar da palavra *precisão* corresponde exatamente àquilo que, em inglês, é representado pela palavra *reliability*, quando usada em estatística, quer de um modo geral, quer em referência aos

testes. Não julgaria isto sujeito a contestação si não tivesse sido contestado. Analisemos entretanto um caso concreto para justificar meu ponto de vista. A uma pergunta sobre a data do nascimento de Dante, suponhamos terem sido dadas as duas seguintes respostas: (1) a 11 de julho de 1272; (2) entre 1260 e 1270. A primeira é muito mais *precisa* que a segunda; entretanto, embora mais *precisa* é menos *fidedigna* (ou menos merecedora de fé) que a outra, pois sabe-se que Dante nasceu em meados de maio de 1265. Ora, o conceito de *reliability* de um teste corresponde exatamente à noção de *precisão* que caracteriza a primeira resposta em relação à segunda, isto é, diz respeito à exiguidade da amplitude de variação de seus resultados, sendo ao conceito estranha a noção de *verdade* (ou de adequação aos fins visados) desses resultados, aspecto este que se considera na noção de *validade dos testes*.

Como prova de que a noção de *reliability*, em estatística e na teoria dos testes, tem realmente o sentido indicado, transcrevo a seguir alguns trechos de livros de referência e de escritores americanos que me parecem de autoridade no assunto.

"The *precision* or reliability of an estimate "is dependent on the variability of its sampling distribution" (Lindquist, *Statistical analysis in educational research*, pg. 10. O grifo é do autor).

"**sampling reliability:** The accuracy with which a given sample characterizes the universe from which it is drawn; the stability of a statistic from sample to sample. . . . Also called reliability, 2. Not to be confused with reliability, 1." (Kurtz e Edgerton, *Statistical dictionary of terms and symbols*, pg. 150).

"**reliability: 1.** The accuracy with which a measuring device measures whatever it does measure. . . Reliability is usually measured in terms of the coefficient of reliability, the index of reliability, the standard error of estimate, or the standard error of measurement. Not to be confused with sampling reliability or with validity. . . 2. Same as sampling reliability." (Kurtz e Edgerton, op. cit., pg. 146).

"A good test is reliable, which means that "the results obtained from it are accurate." (Orleans, *Measurement in education*, pg. "50).

"... the question of reliability is that of "how accurately a test measures the thing "which it does measure." (Kelley, *Interpretation of educational measurements*, "pg. 14).

"**reliability coefficient. . . Statistics.** A measure of the accuracy of a test or measuring instrument, obtained by measuring "the same individuals twice and computing the correlation of the two sets of "measures." (Webster's *New International Dictionary*, pg. 2.104).

"**2. Reliability.** The second important question about a test is the consistency or "accuracy with which a test measures what "it does measure." (Rinsland, *Constructing tests and grading*, pg. 14).

"The concept of reliability refers to the "chance variation in any measurement." (Thurstone, *The reliability and validity of tests*, pg. 8).

"The reliability coefficient of a test is an "index of the accuracy with which a test "measures whatever it measures." (University of Chicago, *Manual of examination methods*, pg. 11).


"The reliability of a test. . . is a measure "of the degree of success in overcoming "variable errors in the test scores." (Greene e Jorgensen, *The use and interpretation of elementary school tests*, pg. 67).

As duas primeiras transcrições se referem ao conceito de *reliability* em sua acepção geral em estatística e as demais ao conceito de *reliability* relacionado aos testes ou outros instrumentos de medida educacional. Julgo-as suficientemente claras para dispensar uma análise interpretativa tendente a demonstrar que o conceito corresponde ao que é traduzido pela palavra *precisão* referente a qualquer processo de medida.

Isto quanto à tradução da palavra *reliability*. Quanto aos seus derivados (coeficiente de precisão, índice de precisão e erro de precisão), sendo eles suscetíveis de definição matemática, não me parecem sujeitos às dúvidas que poderiam ocorrer em relação a um conceito mais geral. Poder-se-á dizer que o coeficiente de precisão e o índice de precisão não constituem medida conveniente da precisão de um teste e que portanto seria vantajosa outra designação mais significativa. Concordo inteiramente com a primeira parte do argumento, e o estudo que publiquei ressaltou especialmente este ponto. Quanto à conclusão, aceita-la-ia de bom grado no caso de se tratar de elementos a serem agora batisados, caso em que adotaria *coeficiente de constância* e *coeficiente virtual* (ou *teórico*) de preferência a *coeficiente de precisão* e a *índice de precisão*. Entretanto, a literatura americana domina praticamente o campo da estatística educacional e já fixou as designações de *coefficient of reliability* e de *index of reliability* (embora a segunda tenha sido adotada por acaso, como assinalou Walker, *op. cit.*, pg. 118). Uma vez que a noção de *reliability* deve ser, como penso, traduzida por *precisão*, as vantagens de se conservar a analogia com a nomenclatura americana me parecem sobrepujar de muito o inconveniente apontado, que existe aliás com maior força em relação à expressão *coeficiente de validade*, que entretanto deve ser conservado pelos mesmos motivos e que não parece ter sido sujeito a controvérsias.

#### REFERÊNCIAS

1. BRAGA, MURILO, *Validade e fidedignidade nos testes coletivos de inteligência*, Rio de Janeiro: 1938.
2. GREENE, H. A. e JORGENSEN, A. N., *The use and interpretation of elementary school tests*, New York: Longmans, Green & Co., 1938.
3. GUILFORD, J. P., *Psychometric methods*, New York: McGraw-Hill Book Co., 1936.
4. KELLEY, T. L., *Interpretation of educational measurements*, Yonkers-on-Hudson, N. Y.: World Book Co., 1927.
5. KURTZ, A. K. e EDGERTON, H. A., *Statistical dictionary of terms and symbols*, New York: John Wiley & Sons, 1939.
6. LINDQUIST, E. F., *Statistical analysis in educational research*, Boston: Houghton-Miffling Co., 1940.
7. MARTINS, OTÁVIO A. L., "Medidas de precisão e de validade dos testes". *Revista do Serviço Público*. Vol. IV. N. 3, dezembro de 1940, pgs. 42 a 54.
8. ORLEANS, J. S., *Measurement in education*, New York: Thomas Nelson & Sons, 1938.
9. RINSLAND, HENRY D., *Constructing tests and grading*, New York: Prentice-Hall, 1938.
10. SILVEIRA, F. RODRIGUES DA, "Coeficientes, índices e equações nas funções de ligação dos fenômenos", *Revista do Serviço Público*, Vol. IV, N. 3, dezembro de 1940, pgs. 29 a 34.
11. The Technical Staff of The Board of Examinations, University of Chicago, *Manual of examination methods*, Chicago: The University of Chicago, 1937.
12. THURSTONE, LOUIS L., *The reliability and validity of tests*, Ann Arbor, Mich.: Edwards Brothers, Inc., 1939.
13. WALKER, HELEN M., *Studies in the history of statistical method*, Baltimore: The Williams & Wilkins Co., 1931.
14. *Webster's New International Dictionary of the English Language*, Second Edition, Unabridged, Springfield, Mass.: G. & C. Merriam Co., 1939.
15. YULE, G. UDNY e KENDALL, M. G., *An introduction to the theory of statistics*, (11.<sup>a</sup> edição), Londres: Ch. Griffin & Co., 1937.


**TRABALHE COM ENTUSIASMO: O BOM HUMOR**  
**AJUDA A PRODUÇÃO**
