

OS DESAFIOS DA APRESENTAÇÃO DOS RESULTADOS DA AVALIAÇÃO DE SISTEMAS ESCOLARES A DIFERENTES PÚBLICOS

Nilma Santos Fontanive, Ligia Gomes Elliot y Ruben Klein

1. INTRODUÇÃO

A Avaliação vem suscitando um crescente interesse nos Sistemas Escolares contemporâneos. Desde a metade da década de 80 e ao longo dos anos 90, países da Europa, América do Norte e América Latina criaram mecanismos institucionais, centros ou programas de avaliação de seus Sistemas Educativos.

Nos dias atuais, há um grande consenso em conferir a avaliação um papel estratégico nos processos de mudanças e nas reformas educacionais.

O Brasil, desde 1990, realiza avaliações nacionais por intermédio do Sistema de Avaliação da Educação Básica – SAEB – do Ministério da Educação e, a partir de 1995, a adoção de uma Metodologia de Avaliação permitiu realizar análises comparativas do desempenho dos alunos, colocados em escalas de proficiências e também, avaliar os progressos da educação ao longo dos anos.

Um dos desafios enfrentados pelos profissionais que conduzem os processos de Avaliação de Sistemas Escolares é o de desenvolver estratégias de apresentação dos resultados capazes de facilitar a compreensão e a apropriação desses resultados pelas diferentes equipes das Secretarias de Educação, Gestores das Escolas, Professores e demais membros da comunidade.

As Avaliações Externas dirigidas aos Sistemas Escolares são, pela natureza dos seus objetivos, diferentes dos conhecidos processos de avaliação da aprendizagem realizados cotidianamente por professores e equipes pedagógicas das escolas.

Essas diferenças se expressam, de um lado, pela utilização de um instrumental técnico e estatístico pouco familiar aos professores de sala de aula e, também, aos gestores e equipes técnicas das Secretarias de Educação. De outro lado, as avaliações externas, embora bastante populares no mundo e no Brasil e não tão recentes —nos últimos 15 anos freqüentaram as pautas das políticas públicas implementadas pelos governos— dispõem de tecnologia que parece não ter sido suficientemente sistematizada e difundida no cenário educacional, e em particular, no cenário brasileiro. Não se encontra ainda hoje suficiente literatura disponível assim como também não há expressivas iniciativas de introdução de disciplina versando sobre essa temática nos cursos de formação de professores e nas licenciaturas, por exemplo.

Como ilustração tomemos como referência o Sistema de Avaliação da Educação Básica – SAEB, que vem sendo realizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP, a cada dois anos, nos últimos 10 anos, cujos resultados do desempenho dos alunos apresentados em escalas de proficiências, ainda causam perplexidade e estranheza à platéias de professores que acorrem aos seminários ou que têm nas mãos seus relatórios técnicos.

Não é nossa intenção minimizar a complexidade inerente aos procedimentos técnicos de obtenção e descrição dos desempenhos dos alunos em escalas, mas trazer para o debate formas de tradução desses procedimentos em linguagens facilitadoras da sua compreensão, apoiadas em

analogias, recursos visuais, ilustrações e exemplificações que concretizem os conceitos e aproximem a tecnologia de avaliação em larga escala do cotidiano, do próximo e do familiar ao contexto escolar.

2. UMA MANEIRA DE CONCRETIZAR O CONCEITO DE ESCALA DE PROFICIÊNCIAS

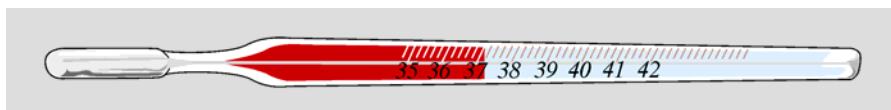
Desde 1995 o SAEB utiliza a Teoria da Resposta ao Item (TRI) para obter as escalas de proficiências (também chamadas de escala de desempenho) dos alunos avaliados. Desde então, outros Sistemas de Avaliação Estaduais e/ou Municipais também vêm adotando a TRI para descrever seus resultados.

As escalas de proficiências ordenam o desempenho dos alunos do menor para o maior em um *continuum* e elas são cumulativas, ou seja, o que os alunos sabem, compreendem e são capazes de fazer quando seu desempenho situa-se em um nível da escala, são capazes também de demonstrar as habilidades descritas no(s) nível(eis) anterior(es) dessa escala.

Nesses últimos 11 anos, os relatórios técnico-pedagógicos elaborados pelos autores deste trabalho, tanto no SAEB (Fundação Cesgranrio, 2003), quanto nas outras avaliações que têm conduzido, procuram concretizar o conceito de escala —origem, intervalos, valores numéricos, níveis— estabelecendo uma analogia desta com uma escala conhecida: a escala de temperatura corporal medida pelo termômetro. Em geral, ao lado do desenho, procura-se informar em uma linguagem coloquial que a escala de proficiências do SAEB, ou outra, é parecida com a escala de temperatura medida pelo termômetro, pois atribuem valores numéricos ao desempenho dos alunos. A ilustração do termômetro, com sua marcação dos pontos da escala serve também para explicar a origem e os níveis de interpretação, assim como a sua cumulatividade.

A explicação da origem e dos intervalos da escala Celsius possibilita estabelecer uma analogia com a origem e aos intervalos da escala de proficiências. Na escala Celsius a origem é o ponto de fusão da água (0 graus) e o seu extremo – 100 graus é o ponto de ebulição. Esta escala é graduada em centígrados. O termômetro, utilizado para medir a temperatura corporal de uma pessoa – cuja temperatura basal é aproximadamente de 36 graus, costuma apresentar os valores que vão dos 35 graus aos 42 graus.

TERMÔMETRO



Assim, se em uma situação o termômetro acusar uma temperatura de 37 graus interpreta-se que a pessoa em questão está febril, mas se o resultado obtido fosse 40 graus a interpretação seria outra, indicando a necessidade de medidas adequadas para a temperatura voltar aos níveis de normalidade.

No SAEB, a origem e a unidade de medida da escala foi arbitrada como a média e o desvio padrão da distribuição do desempenho dos alunos da 8ª série, no ano de 1997, ou seja, o valor de 250 para a média e o desvio padrão de 50.

A exemplo do termômetro (na escala Celsius), cujos pontos marcados vão de 35 graus a 42 graus, a escala do SAEB vai de 0 a 500. Esses valores numéricos são arbitrados e poderiam ser escolhidos outros. No caso do SAEB, na primeira vez em 1995, que os resultados foram apresentados em escalas, evitou-se utilizar escalas numéricas usualmente empregadas pelos professores as de 0 a 100 ou de 0 a 10 – para marcar as diferenças do seu significado.

A cumulatividade e o sentido da ordenação de escala de proficiência são conceitos que também podem ser ilustrados com os níveis da temperatura, pois se uma pessoa tem uma temperatura corporal medida em 38 graus, significa que sua temperatura saiu dos níveis de aproximadamente 36.5 graus e chegou ao valor medido. A escala de proficiências do SAEB (ou de outras avaliações de desempenho de alunos que utilizam a TRI) também apresenta valores numéricos para ordenar o desempenho dos alunos e quanto maior o ponto da escala, melhor o desempenho.

Nas escalas de proficiências, são escolhidos pontos ou níveis para interpretar as habilidades que os alunos demonstram possuir quando seus desempenhos estão situados ao redor daquele ponto ou nível. Os pontos da escala do SAEB foram arbitrados para conter o ponto de 250 e a distância entre si de meio desvio padrão.

3. A EXPLICAÇÃO DA METODOLOGIA DE INTERPRETAÇÃO DAS ESCALAS DE PROFICIÊNCIA

Nas escalas das proficiências, os pontos ou níveis contêm enunciados verbais que descrevem as habilidades ou os desempenhos demonstrados pelos alunos em cada um desses níveis. É importante explicar como essas descrições são obtidas. Em primeiro lugar, a adoção da Teoria da Resposta ao Item (TRI) permite colocar em uma mesma escala os itens de teste utilizados e o desempenho dos alunos. A interpretação da escala é feita selecionando-se (agrupando-se) os itens posicionados em um nível da escala e descrevendo-se o desempenho dos alunos naquele nível a partir da análise das respostas dadas por eles àqueles itens. O item é utilizado para interpretar o desempenho dos alunos em um nível da escala quando o percentual de acertos dos alunos com desempenhos ao redor daquele nível é igual ou superior a 65%. Isto significa que nos níveis mais altos da escala esse percentual de acertos é maior. A seguir, os itens são interpretados por especialistas das áreas curriculares ou disciplinas avaliadas, em reunião técnicas ou painéis. Nos painéis, os especialistas além de analisarem os itens de teste podem enriquecer a descrição das habilidades retomando as matrizes de habilidades ou de referência utilizadas para elaborar os itens de teste. No SAEB, ou em outras avaliações de sistemas escolares que utilizam as escalas do SAEB para apresentar seus resultados, as escalas de proficiência são comuns às três séries avaliadas: 4^a e 8^a séries do Ensino Fundamental e 3^a série do Ensino Médio. A comunalidade entre séries é obtida mediante a introdução nos cadernos de teste, de blocos de itens de uma série para a série subsequente.

É importante esclarecer que as Avaliações de Sistemas Escolares, como o SAEB ou outras Avaliações Estaduais ou Municipais costumam apresentar à população de alunos um grande número de itens para avaliar uma disciplina ou área curricular. Esta estratégia possibilita alcançar maior cobertura das habilidades descritas nas Matrizes de Referência e ainda apresentar itens com diferentes graus de dificuldade, aspecto particularmente crítico em avaliações externas. No SAEB, por exemplo, são utilizados 169 itens de teste por série e disciplina, divididos em blocos de 13 itens que dão origem a 26 cadernos distintos.

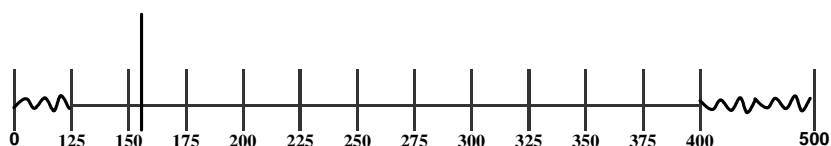
Cada aluno, individualmente, recebe um caderno contendo 3 blocos de 13 itens. Na 70% Turma, os alunos, no seu conjunto, respondem a todos ou a quase todos os itens. Para a obtenção da

escala comum do SAEB, blocos de itens da 4ª série, aparecem nos cadernos de teste da 8ª série e blocos da 8ª série são respondidos pelos alunos da 3ª série do Ensino Médio. Com esse planejamento de itens, obtêm-se uma escala de comum às três séries avaliadas e não se estabelece, *a priori*, os padrões de desempenho dos alunos. Assim, grupos de alunos de 4ª série podem apresentar desempenho nos níveis mais altos da escala do que outros alunos de 8ª série, por exemplo.

4. O ESTABELECIMENTO DAS RELAÇÕES ENTRE AS HABILIDADES DESCRITAS NOS NÍVEIS DAS ESCALAS COM OS ITENS DE TESTE E A DISTRIBUIÇÃO DOS ALUNOS NOS SEUS NÍVEIS

Um outro aspecto importante a ser tratado quando são apresentados os resultados do desempenho dos alunos em escalas diz respeito à distribuição percentual dos alunos nos diferentes níveis dessa escala, visando a favorecer a interpretação correta dessa informação. De uma forma geral, é feita a apresentação da escala onde o nível da média é exibido, como no exemplo (Fundação Cesgranrio, 2004):

ESCALA DE DESEMPENHO. LÍNGUA PORTUGUESA – SAEB 2003 – BRASIL



Além da média, costuma-se também apresentar, nos casos de avaliações estaduais ou municipais, a distribuição dos alunos pelos níveis da escala para cada turma da Escola e do Município ou do Estado. A análise dessa distribuição mostra se o desempenho dos alunos de uma turma está concentrado em torno da média, ou se está muito espalhado nos níveis da escala, como no exemplo a seguir.

QUADRO 1. DISTRIBUIÇÃO DOS ALUNOS DA TURMA X PELOS NÍVEIS DA ESCALA

NÍVEIS DA ESCALA	Abaixo 125	125	150	175	200	225	250	275	300
% de alunos nos níveis	5	20	30	20	20	5			

Como se pode observar o desempenho desta turma está concentrado nos níveis abaixo de 200 (a média geral do SAEB em Língua Portuguesa é 169,4) – cerca de 75% dos alunos. Somente 25% dos alunos estão situados acima do nível 200.

Para efeitos da análise pedagógica e interpretação da distribuição dos alunos nos níveis da escala, as equipes escolares devem considerar que:

a) Quanto maior o percentual de alunos nos níveis mais altos da escala e menor esse percentual nos níveis mais baixos da escala, melhor é o desempenho da turma. Por exemplo:

NÍVEIS DA ESCALA	Abaixo 125	125	150	175	200	225	250	275	300
% de alunos nos níveis	0	0	0	20	70	10			

b) Uma turma que apresenta uma distribuição dos alunos espalhados em quase todos os níveis da escala é muito heterogênea quanto ao desempenho. Seria este um exemplo.

Níveis da Escala	Abaixo 125	125	150	175	200	225	250	275	300
% de alunos nos níveis	5	10	15	20	30	15	5		

O resultado do desempenho dos alunos em avaliações estaduais e/ou municipais costumam ser divulgados tanto para as Secretarias de Educação quanto para as Escolas mediante a elaboração de dois materiais: boletins escolares e relatório técnico pedagógico.

Os boletins escolares oferecem aos diretores, professores e equipes pedagógicas informações sobre o desempenho dos alunos das turmas acompanhados de um padrão de referência para análise desses resultados. A seguir, recolheu-se um exemplo de um boletim confeccionado para uma avaliação realizada pela Fundação Cesgranrio (2006), abrangendo 75 municípios do Rio Grande do Sul (o SAERS 2005), em 42.883 alunos de 2ª série e 5ª série do Ensino Fundamental. O exemplo apresentado refere-se apenas ao resultado da 2ª série em Língua Portuguesa e em Matemática.

QUADRO 2. BOLETIM ESCOLAR DO DESEMPENHO DOS ALUNOS

Escola:				
Município:				
Rede: Municipal				
Código:				
APRESENTAÇÃO DOS RESULTADOS				
Este boletim apresenta os resultados alcançados pela Escola no SAERS 2005, na 2ª e na 5ª série do Ensino Fundamental. As provas aplicadas, de Língua Portuguesa e Matemática, eram compostas por itens especialmente elaborados para avaliar o desempenho dos alunos e por uma redação. Na 5ª série, além dos itens construídos, foram utilizados itens calibrados nas escalas de desempenho do Sistema de Avaliação da Educação Básica (SAEB).				
Número de Alunos Avaliados no SAERS 2005				
	SAERS Estadual*	SAERS MUNICIPAL**	SAERS NA REDE DA ESCOLA	ESCOLA
2ª Série EF	1.865	21.812	56	18
5ª Série EF	2.666	21.082	48	23
*Na Rede Estadual, participaram as Coordenadorias 25 e 32 com 56 e 63 escolas, respectivamente.				
**Participaram 75 municípios do Rio Grande do Sul.				
O Boletim descreve os resultados obtidos pelas séries avaliadas, apresentando:				
<ul style="list-style-type: none"> - a nota final da redação e as notas das competências que a compõem, por SAERS, Rede Estadual ou Municipal Geral, Rede da escola, escola e turmas; - a média dos alunos nas séries, por SAERS Estadual ou Municipal, Rede da escola, escola e turmas; - a distribuição percentual dos alunos da escola pelos níveis das escalas de desempenho de Língua Portuguesa e de Matemática, por série. 				
Para a equipe escolar avaliar o resultado apresentado neste Boletim, deve:				
<ul style="list-style-type: none"> - analisar cuidadosamente a descrição das habilidades que compõem cada um dos níveis das escalas apresentadas a seguir; - verificar os percentuais de alunos situados em cada nível da escala de desempenho, por série. Essas informações têm a seguinte interpretação <ul style="list-style-type: none"> • quanto maior o percentual de alunos nos níveis mais altos da escala e menor o percentual nos níveis mais baixos, melhor é o resultado da escola; • se os percentuais de alunos se distribuem em todos os níveis da escala, com valores aproximados, esta situação configura um resultado heterogêneo, que requer um tratamento pedagógico; • se os percentuais de alunos da escola predominam nos níveis mais baixos da escala, é preciso verificar se as habilidades descritas nos níveis mais altos da escala foram trabalhadas com os alunos. 				
Nessas duas situações, é necessário que a equipe escolar crie condições favoráveis para uma possível revisão do projeto pedagógico da escola.				

QUADRO 2. CONT.

MÉDIA GLOBAL DE DESEMPENHO DOS ALUNOS EM LÍNGUA PORTUGUESA E MATEMÁTICA

2ª SÉRIE EF

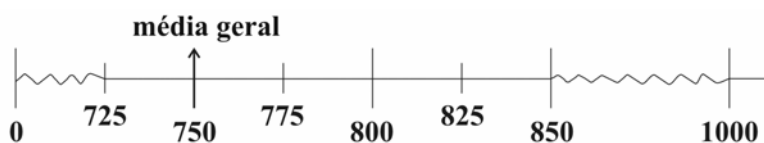
INSTÂNCIA	Língua Portuguesa	MATEMÁTICA
SAERS 2005 Geral	750,0	750,0
Rede Estadual	765,7	771,1
Rede Municipal	748,7	748,2
Rede da Escola	750,3	753
Escola	761,1	768,3
Turma U	761,1	768,3

Média Geral por Idade* em Língua Portuguesa e Matemática, 2ª Série EF

Instância	Língua Portuguesa										Matemática								
	6a	7a	8a	9a	10a	11a	12a	13a	14a		6a	7a	8a	9a	10a	11a	12a	13a	14a
SAERS 2005	748,2	757,2	758,5	745,7	725,8	718,9	721,1	713,9	723,9		756,3	757,8	758,0	746,4	728,1	719,9	723,9	716,7	724,7
Rede Estadual	774,4	771,6	773,8	757,9	740,8	725,8	744,9	753,6	721,0		772,0	775,0	777,7	768,1	750,1	742,6	763,9	741,4	745,0
Rede Municipal	738,6	754,6	757,1	745,0	724,7	718,5	719,5	711,2	724,0		750,6	754,7	756,1	745,2	726,5	718,3	721,1	715,0	723,9

* Idade em anos

⇒ Escala de Desempenho do SAERS em Língua Portuguesa



Escala de Língua Portuguesa do SAERS 2005 – 2ª Série EF

Nível	Descrição dos Níveis da Escala
Abaixo do 725	
725	Os alunos neste nível inferem informações implícitas em desenhos e em textos narrativos.
750	Neste nível os alunos são capazes de localizar informações explícitas não sequenciadas em textos narrativos e inferir conteúdos implícitos em textos narrativos. Inferem, também, o sentido de uma palavra a partir de conhecimentos prévios e compreendem informações contidas em desenhos. Reconhecem o tema de um texto narrativo. Estabelecem relação entre o texto verbal e os recursos gráfico-visuais em histórias em quadrinhos como também são capazes de estabelecer relações de causa/consequência em textos verbais e não-verbais.
775	Os alunos neste nível acrescentam às habilidades de leitura descritas nos níveis anteriores o domínio da habilidade de localizar informações explícitas em texto verbal e não-verbal. Em textos narrativos estabelecem relações de causa e consequência, identificam o personagem e inferem informação implícita. Em textos descritivos são capazes de identificar a finalidade e reconhecer o tema. Os alunos também identificam o efeito de sentido decorrente do uso de onomatopéias.

QUADRO 2. CONT.

Escala de Língua Portuguesa do SAERS 2005 – 2ª Série EF (cont.)							
Nível	Descrição dos Níveis da Escala						
800	Neste nível, os alunos demonstram habilidades de identificar marcas linguísticas que evidenciam o locutor e o interlocutor, como também de inferir informação implícita em textos narrativos e em histórias em quadrinhos. São capazes de estabelecer relações entre o texto verbal e os recursos visuais em histórias em quadrinhos.						
825	Os alunos neste nível, reconhecem os elementos que constroem uma narrativa, como a identificação de personagens, inferem conteúdos implícitos em textos narrativos e em desenhos. Reconhecem o efeito de sentido decorrente da escolha de uma determinada palavra ou expressão.						
850	Neste nível, os alunos demonstram habilidades de associar a um termo seu referente no texto e de identificar os efeitos de sentido decorrentes do uso da pontuação. São capazes também de reconhecer os recursos empregados para assinalar a transcrição da fala de personagens em histórias em quadrinhos.						
⇒ DISTRIBUIÇÃO DOS ALUNOS PELOS NÍVEIS DA ESCALA							
Instância	Abaixo de 725	725 a 750	750 a 775	775 a 800	800 a 825	825 a 850	Acima de 850
SAERS 2005 Geral	32,2	17,2	17,8	15,5	10,6	5,1	1,5
Rede Estadual	22,3	15,6	16,8	18,3	15,2	8,0	3,9
Rede Municipal	33,1	17,4	17,9	15,3	10,3	4,9	1,3
Rede da Escola	25	21,4	30,4	14,3	5,4	3,6	0
Escola	11,1	22,2	33,3	16,7	11,1	5,6	0,0
2ª SÉRIE EF - MATEMÁTICA							
⇒ Escala de Desempenho do SAERS							
Escala de Matemática do SAERS 2005 – 2ª Série EF							
Nível	Descrição dos Níveis da Escala						
Abaixo do 675							
675	Neste nível os alunos são capazes de contar os elementos de uma coleção de objetos.						
725	Os alunos, neste nível, além da habilidade descrita no nível anterior, são capazes de resolver situações-problema que envolvam a adição de números naturais e de identificar a localização de um número natural na reta numérica. Eles lêem números seguindo as regras do sistema de numeração decimal até a ordem das centenas e efetuam a adição de números naturais sem reagrupamento. Neste nível os alunos também resolvem situações-problema que envolvam a adição de números naturais.						
750	Neste nível, os alunos acrescentam ao domínio das habilidades descritas nos níveis anteriores, às habilidades de ler e interpretar calendário e horas exatas em relógios de ponteiros. Identificam a unidade de medida adequada para medir uma determinada grandeza. Interpretam dados apresentados em tabelas simples, como também distinguem entre vários objetos aquele que tem a forma de um quadrado. Eles são capazes também de identificar um número, dada sua decomposição segundo suas ordens e efetuar multiplicação de números naturais, onde o segundo fator tem apenas um algarismo. Os alunos ainda identificam o número de objetos que compõem uma dúzia.						

QUADRO 2. CONT.

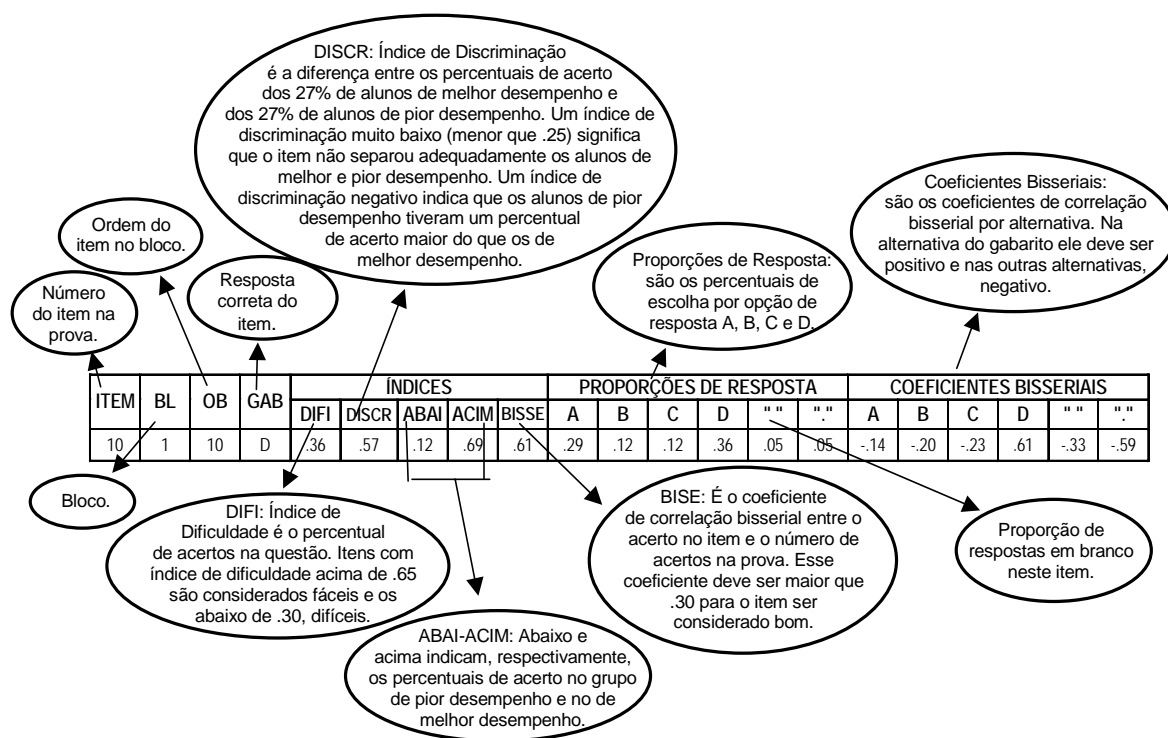
Escala de Matemática do SAERS 2005 – 2ª Série EF (cont.)								
Nível	Descrição dos Níveis da Escala							
775	Os alunos, neste nível, são capazes de resolver situações-problema que envolvam a subtração de números naturais e de adição de números naturais com reagrupamento. São capazes também de perceber regularidades em seqüências numéricas e de identificar a decomposição de um número natural segundo suas ordens. Resolvem problemas de subtração de números naturais sem reagrupamento.							
800	Neste nível os alunos são capazes de efetuar subtração de números naturais com reagrupamento e de reconhecer a quantidade de objetos que compõem meia dúzia. Lêem e interpretam dados apresentados em gráfico de coluna. Resolvem situações-problema envolvendo a idéia de comparação na subtração e que envolvam a troca de notas do sistema monetário brasileiro. Resolvem também problemas de multiplicação de números naturais. São capazes de reconhecer regularidades em seqüências numéricas e identificar a decomposição de um número natural em unidades, segundo as características do sistema de numeração decimal.							
825	Neste nível, os alunos são capazes de escrever números naturais segundo as características do sistema de numeração decimal, comparar e ordenar números naturais. Resolvem ainda, situações-problema envolvendo a adição como operação inversa da subtração.							
850	Os alunos, neste nível, são capazes de identificar entre várias figuras planas o retângulo e de reconhecer em uma lista de objetos, aquele que tem a forma esférica.							
⇒ Distribuição dos Alunos pelos Níveis da Escala								
Instância	Abaixo de 675	675 a 725	725 a 750	750 a 775	775 a 800	800 a 825	825 a 850	Acima de 850
SAERS 2005 Geral	8,3	22,3	17,8	19,1	15,9	10,4	4,8	1,5
Rede Estadual	4,9	13,4	14,1	17,5	17,9	16,7	10,5	5,0
Rede Municipal	8,5	23,0	18,2	19,2	15,7	9,9	4,3	1,2
Rede da Escola	1,8	19,6	30,4	16,1	16,1	12,5	3,6	0
Escola	5,6	5,6	22,2	22,2	16,7	22,2	0,0	0,0

Os resultados de desempenho dos alunos são também divulgados nos relatórios técnicos pedagógicos destinados aos professores e às equipes escolares, e apresentam os itens de teste acompanhados de suas estatísticas (Clássica e TRI) e de comentários pedagógicos sobre os desempenhos dos alunos nos itens. Nesses relatórios todos os itens podem ser divulgados ou apenas uma parcela deles e, essa decisão depende exclusivamente de um acordo com a equipe coordenadora da avaliação da Secretaria de Educação de um Estado ou Município. Quando se opta por divulgar apenas uma parte dos itens, estes são apresentados como exemplo de desempenho dos alunos nos diferentes níveis na escala. Nesse caso, após a descrição dos níveis com os percentuais de alunos com o desempenho naquele nível são apresentados exemplos típicos, três ou quatro itens de teste com as suas estatísticas e análises pedagógicas.

Os itens de teste são divulgados na seqüência em que aparecem nos blocos apresentados aos alunos nos cadernos de teste, caso se opte por apresentar todos os itens utilizados naquela avaliação.

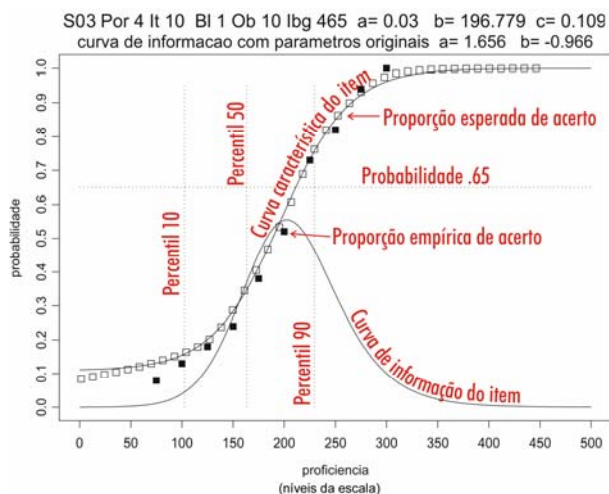
As estatísticas clássicas e as curvas de informação dos itens obtidas pela TRI não são de imediata compreensão por profissionais pouco acostumados a lidar com essas informações. No entanto, é importante que as equipes que conduzem as avaliações externas contribuam para formar uma cultura de avaliação divulgando seus procedimentos técnicos. Nessa intenção, os autores desse trabalho vêm se empenhando a apresentar os resultados estatísticos dos itens e procurado aperfeiçoar as formas de comunicação com os professores. Para as estatísticas clássicas o formato escolhido foi o similar ao existente no formulário do imposto de renda no qual o conceito a ser explicado é destacado como no exemplo abaixo (Fundação Cesgranrio, 2004):

ESTADÍSTICAS CLÁSSICAS



Além das estatísticas clássicas serão apresentadas também as estatísticas obtidas pela Teoria da Resposta ao Item (TRI). Esta teoria modela a probabilidade de acerto em função da proficiência (habilidade) do aluno e das características do item. Esta função deve ser crescente, isto é, quanto maior a proficiência, maior a probabilidade de acerto do item. A modelagem utilizada no SAEB para o item de múltipla escolha é uma função logística de três parâmetros chamada de curva característica do item, que pode ser vista no gráfico ao lado.

CURVA CARACTERÍSTICA DO ITEM

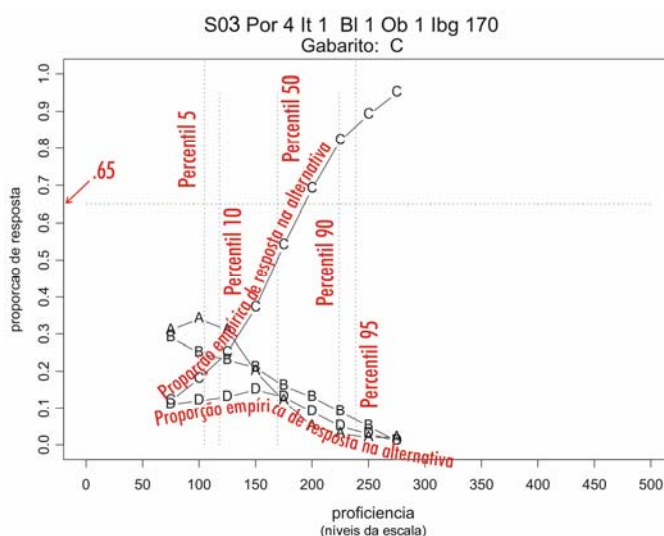


Legenda: Por 4 = Língua Portuguesa – 4ª Série; It 10 = Item 10; BI 1 = Bloco 1; Ob 10 = Ordem 10 no bloco; lbg 465 = Número do item no Programa Bilog; a, b e c = Parâmetros da função logística de 3 parâmetros.

O eixo horizontal no gráfico é a proficiência e o eixo vertical é a probabilidade de acerto que varia de 0 a 1. Traçando-se uma linha vertical em uma proficiência, na intersecção desta linha com a curva característica do item, obtém-se o valor da probabilidade de acerto no item para um aluno com aquela proficiência. O percentil 10 da distribuição de proficiências é o ponto abaixo do qual estão 10% da população de alunos e acima dele 90%. Por exemplo, entre o percentil 10 e o percentil 90 encontram-se 80% dos alunos. É importante acrescentar que quanto mais para a direita está a curva característica do item, mais difícil é o item.

O outro gráfico apresentado junto com os exemplos de itens mostra as curvas de proporção de respostas por alternativa (A, B, C, D ou E).

CURVAS DE PROPORÇÃO DE RESPOSTAS POR ALTERNATIVA

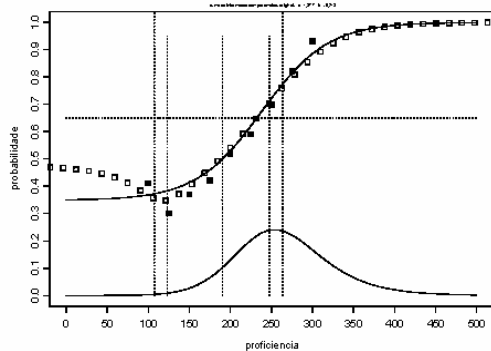


Para concluir são apresentados dois exemplos de estatísticas de itens comentadas junto com a análise pedagógica do desempenho dos alunos (Fundação Cesgranrio, 2006):

QUADRO 3. EXEMPLO DE ANÁLISE PEDAGÓGICA DO ITEM – LÍNGUA PORTUGUESA

ITEM	BL	OB	GAB	ÍNDICES					PROPORÇÕES DE RESPOSTAS					COEFICIENTES BISSERIAIS						
				DIFI	DISCR	ABAI	ACM	BISSE	A	B	C	D	'' ''	'' ''	A	B	C	D	'' ''	'' ''
34	3	10	B	0.51	0.44	0.31	0.75	0.45	0.18	0.51	0.06	0.23	0.02	0.01	-0.11	0.45	-0.26	-0.25	-0.69	-0.80

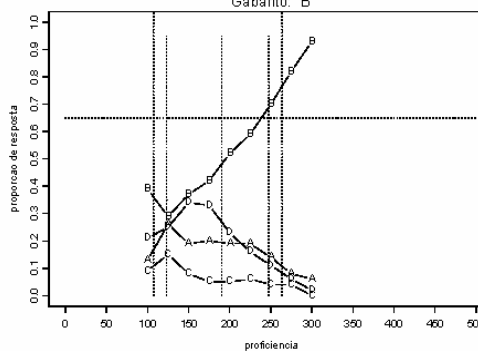
RS05 Por 5 It 34 BI 3 Ob 10 Ibg 125 a= 0.025 b= 238.234 c= 0.348



Nível: 250

RS05 Por 5 It 34 BI 3 Ob 10 Ibg 125

Gabarito: B



Este item avalia a habilidade do aluno de reconhecer uma informação explícita. Para responder corretamente, bastava acompanhar a seqüência narrativa em que, no primeiro quadro da tira, o avô, por ter 2000 anos, é identificado como contemporâneo Jesus Cristo.

Com um percentual de acertos de 51%, o item pode ser considerado de dificuldade média. O índice de discriminação (0.44) indica que o item separou bem os alunos de melhor e de pior desempenho.

Os alunos que optaram pelas alternativas erradas A (18%) e C (6%) mostram dificuldade na compreensão do diálogo e na não caracterização dos personagens. Já a escolha da alternativa D (23%) revela que os alunos foram motivados pela explicitação do nome Jesus Cristo no segundo quadro.

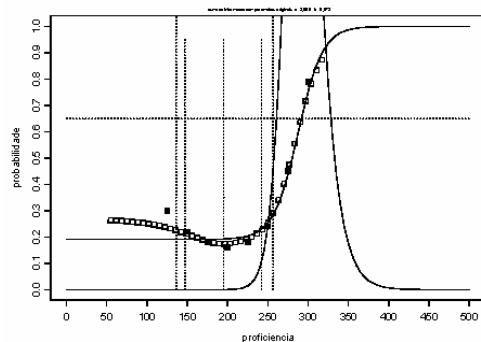
Analisando o gráfico da direita, pode-se constatar que os alunos de baixo e médio desempenho na prova dispersaram-se pelas 4 alternativas, enquanto a alternativa correta B, destacou-se a partir do percentil 50.

O item está situado no nível 250 da escala de proficiência, revelando que grande parte dos alunos nesse nível consegue identificar uma informação explícita em história em quadrinhos.

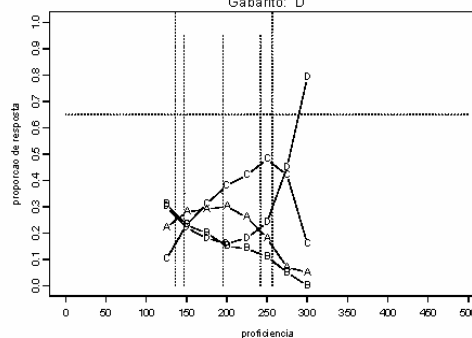
QUADRO 4. EXEMPLO DE ANÁLISE PEDAGÓGICA DO ITEM – MATEMÁTICA

ITEM	BL	OB	GAB	ÍNDICES					PROPORÇÕES DE RESPOSTAS					COEFICIENTES BISSERIAIS						
				DIFI	DISCR	ABAI	ACIM	BISSE	A	B	C	D	“ “	“ ”	A	B	C	D	“ “	“ ”
28	3	4	D	0.20	0.14	0.14	0.28	0.33	0.26	0.17	0.34	0.20	0.02	0.00	-0.15	-0.22	0.09	0.33	-0.42	-0.62

RS05 Mat 5 It 28 BI 3 Ob 4 lbg 119 a= 0.066 b= 287.564 c= 0.191



RS05 Mat 5 It 28 BI 3 Ob 4 lbg 119
Gabarito: D



Nível: 300

O item foi difícil, com apenas 20% de acerto. O índice de discriminação e o coeficiente bisserial foram baixos (0,14 e 0,33), indicando que o item não funcionou bem para discriminar os alunos quanto ao desempenho.

Além disso, o distrator da alternativa C apresentou coeficiente bisserial positivo: seu gráfico de proporções de respostas mostra uma preferência crescente com o desempenho, exceto no grupo de melhor desempenho.

Para resolver o item, os alunos deveriam efetuar uma soma de um inteiro com uma fração ($4 + \frac{1}{2}$) e expressar o resultado sob forma fracionária.

O distrator que mais atraiu foi o da alternativa C ($\frac{5}{2}$), com 34% das respostas. Os alunos que escolheram esta alternativa simplesmente somaram o número inteiro com o numerador da fração. A alternativa A ($\frac{4}{3}$) também foi muito escolhida (26% das respostas), mas é difícil entender o motivo desta preferência. De modo geral, o item sugere que os alunos não dominam somas de frações.

O item posiciona-se no nível 300 da escala de proficiência do SAEB.

REFERENCIAS BIBLIOGRÁFICAS

Fontanive, N.S. (1997). Avaliação em Larga Escala no Brasil: Sistema Nacional de Avaliação da Educação Básica (SAEB/95). In: Bomeny, H. (Org.). *Avaliação e Determinação de Padrões na Educação Latinoamericana*. Rio de Janeiro: Fundação Getúlio Vargas, PREAL, 1997.

- Fontanive, N.S. (2005). O uso pedagógico dos testes. In: Alberto de Mello e Souza. (Org.). *Dimensões da Avaliação Educacional*. Petrópolis, RJ: Vozes, 2005, p. 139-173.
- Fundação Cesgranrio (2005). *Avaliação da 4ª série do município do Rio de Janeiro e a avaliação dos alunos brasileiros*: informativo para pais e comunidade escolar. Rio de Janeiro, 2003.
- Fundação Cesgranrio (2005). *Relatório técnico do SAEB*. Rio de Janeiro, 2004.
- Fundação Cesgranrio (2005). *Sistema de Avaliação do Rendimento Escolar do Grande do Sul: SAERS 2005: relatório geral técnico-pedagógico: 2ª e 5ª séries [do] ensino fundamental*. Rio de Janeiro, 2006.