

UTILIZAÇÃO DO MÉTODO BOOK-MARK PARA A INTERPRETAÇÃO PEDAGÓGICA DA ESCALA DE FÍSICA NO EXAME NACIONAL DO ENSINO MÉDIO (ENEM)

Eduardo Carvalho Sousa

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP)

eduardo.sousa@inep.gov.br

RESUMO: Neste trabalho, propomos a utilização método Bookmark para a elaboração de uma interpretação da escala da área de Física com base nos itens de Ciências da Natureza e suas Tecnologias do Exame Nacional do Ensino Médio (ENEM) com o objetivo de constatar a utilidade prática desse método na análise pedagógica de itens, que não seja baseada exclusivamente em critérios estatísticos. Os resultados mostram que a utilização da Taxonomia de Bloom associada ao método Bookmark é promissora para a criação de uma metodologia de criação de uma Interpretação de Escalas de Proficiências. O método viabiliza a produção de uma interpretação potencialmente mais clara da escala de medida e, conseqüentemente, mais acessível à comunidade escolar e aos pais dos alunos.

PALAVRAS-CHAVE: Escalas de Proficiência, Física, ENEM.

OBJETIVO

No Brasil, o acesso ao Ensino Superior apresenta uma problemática, pois os exames de acesso vestibulares devem definir de forma igualitária quais indivíduos terão direito a uma oportunidade não disponível para todos. Esses exames sempre foram realizados de forma descentralizada e dessa forma, limitando o pleito uma vez que, favoreciamos candidatos com maior poder aquisitivo, e restringiam a capacidade de seleção das Instituições de Ensino Superior (IES), desfavorecendo aquelas localizadas em centros menores.

No ano de 2009, o Ministério da Educação brasileiro propôs a reformulação do Exame Nacional do Ensino Médio (ENEM), com o intuito de que ele fosse utilizado pelas IES para substituir seus processos seletivos e dessa forma:

- democratizar as oportunidades de concorrência às vagas federais de Ensino Superior;
- ser instrumento de indução na construção e revisão dos currículos do Ensino Médio;
- certificar jovens e adultos no Ensino Médio.

Assim sendo, um cuidado especial deve ser tomado quanto à complexidade dos itens que comporão os testes. Tendo por base a finalidade de seleção que o ENEM assumiu e com a expectativa de candidatos extremamente preparados, é fundamental que o delineamento dos testes comporte um número razoável de itens de alta complexidade, capaz de discriminar alunos de altíssima proficiência. Para tanto, foi introduzida, também em 2009, no ENEM, a Teoria de Resposta ao Item (TRI) como metodologia de análise dos resultados utilizada no Exame. Contudo, os resultados pedagógicos oriundos do Exame apresentam-se como uma incógnita para grande parte dos interessados: alunos, pais, professores e gestores do Ensino Médio. A falta de divulgação de uma interpretação pedagógica dos níveis de proficiência com base no conjunto de habilidades características para cada nível da escala de proficiência proposta ao exame reforça a falta de entendimento pedagógico sobre o Exame.

O objetivo deste trabalho é propor uma interpretação pedagógica da escala de proficiências da disciplina de Física que compõe a prova de Ciências Natureza e suas Tecnologias do ENEM utilizando o método Bookmark. Outro aspecto a ser destacado na relevância desse tema é a falta de referencial para a produção da interpretação pedagógica da Física, ensejando o surgimento de suporte ao trabalho de gestores e docentes que atuam no Ensino Médio.

MARCO TEÓRICO

No ENEM, os resultados dos alunos avaliados são calculados por meio da TRI que permite que tanto os examinandos quanto os itens que constituem o teste estejam posicionados em uma mesma escala de proficiência. Desta forma, dentre outras, podemos compará-los entre si as proficiências obtidas para os diferentes respondentes. No entanto, essas proficiências não possuem isoladamente qualquer significado prático em termos pedagógicos. Assim, a menos que se efetue uma ligação desses valores com os conteúdos envolvidos na avaliação, pode-se dizer apenas que um indivíduo com proficiência igual a 680, por exemplo, deve possuir um conhecimento muito maior do conteúdo avaliado do que um indivíduo com proficiência 450 (ANDRADE et al, 2000).

Por esses motivos é urgente a criação e divulgação de escalas de conhecimento — também chamadas de escalas de proficiência — que tornam possível a interpretação pedagógica dos valores das proficiências de acordo com a Matriz de Competências em que se fundamenta a elaboração dos itens que constituem o Exame. Essas escalas são definidas por *níveis âncora* que por sua vez são caracterizados por conjuntos de itens denominados *itens âncora*. *Níveis âncora* são pontos selecionados a partir de critérios psicométricos na escala de proficiência para serem interpretados pedagogicamente. Já os *itens âncora* são os itens selecionados para compor esses níveis, como exemplo citemos o SAEB¹ (Sistema Nacional de Avaliação da Educação Básica): Klein (2003) faz uso da Probabilidade de Resposta Correta (PRC) para alocação dos itens nos níveis de proficiência; como a relação entre o item e a escala é probabilística, ou seja, qualquer item pode ser acertado ou errado por aluno de qualquer proficiência, definiu-se em 65% o valor percentual de acerto para o ponto da escala para onde deve ser alocado cada item. Se um item é fácil, o valor da PRC atinge 65% logo nos pontos de menor valor da escala; por outro lado, se o item é difícil, ele demanda maior proficiência para sua resolução e, portanto, fica alocado nos pontos finais da escala. Ainda, Klein (2003) propõe que um nível somente seja denominado *âncora* se há, pelo menos, 50 respondentes no nível, para que se evite uma avaliação incorreta do nível.

Os conceitos apresentados solicitam um esforço da comunidade escolar. Por conseguinte, para facilitar e garantir uma boa comunicação das informações contidas nas escalas de proficiência faz-se

1. O Sistema Nacional de Avaliação da Educação Básica (Saeb), realizado pelo Inep/MEC, abrange estudantes das redes públicas e privadas do país, localizados em área rural e urbana, matriculados na 4ª e 8ª séries (ou 5º e 9º anos) do Ensino Fundamental e também no 3º ano do ensino médio. São aplicadas provas de Língua Portuguesa e Matemática.

necessário realizar uma interpretação pedagógica, ou seja, evidenciar, além dos dados numéricos, o que cada aluno sabe e pode fazer. A pretensão com essa interpretação é a de proporcionar dados para que as equipes pedagógicas das escolas possam analisar e refletir sobre os resultados de seus alunos, bem como planejar intervenções pedagógicas orientadas para a melhoria da qualidade do ensino e da aprendizagem (UBRIACO, 2012).

Assim, ao interpretar pedagogicamente uma escala de proficiência é possibilitado um melhor aproveitamento dos resultados e o desenvolvimento de diversas atividades. Porém para que isso seja realizado, uma importante decisão é qual método orientará os trabalhos de descrição dos níveis de proficiência, ou seja, traduzir as medidas de proficiência em diagnósticos qualitativos de aprendizagem, que se espera, em termos de conhecimentos, de um estudante de determinada etapa ou série submetida ao teste.

O Método Bookmark descrito por PERIE & ZIEKY (2006), utilizado neste trabalho propõe que os itens sejam alocados em um nível e, a partir disso se realize a descrição de todos os níveis da escala. Desta forma, possibilita-se ao participante uma informação precisa do que ele é capaz de realizar em determinada área do conhecimento.

METODOLOGIA

Nesta pesquisa foram utilizados os itens da prova de Ciências da Natureza e suas Tecnologias aplicadas no ENEM dos anos de 2009, 2010 e 2011 e divulgadas. Para a obtenção dos dados para o cálculo dos valores de PCR dos *itens âncora*, foram utilizados os microdados fornecidos pelo Inep em seu site². Como a quantidade de alunos que realizam o ENEM todos os anos é muito grande, foi feito um recorte da amostra (Tabela 01) de acordo com os seguintes critérios:

- a) Somente alunos presentes na prova que responderam ao caderno regular azul.
- b) Somente concluintes regulares dos anos de 2009, 2010 e 2011.
- c) Somente alunos que possuem proficiência na área de Ciências da Natureza e Suas Tecnologias

Tabela 1.
Quantidades de respostas utilizadas
para o cálculo dos valores de PRC

Ano	Quantidade de Alunos
2009	671.174
2010	883.185
2011	1.038.047

Conforme ressaltado anteriormente, interpretar a escala significa escolher alguns pontos ou níveis da escala e descrever os conhecimentos e habilidades que os alunos demonstram possuir quando situados em torno desses pontos.

A metodologia para a interpretação das escalas inclui dois procedimentos principais: identificação de itens âncoras e a apresentação desses itens a um painel de especialistas (Klein, 2005).

Para cada nível a ser interpretado, a cada item estima-se o percentual de acertos dos alunos no nível, calculando o percentual de acerto dos alunos com uma proficiência em um intervalo de meio desvio-padrão dentro da escala que contém o nível.

2. <http://www.inep.gov.br>

Assim sendo, para definirmos *os itens ancora* dos pontos da escala utilizaremos os critérios apresentados por Mullis (2011) utilizados no TIMSS (Trends in International Mathematics and Science Study) e PIRLS (Progress in International Reading Literacy Study) em 2011, incorporando o número mínimo de observações proposto por Klein (2003):

- 1º. O número de alunos no nível que respondeu ao item é maior que 50. Esse valor é arbitrário e o seu uso é para evitar conclusões com poucas observações.
- 2º. Para os níveis com menores proficiências, que 65% ou mais dos respondentes acertem o item.
- 3º. Para os níveis com maiores proficiências, que 65% ou mais dos respondentes acertem o item e que menos de 50% dos alunos posicionados no nível anterior acertem o nível.

Para a definição dos que seriam os níveis com mais baixas ou altas proficiências foi utilizado neste trabalho um critério puramente estatístico. Desta forma, os itens que se encontravam a dois desvios-padrão da média do posicionamento dos itens nos níveis foram alocados no segundo critério e os restantes no terceiro critério.

RESULTADOS

Primeiramente, utilizando-se os critérios estabelecidos foram calculados os valores de PRC para os itens disponíveis com auxílio do *software* SPSS. De posse dessa ordenação, construímos o catálogo dos itens, peça fundamental para a aplicação do Método do Bookmark.

O catálogo foi composto por 09 itens, de 31 disponíveis e mesmo com um número de itens reduzido a qualidade dos resultados e, conseqüentemente das informações apresentadas é muito boa. A tabela 02 apresenta a alocação dos itens na escala e a sua respectiva descrição pedagógica.

Tabela 2.
Alocação e descrição dos níveis nos pontos da escala de proficiência.

		CODIGO ITEM	DESCRIÇÃO
NÍVEIS	850		
	800	FIS_CN09_17	Calcular raio em trajetória circular
		FIS_CN09_30	Calcular a quantidade de energia consumida por um gerador
	700		
	750		
	650	FIS_CN10_36	Relacionar os conceitos de empuxo e força
	600	FIS_CN10_13	Compreender o papel das transformações gasosas em um sistema termodinâmico
		FIS_CN11_21	Compreender o papel das transformações de energia em um sistema termodinâmico
	550		
	500	FIS_CN09_14	Identificar formas de economia de energia ou aumento da eficiência
	FIS_CN09_20	Identificar formas de aumentar o rendimento de uma usina de energia	
	FIS_CN10_07	Identificar o forno mais eficiente	
450			
400	INTER_CN09_01	Reconhecer formas de reduzir o efeito estufa	

Os resultados mostram que a alocação dos itens respeita a adaptação realizada por ANDERSON et al (2001) na Taxonomia de Bloom, na qual espera-se que nos níveis mais baixos, encontrem-se as ta-

refas cognitivas mais simples (conhecer, compreender, identificar) e nos mais altos, as tarefas cognitivas mais complexas (calcular, analisar etc.).

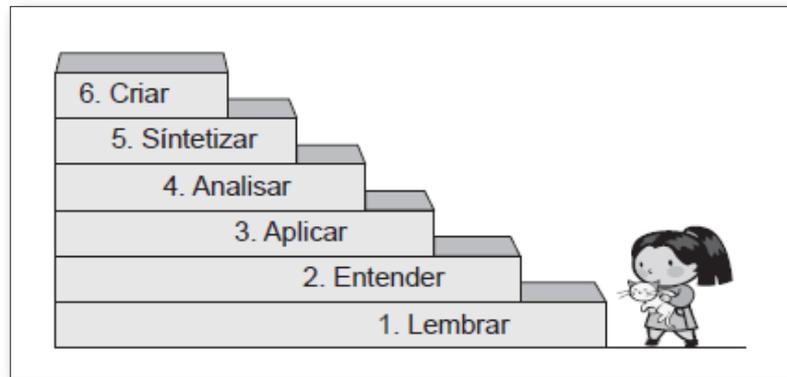


Fig. 1. Categorização atual da Taxonomia de Bloom proposta por ANDERSON et al, no ano de 2001.

CONCLUSÕES

Pode-se verificar que a utilização da interpretação da escala possibilita viabilizar informações pedagógicas mais acessíveis aos vários públicos, ao revelar o que um aluno localizado em certo nível consegue desempenhar, tendo como referência a Matriz de Referência do ENEM. A partir das informações sistematizadas, os pais, professores e gestores da área de educação podem inferir como está a aprendizagem dos alunos e o próprio aluno, diante da alocação de sua proficiência em um nível, pode se conscientizar de sua aprendizagem e se envolver mais em seu processo de formação, desde que a descrição pedagógica do nível consiga comunicar com clareza o que ele demonstra saber e ser capaz de fazer.

O procedimento utilizado demonstra ser uma grande contribuição ao processo de Interpretação de Escalas de Proficiência, mas o maior ganho que esse método oferece é a oferta de informação pedagógica que permite a geração de ações de intervenção que se concretizam na melhoria do ensino brasileiro (UBRIACO, 2012). Por exemplo, de posse dessas informações, os educadores podem repensar e replanejar seus projetos de ensino, almejando um avanço na qualidade do ensino e da aprendizagem.

Neste trabalho, elegeram-se os critérios de ancoramento proposto pelo PRILS/TIMMS e, para a descrição dos níveis, o método de Bookmark devido a sua simplicidade e por sua eficiência já comprovada.

Os resultados mostram que a gradação das tarefas cognitivas nos níveis está de acordo com o proposto pela Taxinomia de Bloom e que a utilização combinada desses métodos é promissora para a Interpretação das Escalas de Proficiências.

Como próximos passos, podem-se realizar estudos com outros critérios de ancoragem para que tenhamos mais itens por nível. Outro caminho a ser abordado é realizar uma descrição completa da área de Ciências da Natureza e suas Tecnologias (Física, Química e Biologia) e assim interpretar uma verdadeira escala de Ciências da Natureza.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANDERSON, L. W. et. al. (2001) *A taxonomy for learning, teaching and assessing: a revision of Bloom's Taxonomy of Educational Objectives*. Nova York: Addison Wesley Longman, p. 336.
- ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. DA C. (2000) *Téoria da Resposta ao Item: conceitos e aplicações*. São Paulo: Associação Brasileira de Estatística (ABE).
- KLEIN, R. (2003) Por uma educação de qualidade. *Avaliação e Políticas Públicas em Educação*, v. 11, n. 38, p. 115-120.
- KLEIN, R. (2005) Testes em rendimento escolar. *Dimensões da Avaliação Educacional*. Rio de Janeiro: Vozes.
- MULLIS, I. (2011) *Using Scale Anchoring to Interpret the TIMSS and PIRLS 2011 Achievement Scales*. Disponível em: http://timss.bc.edu/methods/pdf/TP11_Interpret_Achievement.pdf
- PERIE, Marianne; ZIEKY, Michael. (2006) *A Primer on Setting Cut Scores on Tests of Educational Achievement*. New York: ETS,
- UBRIACO, F. E. C. (2012) Interpretação de escalas de proficiência com utilização do método marcador. *Est. Aval. Educ.*, São Paulo, v. 23, n. 52, p. 86-105.