

ERROS EM TESTES E MEDIDAS

* RENAN M. F. SAMPEDRO

1. INTRODUÇÃO

Parece impraticável, ou demasiadamente idealista, esperar que pesquisas envolvendo testes e medidas sejam apresentadas totalmente livres de erros. O maior problema não é eliminar o erro, ao contrário, é evitar o erro que possa resultar em má interpretação dos resultados. Em uma consideração mais ampla, é possível classificar erros em pesquisa como sendo resultantes de : (a) falta de lógica; (b) generalizações sem justificativa; e (c) falta de exatidão nas medidas. Cada um destes aspectos têm recebido consideração de alguma forma, por parte de vários autores (BROZEK & ALEXANDER, 1947; CURETON, WICKENS & ELDER, 1935).

O propósito da presente discussão é o de, primeiramente, considerar o erro de medidas, sua avaliação e seu controle.

2. CONSIDERAÇÕES SOBRE ERROS DE MEDIDAS

Quando a pesquisa é realizada com seres humanos, o problema de erros em testes e medidas complica pelo fato de que diferenças individuais certamente estarão presentes na função a ser medida. Qual é a força muscular do homem? Uma afirmação perfeita e verdadeira sobre esta quantidade é impossível, muito embora, quase sempre seja possível fazer-se afirmações quantitativas significantes sobre esta função, considerando-se a possibilidade de existirem fontes para prováveis erros. Força, está relacionada com sexo, idade, e até mesmo com o ambiente; torna-se necessário, portanto, considerar-se um grupo amostral homogêneo para este tipo de estudo. Pode ser significativo, por exemplo, afirmar-se que a força da mão de universitários,

* PROFESSOR TITULAR DO DEPARTAMENTO DE MÉTODOS E TÉCNICAS DESPORTIVAS/ CEFD/UFSM.

idades entre 18 e 20 anos, é de 43 Kg com um desvio padrão de 6.5 Kg, ou simplesmente que a força de mão de um certo indivíduo é de 36.4 Kg.

No entanto, afirmações deste tipo, descrevendo o resultado de medidas, são essencialmente aproximações, e não os resultados verdadeiros. Obviamente, estes resultados podem servir quase satisfatoriamente como substitutos dos resultados reais, desde que as discrepâncias entre as aproximações e a realidade sejam conhecidas. Considerando-se o indivíduo cujo escore foi de 36.4 Kg, ele recebeu treinamento e instruções com o dinamômetro, e sabia como usá-lo; o dinamômetro foi ajustado para o tamanho da sua mão, e ele foi muito bem motivado, de forma que empregou o máximo de força na tentativa. Não é possível avaliar se a medida foi realizada em condições ideais, no entanto, tudo foi feito de forma a buscar estas condições.

Mesmo que todas as precauções tenham sido tomadas, ainda assim existe um grande potencial para erros que precisa ser considerado. Qual a consistência do dinamômetro? Uma resposta razoavelmente satisfatória a esta questão pode estar assegurada pelo fato de repetidamente testar-se o dinamômetro, com um peso de 35 Kg e observando-se o resultado. A construção do equipamento foi de tal forma, que a consistência esperada deveria ser muito alta, o que na realidade vamos admitir tenha acontecido neste caso, exemplificando com um desvio padrão das discrepâncias de 0.4 Kg. Qual a consistência do indivíduo? A média de 10 medidas, bem separadas para evitar a fadiga, com o respectivo desvio padrão, pode ser a resposta adequada. No entanto, é preciso considerar, também, uma estimativa de erro para o escore de qualquer indivíduo de um grupo, o que pode ocorrer em função de que os escores de um sujeito podem também variar de um momento para o outro, ou de um dia para o outro, o que é muito comum.

Considerando-se que os resultados podem, nestes casos, em certa extensão ser falhos, o resultado a ser usado (X) deverá, então, ser composto pelo resultado real obtido (T), somado a um erro sistemático (E), constante, de direção fixa, seja positiva ou negativa, e um erro variável (e), que varia aleatoriamente em magnitude e direção quando um certo número de indivíduos está sendo medido, ou quando um indivíduo é medido repetidas vezes: $X = T + E + e$.

a) Erro Sistemático (Constante E) Este erro pode surgir de inúmeras circunstâncias. O dinamômetro tem, por exemplo, uma escala colocada pelo fabricante, marcada em Kg. Mesmo assim, a validade absoluta ou relativa das marcas não são garantidas. O Pesquisador tem a responsabilidade ética de calibrar o instrumento e determinar quais os fatores de correção que deverão ser usados.

As exigências de precisão na calibragem dependem do propósito para os quais os dados serão usados. HENRY, CURETON, HEWIT & JARRETT (1945) dizem que a precisão deve ser suficientemente alta de forma a não se tornar um fator limitante na interpretação dos dados, mas não é necessário uma precisão maior do que a exigida pela própria natureza do problema. Em alguns problemas, o erro sistemático não interfere, desde que ele permaneça constante; por outro lado, não existem normas adequadas se a magnitude do erro sistemático é desconhecida. O ponto crucial é que os pesquisadores tenham a segurança de que seus instrumentos e métodos possuam precisão suficiente para resolver seus problemas específicos de pesquisa. Pesquisadores em Educação Física apresentam tendências em ser descuidados nestes aspectos, em suas pesquisas e investigações (MATHEWS, 1980).

O erro sistemático pode aparecer sob vários prismas, inclusive sob o aspecto psicológico. Tarefas simples como a de administrar um questionário a um indivíduo, ou a um grupo de indivíduos, podem originar erros até mesmo pela simples maneira de apresentar o questionário. O grau de treinamento, e a fadiga, relacionados a sensação subjetiva, também podem ser fontes de erros em medidas baseadas na performance humana (MCARDLE, KATCH & KATCH, 1986). O grau de motivação também é uma potente fonte de erro do tipo psicológico.

Fatores ambientais e fisiológicos também são considerados como fontes potenciais de erro sistemático (ASTRAND & RODAHL, 1986). Por exemplo, o metabolismo basal pode ser influenciado pelo tipo e pela quantidade de alimentos ingeridos; frequência cardíaca e pressão arterial são afetadas pela temperatura ambiente. A velocidade de um corredor pode ser influenciada pela velocidade do vento.

Seria quase impossível listar aqui todas as fontes potenciais de erro sistemático, o importante não é saber se o erro existe ou não, mas sim qual a magnitude do erro, e como este erro pode ter seu efeito reduzido a um grau desejável, ou ser considerado de alguma outra maneira, de forma a não intervir nos resultados. Existem situações

em que o próprio pesquisador pode ser visto como uma fonte de erro, devido a sua técnica ou acuidade sensorial, influenciar nos resultados (BROOKS & FAHEY, 1984).

b) Erro Variável ou randômico ou aleatório (e). Mesmo um instrumento altamente sofisticado pode falhar na consistência de seus resultados. O dinamômetro usado como exemplo anteriormente, quando testado com um peso padronizado, apresentou variações irregulares. A força dos sujeitos submetidos aos testes também varia de forma irregular, próximo a um valor médio, quando as medidas são repetidas.

Estes erros variáveis também podem aparecer através de uma gama muito grande de possibilidades. Todos os instrumentos e métodos de medida praticamente apresentam alguma forma intrínseca de erro deste tipo. Alguns exemplos mais comuns são citados aqui. Um examinador dificilmente consegue aplicar um teste duas vezes, exatamente da mesma maneira; condições ambientais variam de forma irregular, de tempo em tempo; as interpolações entre dois números em uma escala, etc. Existe apenas uma forma de reduzir o erro variável, que é reduzindo sua causa. Na prática, é possível se fazer isto ocasionalmente, e adequadamente, através do uso do senso comum. No entanto, é usualmente aconselhável que os pesquisadores familiarizem-se com os princípios dos métodos ou dos instrumentos que irão usar, através da literatura ou da orientação de especialistas (SAFRIT, 1980).

Os erros variáveis são caracterizados pela sua natureza aleatória. Eles variam em magnitude de forma irregular, e flutuam do positivo para o negativo, de forma assistemática. Há uma tendência de aproximação a uma distribuição estatística normal, ou a chamada distribuição de Gaussian (HENRY et al., 1945). A magnitude total do erro variável é o resultado da combinação de todos os erros randômicos ou aleatórios.

Torna-se, portanto, óbvio que se inúmeros erros de mesma magnitude ou de magnitude similar são encontrados nas medidas, a simples eliminação de um único erro não trará benefício algum na correção do erro total. Mesmo quando existem poucos erros, a eliminação do menor deles em termos de magnitude, também não terá nenhum efeito se os seus valores são muito diferentes. Assim sendo, torna-se aparente que a única forma prática de aumentar a precisão das medidas é a eliminação da maior quantidade possível de erros, ou a redução da

magnitude dos erros de maior valor.

Como salientamos anteriormente, o problema não é simplesmente o de eliminar os erros mas, sim, evitar que erros que possam ser suprimidos venham a contribuir para distorções muito grandes ao final da investigação. Algumas vezes é possível evitá-los quantitativamente através de técnicas estatísticas; outras vezes, torna-se praticamente impossível evitá-los, ficando estes erros portanto, sob a responsabilidade do pesquisador que deve ter uma boa compreensão do problema, conhecer as técnicas adotadas, e saber usar sua sensibilidade para evitar o surgimento destes tipos de erros em seus testes e medidas.

c) Redução do Erro através da Repetição de Medidas. No caso de erro variável ou randômico, a precisão das medidas pode ser aumentada pela repetição das mesmas. Isto, estatisticamente, aumenta a precisão de forma diretamente proporcional à raiz quadrada do número de repetição de cada medida. No entanto, a repetição de medidas não aumenta necessariamente a precisão na determinação da média dos escores de um grupo de indivíduos. Neste caso, aumentando o número de indivíduos (n), ao invés de aumentar o número de repetições da medida, parece ser a única forma prática de reduzir o erro. É importante não confundir o número de indivíduos de uma amostra (n) com o número de testes ou medidas feitas em um único indivíduo que também pode ser representado pelo mesmo símbolo (n). Esta nomenclatura não é universal, o que pode criar alguma confusão para os pesquisadores menos experientes.

O uso de técnicas como a análise de variância, algumas vezes, é útil na determinação da quantidade de erro das várias partes de uma investigação ou processo de testagem que contribuem mais efetivamente para o erro total. É preciso notar, no entanto, que o efeito relativo de um erro na média de um grupo ocorre apenas em função do número de indivíduos na amostra.

d) Influência dos Erros na Correlação. O erro constante, ou sistemático, não é importante em análises de correlação, desde que ele influencie todos os escores proporcionalmente. Se a quantidade de erro for diferente em diferentes partes da escala de um instrumento, por exemplo, afeta a normalidade da distribuição de frequência é afetada. embora, ordinariamente, isto não seja um problema sério, pode facilmente ser corrigido através da utilização de escalas apro-

priadas de normalização (exemplo, a transformação em T-escores).

Os erros variáveis ou randômicos, já apresentam uma certa influência. Como mencionamos anteriormente, os escores reais, ou verdadeiros, dificilmente são obtidos através de testes ou medidas; um determinado escore, portanto, consiste do escore obtido através da medida mais o erro variável. Em alguns tipos de problema envolvendo a simples predição de um resultado contra um resultado de performance, desde que o erro variável seja de pequena magnitude e menor do que a variação das medidas de um mesmo indivíduo, não há com o que se preocupar. Estas variações são consideradas inerentes à situação, e devem ser aceitas. Dependendo do caso, todos os erros variáveis devem ser considerados como parte normal da situação. Nestes casos, estaremos medindo apenas se a predição dos escores, comparada com os escores obtidos, foi ou não correta.

d) Seleção dos Dados. A escolha de dados de forma tendenciosa às vezes ocorre inadvertidamente. Com boa fé, mas em total ignorância científica, o pesquisador observa que um determinado conjunto de dados apresenta resultados que fogem as suas expectativas e procura, então, justificativas que lhe permitam desqualificar aqueles dados. Obviamente, seu juízo está primeiramente condicionado ao fato de que ele não encontrou o que queria e, por isso, eliminou alguns dados. A regra a ser seguida, nestes casos, é a de sempre tomar decisões de eliminar dados ou não, antes de olhar os escores. Indivíduos doentes, lesionados, extremamente obesos, ou atípicos da amostra em geral, devem ser eliminados antes que os dados sejam analisados pelo investigador.

3. SUGESTÕES PARA REDUÇÃO E CONTROLE DOS ERROS

Para a redução e o controle dos erros, podemos sugerir alguns procedimentos que, se levados em consideração, poderão contribuir para minimizá-los.

a) Fatores Psicológicos. Instruções e procedimentos devem ser escritos, e seguidos minuciosamente. O teste, ou antecipações dos resultados, não devem ser discutidos na presença dos sujeitos, além das explicações gerais fornecidas a todos. A motivação deve ser, ao máximo possível, padronizada para todos os sujeitos. Da mesma forma, o aquecimento, quando parte do protocolo de testagem, deve ser padronizado. Deve-se evitar, ao máximo, distrações tais como

ruidos, conversas, etc., que podem influenciar nos resultados. Em exercícios de exaustão, os fatores psicológicos são particularmente importantes.

b) Fatores Fisiológicos. Precauções devem ser tomadas a fim de evitar os problemas relacionados à fadiga, mesmo nos testes mais simples (MATHEWS, 1980). Testes que exigem exercícios intensos, devem ser separados por um intervalo adequado de descanso. Hora do dia, alimentação, temperatura ambiente e condições de ventilação devem ser padronizados, ou controlados pelo "design" experimental de alguma forma. É interessante salientar também, que fatores psíquicos, tais como exitação e ansiedade, influenciam processos fisiológicos, particularmente na frequência cardíaca e na pressão arterial (ASTRAND & RODAHL, 1986).

c) Fatores que Afetam a Performance. A possibilidade do aparecimento de uma grande quantidade de erros deve ser mantida na mente de um pesquisador, quando medindo a performance humana. O tipo de piso de uma pista de atletismo, o tipo de calçado usado por um atleta, e outras variáveis, podem afetar, por exemplo, corridas ou saltos. Na medida da força, o ângulo das articulações deve ser padronizado, etc.

d) Fatores Instrumentais. Certos tipos de instrumentos de medidas são notoriamente inconstantes, e devem ser evitados pelos pesquisadores. Em caso de necessidades especiais exigirem o uso destes instrumentos, os mesmos devem ser frequentemente comparados e calibrados contra alguma outra forma similar, confiável de medida. Todos os instrumentos devem ser examinados com regularidade por técnicos especializados, para evitar correções.

4. CONCLUSÕES

Para concluir, deixamos aqui algumas atitudes que são esperadas de parte dos pesquisadores, principalmente em Educação Física, relacionados com erros de pesquisa.

Quando a magnitude de um erro específico é desconhecida ou mensurável, todo e qualquer esforço deve ser feito para assegurar que seu efeito nos resultados finais sejam insignificantes. Potenciais fontes de erro devem ser antecipadas e eliminadas, ou controladas de alguma forma. Vale a pena padronizar todos os procedimentos, cuidadosamente, entendendo-se que isto pode evitar ou controlar vários

tipos de erros que não podem ser reconhecidos ou antecipados.

Poderá acontecer que o erro permaneça demasiadamente grande, não permitindo inferências ou generalizações, mesmo que todas as considerações possíveis tenham sido analisadas. Embora torne-se redundante, deve ser enfatizado que inferências ou generalizações para o resto da população não podem ser feitas nestes casos. Isto é uma ocorrência relativamente incomum, no entanto, é muito mais prático assegurar-se de que se pode usar os dados coletados na maioria dos casos, se cuidados especiais forem tomados para evitar ou pelo menos controlar os erros.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- 1 ASTRAND, P.O. & RODAHL, K. Textbook of Work Physiology. Physiological Bases of Exercise. 3.ed., New York, McGraw-Hill Book Company, 1986.
- 2 BROOKS, G.A. & FAHEY, T.D. Exercise Physiology. Human Bionergetics and its Applications. New York, John Wiley & Sons, 1984.
- 3 BROZECK, J. & ALEXANDER, H. Components of variation and the consistency of repeated measurements. Research Quarterly, (18): 152-166, 1947.
- 6 CURETON, T.K.; WICKENS, J.S. & ELDER, H.P. Reliability and objectivity of the Springfield postural measurement. Research Quarterly, Supplement, 6(2):81-92, 1935.
- 7 HENRY, F.M.; CURETON, T.K.; HEWITT, J.E. & JARRETT, R.F. Erros in measurements. IN: F.M. Henry (Ed.). Research Methodology, New York, McGraw-Hill Book Company, 1945.
- 8 MATHEWS, D.K. Tests and Measurements in Physical Education. New York, McGraw-Hill Book Company, 1980.
- 9 McARDLE, W.D.; KATCH, F.I. & KATCH, V.L. Exercise Physiology. Energy, Nutrition and Human Performance. 2.ed., Philadelphia, Pa, Lea & Febiger, 1986.
- 10 SAFRIT, M.J. Evaluation in Physical Education. Englewood Cliffs, N.J., Prentice-Hall Inc., 1980.

Recebido para publicação em: 30/6/88.