

IMPORTÂNCIA DAS PROVAS NO PROCESSO EDUCACIONAL

Por Hindenburg Melão Jr.

www.sigmasociety.com

Texto extraído do projeto pedagógico apresentado pelo palestrante à Secretaria da Educação do município de Bom Jesus dos Perdões, e que será implementado em 2006.

Provas são muito mais do que instrumentos para “aprovar” ou “reprovar”. Prestam-se primordialmente à avaliação numérica e objetiva dos alunos, possibilitando a identificação de talentos e o diagnóstico de deficiências, bem como a quantificação da dimensão dos talentos e da gravidade das deficiências. Graças às provas é possível obter preciosas informações sobre a estrutura cognitiva e epistemológica das pessoas examinadas, e assim oferecer aos alunos e a seus respectivos tutores toda a orientação necessária para o melhor aproveitamento e desenvolvimento de suas potencialidades, reforçando os pontos fracos e cultivando os pontos fortes. Por isso a aplicação de uma prova exige que o examinador atenda a determinados quesitos: o examinador deve ter bons valores éticos, bons conceitos pedagógicos e completa compreensão das propriedades do instrumento de avaliação utilizado, a fim de que sua avaliação seja científica, unívoca e, acima de tudo, justa.

O escopo desta palestra não é um estudo exaustivo dos tópicos listados a seguir, mas sim discutir os pontos mais importantes sobre este tema e proporcionar ao educador uma visão panorâmica sobre procedimentos imprescindíveis ao pleno exercício de sua profissão, conhecimentos estes que o ajudarão a evitar os erros mais comuns, habilitando-o para construir melhores provas, tanto no que diz respeito ao conteúdo quanto no que respeita às propriedades pedagógicas e à correta interpretação das notas, além de capacitá-lo para solucionar uma vasta gama de problemas pedagógicos, substituindo os tradicionais “achismos” por procedimentos comprovadamente apropriados e eficientes.

As provas continuam sendo os melhores instrumentos que existem para avaliação pedagógica e psicológica. Há abundantes evidências de que as avaliações subjetivas feitas por professores costumam ser impregnadas de elementos pessoais, em que a relação de empatia ou apatia com os alunos acaba afetando drasticamente o julgamento do educador e levando-o a conclusões incorretas e pareceres iníquos. Eu mesmo já fui beneficiado muitas vezes por professores que simpatizavam comigo e me ofereciam algumas facilidades (em Educação Artística e Educação Física, por exemplo), enquanto outros alunos que talvez merecessem melhores notas, por terem mais talento artístico do que eu ou por serem atletas mais habilidosos do que eu, acabavam não recebendo notas tão boas pelo simples fato de não se darem bem com os professores. As notas nessas disciplinas eram atribuídas com base na avaliação subjetiva dos professores, dando margem a erros deste tipo, mas mesmo em outras disciplinas, quando havia exposição de seminários ou algo assim, a avaliação subjetiva acabava produzindo alguns resultados que talvez não fossem os mais justos. Em contraste a isso, as notas atribuídas com base em provas são muito mais imparciais e ainda por cima constituem documentos que podem ser futuramente examinados por especialistas, nos casos em que haja suspeitas sobre a equanimidade de uma avaliação. Por isso as provas são francamente superiores a qualquer método subjetivo. Alguém poderia contestar que nas provas pode-se “colar”. De fato, mas a “cola” e outros tipos de fraude são problemas causados por

falhas na supervisão, portanto situa-se no âmbito da fiscalização, ou seja, é um problema extrínseco às provas.

Os educadores que costumam se posicionar contra a aplicação de provas, geralmente o fazem por estarem habituados a verem provas de tão baixa qualidade que suas críticas, na maioria das vezes, são plenamente justificadas. Até mesmo em grandes exames, como Fuvest e Enem, há erros graves na metodologia de construção dos itens, na seleção dos itens e no tratamento estatístico dos itens, portanto não é de surpreender que, perante este quadro tão alarmante, muitos educadores contestem a validade destes instrumentos de avaliação. Não obstante, é preciso enxergar onde o problema está de fato, para que a crítica não seja injusta e resulte num preconceito que causa ainda mais danos à qualidade do ensino. É necessário compreender que não há nada de errado em usar provas discursivas ou testes de múltipla escolha com a finalidade de acompanhar a evolução do aluno. Aliás, é importantíssimo valer-se destes meios para obter informações numéricas que representem a evolução dos alunos. Portanto não se pode renunciar ao uso de provas, como se o mal estivesse nas provas em si. O mal está em não haver disciplinas adequadas nos cursos de Pedagogia que ofereçam aos futuros educadores a bagagem necessária para que se tornem capazes de elaborar, aplicar e interpretar corretamente os resultados de questionários em geral, e esse mesmo problema também se observa em cursos de áreas ligadas à Administração, Publicidade, Economia e outras que envolvam questionários (pesquisa de intenção de compra, avaliação de satisfação com produtos e serviços, avaliação do impacto causado por mudanças na política de uma empresa etc.). A atitude correta, portanto, consiste em aprender como construir melhores provas, para que estas sejam capazes de aferir com acurácia e fidedignidade o grau de desempenho dos estudantes, possibilitando diagnosticar corretamente as características cognitivas e personalógicas de cada aluno, identificando pontos fortes e fracos para, em seguida, trabalhar nestes pontos de modo a contribuir para a inserção destes alunos na comunidade e no mercado de trabalho, ajudando-os para que se tornem melhores profissionais e melhores cidadãos. É justamente esta nossa proposta: oferecer um treinamento de alto nível, objetivando instruir e orientar os educadores que estejam sinceramente interessados em fazer o melhor possível para seus alunos.

O problema da má qualidade das provas usadas no Brasil e no mundo certamente não é o mais grave nem o mais urgente, porém é o mais fácil de lidar, é o mais barato e o que requer menos tempo para ser satisfatoriamente resolvido, por isso é estrategicamente interessante que o primeiro passo no sentido de melhorar a qualidade do Ensino tenha como foco a elaboração de melhores provas. Além disso, os mesmos métodos usados para provas também possibilitam criar questionários diversos para avaliação de todos os projetos subseqüentes. Problemas como o uso de narcóticos por alunos, violência familiar, gravidez precoce e outros problemas direta ou indiretamente ligados à Educação, para serem resolvidos precisam antes ser medidos, pois sem uma informação numérica que indique se o problema está aumentando ou diminuindo, fica difícil decidir se um determinado projeto está sendo bem sucedido ou não. Muitas vezes um projeto é aparentemente promissor e todos acreditam que funcionará, então colocam-no em prática e levam o projeto até o final, para depois descobrir que o efeito foi neutro ou mesmo negativo, resultando em desperdício de tempo, esforço e recursos. Para evitar surpresas desagradáveis como essa, depois que o projeto estiver em vigor deve-se periodicamente medir os resultados e avaliar se se está cumprindo o objetivo, ou se convém substituir o projeto por outro, ou adaptá-lo em alguns pontos que não estejam funcionando a contento. E medir a qualidade de um projeto não é algo que se faça a esmo. Muito pelo contrário: para que a avaliação seja coerente e os resultados tenham utilidade, é imprescindível que sejam seguidas determinadas normas. Para avaliar tais projetos usam-se os mesmos métodos que servem para avaliar a evolução dos alunos, e quanto melhores forem os métodos usados na avaliação, tanto mais confiáveis serão os resultados. Por estes motivos, um curso

sobre provas e questionários pode ser útil para outras finalidades além de avaliações individuais.

Algumas importantes utilidades das provas, quando elaboradas, aplicadas e interpretadas seguindo normas que atendam aos nossos padrões de qualidade:

Exemplo 1: No final do ano, um aluno obteve nota média 7,7 em Matemática e nota 9,1 em Língua Portuguesa, e os professores ficaram em dúvida para qual área este aluno apresenta maior vocação. Tanto os professores quanto os alunos acharam que a prova de Matemática estava mais difícil do que a de Português, portanto a nota mais alta em Português não era indicio suficiente de que o aluno apresentava mais talento para a Comunicação. Como proceder nesse caso, a fim de poder orientar corretamente este aluno na escolha de uma carreira? A palestra ensina uma maneira muito simples e segura para resolver este impasse.

Exemplo 2: Tenho uma lista de exercícios com 58 perguntas e desejo elaborar uma prova que contenha as 10 perguntas que melhor representem o conhecimento global de todas as 58. Como devo escolher estas 10 perguntas?

Exemplo 3: Apliquei uma prova de múltipla escolha com 10 perguntas e constatei que para 9 destas perguntas a alternativa mais escolhida era a alternativa certa, mas uma das perguntas teve predominância de escolhas numa alternativa errada. Isto é bom ou mau? Por quê?

Exemplo 4: Quando eu elaboro uma prova de múltipla escolha, é melhor formular alternativas em que o aluno não consiga descartar opções improváveis? Por quê? [Sugerimos leitura do artigo do palestrante “Pontos fracos na provas da Fuvest”, disponível para download em formato PDF na Sigma Magazine, em www.sigmasociety.com]

Exemplo 5: Desejo atribuir pesos diferentes às provas de cada bimestre. Costumo usar peso 1 no primeiro bimestre, peso 2 no segundo e no terceiro, peso 3 no quarto. Mas tenho notado que algumas vezes as provas com peso 2 são mais representativas da totalidade da matéria e deveriam receber maior peso. Como resolver este problema?

Exemplo 6: Temos um programa de incentivo aos estudantes mais dedicados e, para tanto, oferecemos prêmios aos alunos com melhores notas, além de enviá-los para representar a escola em Olimpíadas da Matemática, Física, Química etc. Qual é a maneira mais justa de atribuir estes prêmios e selecionar os representantes da escola, de modo que os premiados sejam efetivamente aqueles que demonstraram melhor desempenho global, tendo em conta fatores tais como o diferente nível de dificuldade das provas?

Exemplo 7: Para os alunos do 3º ano do Ensino Médio, costumamos aplicar provas que visam a simular o desempenho que os alunos teriam no vestibular, e para isso incluímos várias questões extraídas de vestibulares passados e algumas questões novas. No entanto, o desempenho dos alunos nesses simulados costuma ser significativamente diferente do desempenho nos vestibulares propriamente ditos, de modo que estas provas não estão se mostrando como bons preditores de desempenho no vestibular. Como proceder para que estes simulados sejam preditores eficientes? (este problema é enfrentado por praticamente todos os melhores colégios e melhores cursos pré-vestibular, e é resolvido nas palestras dos módulos II e IV).

Sugestões de leitura de temas afins:

http://www.sigmasociety.com/artigos/fuvest_2004_artigo.pdf

<http://www.sigmasociety.com/artigos/historia.pdf>

http://www.sigmasociety.com/sigma_teste/sigma_teste.asp

Conteúdo programático dos módulos I, II, III e IV:

I. Propriedades de um teste:

1. Validade de construto (o teste mede o que propõe medir?).
2. Coeficiente de homogeneidade dos itens (o teste inteiro mede a mesma coisa?).
3. Estabilidade temporal das variáveis medidas pelo teste nos sujeitos da amostra (o mesmo teste, aplicado na mesma amostra, produz mesmos resultados?).
4. Norma padronizada (transformação linear das notas em escores z).
5. Norma padrão normalizada (transformação não-linear das notas em EPN).
6. Normas requintadas (como escalonar os escores com precisão e robustez?).
7. Construindo uma prova na prática. (Participação dos professores, sob orientação do palestrante).
8. Normatizando uma prova na prática. (Participação dos professores, sob orientação do palestrante).

II. Propriedades dos itens (pré-requisito: conceitos básicos de Estatística):

1. Parâmetro de dificuldade ("b" em TRI).
2. Parâmetro de discriminação ("a" em TRI).
3. Probabilidade de acerto mediante "chute" ("c" em TRI).
4. Coeficiente de Informação ("I" em TRI).

III. Avaliação e diagnóstico (pré-requisito: conceitos básicos de Estatística):

1. Interpretação dos escores nas provas.
2. Habilidades gerais e específicas.
3. Identificação de portadores de dificuldades de aprendizado.
4. Identificação de portadores de altas habilidades.

IV. Falhas gravíssimas que um teste não pode ter, mas praticamente todos têm:

1. Cuidados essenciais com a validade de construto.
2. Distribuição teórica e limites de sua representatividade da distribuição experimental.
3. Pontos fracos na prova da Fuvest (baseado no artigo de mesmo nome *).
4. Pontos fracos no *Raven Standard Progressive Matrices* (teste mais usado no mundo).
5. Erros frequentemente cometidos nos principais grandes exames do mundo.
6. Erros frequentemente cometidos nos principais grandes exames do Brasil.
7. Erros frequentemente cometidos nas provas escolares.

CURSOS COMPLEMENTARES PARA ELABORAÇÃO DE PROVAS:

Conceitos elementares de Estatística:

Conceito de medida
Amostra e população
Assíntota
Conceito de aleatoriedade e determinismo
Noções de probabilidades
Incerteza e Intervalo de confiança
Nível de significância e grau de significância
Acurácia e Precisão
Funções discretas e contínuas
Escala nominal, ordinal, intervalar e de proporção
Tendência Central
Dispersão
Assimetria
Curtose
Como fazer usando Excel?
Como fazer usando software dedicado?

Curso sobre distribuições:

Uniforme
Triangular
Binomial
Poisson
Normal
Lognormal
Beta e Beta inversa
Hipergeométrica
Chi-quadrado (1 e 2 parâmetros)
Student t
Fisher F
Erlang
Exponencial
Logística
Laplace
Cauchy (1 e 2 parâmetros)
Gamma e Gamma inversa
Gompertz
Rayleigh
Fréchet
Gumbel
Weibull
Pareto
Como fazer usando Excel?
Como fazer usando software dedicado?

Curso sobre testes de aderência ou bondade de ajuste (goodness fit):

Olhómetro
Jarque-Bera (Chi-quadrado)
Shapiro-Wilk (Chi-quadrado)
Wald-Wolfowitz
Lilliefors
Kuiper
Watson
Cramer-von Mises
Kolmogorov-Smirnov
Anderson-Darling
Como fazer usando Excel?
Como fazer usando software dedicado?

Curso sobre estimadores de tendência central e dispersão:

Moda
Mediana
Média Aritmética
Média Geométrica
Média Harmônica
Média Ponderada
Trimmed mean
Ramsay
Rogers
H de Hüber
M de Hampel
Biweight de Tukey
Andrews wave
Desvio-padrão
Variância
Stahel
Hampel
Rousseeuw
Ronchetti
Como fazer usando Excel?
Como fazer usando software dedicado?

TÓPICOS PARA CURSOS AVANÇADOS:

Curso sobre fundamentos da Metodologia Científica:

Cognição
Realidade senciente e Meta Realidade
Gnosiologia (Teoria do Conhecimento)
Epistemologia
Conceito de Linguagem (Lógica, Matemática, Programação, Esperanto, Inglês, Morse, Braille...)
Conceito de Ciência (Física, Química, Biologia, Economia, Jurisprudência, Psicologia...)
Indução e dedução
Corroboração e demonstração
Verificabilidade e falseabilidade
Rudimentos de Lógica Clássica (axiomas, proposições, conclusões, falácias típicas)
Lógica formal, Lógica formalizada e Lógica Difusa

Curso sobre operacionalização da Metodologia Científica:

Placebo e elemento ativo

Experimentos cego e duplo cego

Grupo de controle

Precauções diversas

Ferramentas estatísticas necessárias ao processo decisório científico

Coeficientes de correlação, similaridade e distância (dissimilaridade)

Regressão linear, polinomial, funções bayesianas

Testes de contraste entre médias

Testes de homoscedasticidade / heteroscedasticidade

Análise de Variância univariada e multivariada

Análise concorrente / discriminante

Análise Fatorial e outros métodos para estratificação

Coeficientes de homogeneidade

Bases da Estatística Robusta (outliers e forma)

Como fazer usando Excel?

Como fazer usando software dedicado?

Aperfeiçoamento em Teoria de Resposta ao Item (TRI):

Modelo de Rasch (1 parâmetro)

Sistema Elo de rating em Xadrez

Sistema Sonas de rating em Xadrez

Modelo Logístico de 2 parâmetros

Modelo Logístico de 3 parâmetros

Coeficiente de Informação

Métodos para ajustes simultâneos

Newton-Raphson

Funções Bayesianas

Bootstrap

Método de Monte Carlo

Cadeias de Markov

Como fazer usando Excel?

Como fazer usando software dedicado?

Sistema adotado na norma do Sigma Test:

Balanceamento de itens

Métodos para interpolação e extrapolação

Séries de Fourier

Expansões de Taylor

ARMA, ARIMA e ARFIMA

Wavelets

Variações dos parâmetros a e c em função de θ

Modelo Logístico de 4 parâmetros

Modelo Logístico de 5 parâmetros

Determinação não-simultânea dos parâmetros verdadeiros

Conceito de proporção de potencial

Como fazer usando Excel?

Como fazer usando software dedicado?

IMPORTANTE: cada item de cada curso avançado pode ser amplamente desmembrado em vários outros cursos. Por exemplo: no curso sobre “Operacionalização da Metodologia Científica”, temos o item “Ferramentas estatísticas necessárias ao processo decisório científico”, entre os quais temos o sub-item “Análise Fatorial e outros métodos para estratificação”. No curso sobre operacionalização da Metodologia Científica veremos para que serve Análise Fatorial e outras ferramentas similares, entretanto um estudo sistemático do assunto demandaria muito mais tempo. Um curso mais aprofundado deveria abordar Nuvens Dinâmicas, Análises de Clusters, Análise Fatorial Hierárquica etc., e cada um destes itens pode ainda ser amplamente dissecado, como no caso de Análise Fatorial Hierárquica podemos estudar as distâncias Euclidianas, distâncias de Shebychev, distâncias de Minkowski, distâncias de Mahalanobis, com ligações totais, ligações ward etc. Mas este nível de profundidade só se torna necessário quando se está desenvolvendo um trabalho específico em que pequenas nuances da ferramenta utilizada sejam importantes para maximizar a exatidão dos resultados. Para a grande maioria dos estudos, é suficiente uma estratificação “default”, com Nuvens Dinâmicas, por exemplo.

A idéia do curso é proporcionar o conhecimento necessário para que se possa recorrer à ferramenta certa, quando tal ferramenta se fizer necessária. E então, se for preciso atingir um grau de rigor mais elevado, pode-se investigar o assunto com mais profundidade. Por exemplo: temos um grupo de crianças que chegaram de diferentes cidades brasileiras, com diferentes idades, em diferentes séries, e como o padrão de ensino não é o mesmo nestas cidades, algumas crianças estão mais adiantadas em Matemática, mas atrasadas em Português e História, enquanto outras estão adiantadas em Português e Matemática, mas atrasadas em História, e outras estão adiantadas em tudo, outras estão atrasadas em tudo, outras estão adiantadas só em História etc. Desejamos agrupar estas crianças de modo a formar grupos tão homogêneos quanto possível em cada disciplina ao mesmo tempo evitar que crianças com idades muito diferentes fiquem juntas. Como resolver esse problema? Isso pode ser resolvido no olhometro, mas esta não é a melhor maneira. Uma das soluções apropriadas é usar Análise Fatorial. Sabendo operar um software dedicado que ofereça este recurso, e disponho dos dados sobre as crianças numa planilha do Excel, basta importar os dados e, em alguns segundos, a solução estará pronta. Isso é ensinado em nosso curso. Porém se a pessoa quiser configurar com mais detalhes a solução que deseja para o problema, rotacionando os fatores por diferentes métodos, ou usando diferentes sistemas de cálculo de ligações, distâncias, correlações etc., isso requer um conhecimento mais profundo.

<http://www.itl.nist.gov/div898/handbook/eda/section1/eda121.htm>

<http://www.itl.nist.gov/div898/handbook/eda/section1/eda16.htm>