

Construção de Instrumentos para a avaliação de larga escala e indicadores de rendimento: o modelo SAEB¹

IZA LOCATELLI

Diretora de Avaliação da Educação Básica do Instituto Nacional de Estudos e Pesquisas Educacionais, vinculado ao Ministério da Educação.

Doutora em Educação pela PUC-Rio

iza@inep.gov.br

Resumo

Este estudo tem como finalidade discutir o conceito de avaliação de larga escala e sua possível e necessária apropriação pelos professores como um indicador para melhoria de seu próprio trabalho pedagógico, complementando as avaliações realizadas em sala de aula. Propõe-se, ainda, a discutir a forma de construção de instrumentos de avaliação de larga escala, suas possibilidades e limites, as formas de construção de sistemas informatizados de avaliações padronizadas e de utilização dos dados coletados através de indicadores de rendimento e de questionários contextuais, bem como as formas de disseminação dos resultados. O foco ilustrativo será aquele utilizado pelo Sistema Nacional de Avaliação da Educação Básica do Brasil (SAEB).

Palavras-chave: Avaliação de larga escala. Sistema Nacional de Avaliação da Educação Básica (SAEB). Construção de instrumentos e de indicadores de rendimento. Disseminação de resultados.

Resumen

Este estudio tiene por finalidad discutir el concepto de evaluación en gran escala y su posible y necesaria captación por los profesores, como un elemento que les permitirá mejorar su trabajo pedagógico, auxiliándoles las evaluaciones realizadas en la sala de clase. Pretende, también, discutir la forma de construir instrumentos para evaluaciones en gran escala, sus posibilidades y limitaciones, el modo de constituir las bases informatizadas producto de evaluaciones padronizadas y la utilización de esos datos recolectados por medio de indicadores de rendimiento y de cuestionarios contextuales, como también las formas de divulgar los resultados. El foco ilustrativo será el utilizado en el Sistema Nacional de Avaliação da Educação Básica do Brasil - SAEB.

Palabras-clave: Evaluación en gran escala. Sistema Nacional de Avaliação da Educação Básica (SAEB). Construcción de instrumentos y de indicadores de rendimiento. Divulgación de resultados.

Abstract

This text discusses the concept of large-scale assessment. It is suggested that teachers incorporate large-scale assessment principles into their work. In so doing, they will be able

¹ Texto apresentado no 1 Fórum Hemisférico de Avaliação Educacional (Cúpula das Américas - Linha 2: Avaliação Educacional), realizado em Brasília de 12 a 14 de março de 2002.

to complement school evaluations. The text also discusses several strategies to build up large-scale assessment practices, investigating their possibilities and limits, as well as ways of consolidating information systems related to them. It shows that large-scale assessment also involves using data collected from students, teachers, principals and school contextual questionnaires. The Brazilian National System for Basic Education Evaluation (SAEB) is presented as a case study.

Keywords: Large-scale assessment; Brazilian National System for Basic Education Evaluation; Construction of Achievement tools and Indicators; Dissemination of Results.

Introdução

A montagem de sistemas padronizados de avaliação de larga escala é um fenômeno relativamente novo no Brasil. O Brasil implementou, em 1990, o Sistema Nacional de Avaliação da Educação Básica (SAEB) e, gradativamente, começou a implantar, desde meados da década de 90, a avaliação do Ensino Médio e do Ensino Superior. Hoje, essas três modalidades de avaliação já se encontram consolidadas. A avaliação tem, portanto, assumido papel central na formulação e implementação de políticas públicas em educação.

No caso específico da avaliação da Educação Básica, o objetivo é o de monitorar a qualidade desta, fornecendo aos estados brasileiros informações não só sobre o desempenho dos alunos como também sobre os fatores associados a este desempenho, com a finalidade de reorientar políticas públicas voltadas para a educação. A ênfase, aqui, é dada ao uso das informações geradas pelas avaliações muito mais do que ao resultado em si.

No momento atual de consolidação do SAEB, o grande desafio é o de envolver professores, diretores, pais, alunos, membros das universidades e da sociedade civil, em geral, na participação do processo em seus vários momentos. Busca-se levar diferentes atores a incorporar a avaliação externa ao seu cotidiano, encarando-a como uma forma de apoio para melhorar a qualidade do ensino. Procura-se, também, discutir o papel complementar das avaliações de larga escala e as possibilidades de comparabilidade que elas propiciam, entendendo, no entanto, que estas não substituem as avaliações dos professores em sala de aula.

As avaliações de sistemas educacionais têm cunho totalmente diferente daquelas realizadas pelos professores, no cotidiano das salas de aulas. Se nessas avalia-se o processo de construção do conhecimento, nas avaliações de larga escala tende-se a avaliar o produto da aprendizagem. No entanto, se bem construídas, essas poderão, também, discutir diferentes momentos dos processos de construção do conhecimento.

Para obter a adesão dos professores às avaliações de larga escala, é preciso que os professores entendam que essas podem ajudá-los, fornecendo informações complementares às avaliações que realizam. Para tanto, devem ser bem definidos e divulgados os objetivos das avaliações.

É importante, ainda, que os professores participem da discussão sobre as matrizes de referência que informarão o que avaliar. Torna-se indispensável, também, uma ampla discussão que mostre, com transparência, todos os passos a serem dados na organização da amostra, na construção dos itens, na organização das provas e na análise de seus

resultados. Além disso, para que as avaliações tenham utilidade para os professores, é fundamental que esses recebam, em tempo hábil, informações e subsídios que os ajudem a reorientar suas práticas pedagógicas

Quando os sistemas de avaliação incorporam elementos da cultura escolar e os professores entendem que as avaliações externas podem ajudá-los, e que não são realizadas para controlar seu trabalho, terminam por assimilá-las como um componente complementar ao seu próprio trabalho. Deve-se, portanto, insistir no caráter complementar dessas duas modalidades de avaliação, as avaliações de larga escala mais padronizadas e precisas e com condições de comparabilidade e as avaliações dos professores, de caráter mais formativo, que contemplam aspectos não observáveis por aquelas primeiras.

As avaliações de larga escala: críticas e possibilidades

As avaliações realizadas com a finalidade de acompanhar os sistemas de ensino sofrem, ainda, algumas críticas. Argumenta-se, por um lado, que essas avaliações serviriam para aumentar o controle governamental sobre as escolas; por outro, que essas avaliações introduziriam a competitividade entre as escolas ou estados. Argumenta-se, ainda, que essas avaliações induziriam à redução dos currículos às áreas e tópicos abrangidos pela avaliação, servindo, também, para adequar os alunos aos novos modelos de organização da produção, para monitorar o trabalho dos professores e diretores e para pressionar os países a adotarem um currículo nacional.

De outra parte, critica-se a implementação de sistemas de avaliação de larga escala por seu custo elevado, pela necessidade de pessoal qualificado, pelo volume de investimentos na compra de equipamentos e na melhoria da competência técnica para utilização de modernos recursos de informática.

Critica-se, também, o gasto desnecessário de recursos, já que os dados gerados por pesquisas deste tipo viabilizam tão grande quantidade de informações que estas terminam por não serem eficazmente utilizadas nem pelas instâncias produtoras do conhecimento, nem pelos gestores dos sistemas. Alega-se, finalmente, que, em geral, não são implementadas medidas derivadas da análise dos dados coletados.

Na realidade, o desenvolvimento de padrões em educação, com a finalidade de servir de referência para os sistemas de ensino, baseia-se em dois grandes pólos: a questão da equidade em educação e o impacto desses padrões sobre a aprendizagem dos alunos.

Embora os críticos das avaliações de larga escala afirmem que o estabelecimento de padrões conduz a uma tendência homogeneizadora, na verdade é através desta definição que se pode enfrentar o problema da equidade com mais eficiência.

Quando se definem os padrões de proficiência desejáveis, ao término de cada etapa da escolarização, levando-se em consideração a diversidade cultural e os fatores socioeconômicos intervenientes na trajetória dos alunos, aponta-se para as escolas, para os professores e para a sociedade em geral aquilo que a escola deve ensinar e aquilo que os alunos devem aprender.

Estratégias e ações implementadas para divulgar a realização do SAEB 2001

Em constante aperfeiçoamento desde sua implementação em 1990, a equipe responsável pela execução do SAEB 2001 promoveu, antes de sua realização, encontros locais em diversos estados brasileiros com a professores, diretores e técnicos das equipes de avaliação dos estados, para discutir os objetivos do SAEB e a necessidade de um trabalho integrado entre o Inep e as Secretarias Estaduais de Educação. Nestes encontros foi distribuído o documento *SAEB 2001 - Novas perspectivas*, no qual várias questões relacionadas à avaliação são tratadas.

Tais reuniões tiveram como objetivo apresentar a versão final das Matrizes de Referência para o SAEB 2001; os passos utilizados para construção revisão e calibração de itens; o modelo de prova; e as prioridades definidas para cada uma das séries a serem avaliadas. Discutiram-se, ainda, aspectos enfocados nos questionários contextuais.

Foram realizados, também, seminários regionais com a participação de coordenadores de avaliação e membros das equipes técnicas dos estados de cada região, e um Seminário Nacional, com a presença de coordenadores dos sistemas de avaliação de todos os estados, para discutir aspectos técnicos e metodológicos que envolvem a realização de uma avaliação nacional.

Nesses seminários, foram discutidos: a construção do plano amostral; a constituição da base de dados; o uso da teoria da resposta ao item (TRI) na organização de escalas de proficiência, na equalização das provas e na calibração de itens; e as formas de apresentação e disseminação dos resultados. Foi distribuído aos participantes desses seminários um conjunto de documentos denominado *SAEB 2001 - Estratégias para a ação*.

Além dessas atividades, no esforço de melhor divulgar as finalidades do SAEB e a importância da participação de todos os

envolvidos, distribuíram-se cartazes e *folders* para escolas, diretores e participantes da amostra selecionada. Veiculou-se, também, por meio do programa interativo "*Salto para o Futuro*", produzido pela TV Escola, uma série sobre o SAEB 2001, discutindo seus objetivos, conceitos básicos, formas de construção das Matrizes de Referência, dos instrumentos, de coleta de dados e as formas de disseminação dos resultados. Finalmente, realizou-se uma teleconferência em que, além dos temas referidos, discutiu-se como seria realizada a execução do SAEB 2001, enfocando-se os procedimentos necessários para distribuição, aplicação e controle dos instrumentos utilizados.

Diante desses cuidados, a execução do SAEB 2001 não apresentou problemas, tendo sido bastante positiva a participação de todas as escolas, turmas, professores, diretores e alunos envolvidos.

A construção de avaliações de larga escala: o modelo SAEB

A construção de um programa de avaliação de larga escala como o SAEB envolve inúmeros passos, muitos deles concomitantes. Necessita-se definir os objetivos da avaliação, a periodicidade, as séries e disciplinas que serão avaliadas, a amostra de alunos e as matrizes de referência especificando conteúdos, competências e habilidades a serem avaliadas.

Necessita-se ainda, elaborar, revisar e pretestar itens, bem como revisar e pretestar as questões que servirão à coleta de dados dos questionários contextuais. A seguir, é preciso organizar um modelo de prova com as especificações de descritores e número de itens.

O passo seguinte é o de organizar os cadernos de prova, tendo-se o cuidado, no caso do SAEB, de intercalar itens e blocos comuns que permitam a comparabilidade entre diferentes anos e séries.

Antes da aplicação, faz-se a checagem da amostra, quando se verifica se todas as escolas selecionadas continuam a manter as mesmas turmas e séries anteriormente definidas no Plano Amostral, substituindo algumas por escolas-reservas, se necessário.

Outro passo a ser dado é o treinamento de aplicadores, supervisores e de coordenadores da aplicação, bem como a distribuição das provas e questionários para os diferentes pontos de aplicação, resguardando-se o sigilo do material.

Após a aplicação, cuidados devem ser tomados para a devolução do material empregado, enviando-se o mesmo para a instituição onde será realizadas a constituição das bases de dados e sua análise.

Cumpra assinalar que parte das tarefas aqui descritas sumariamente é realizada no âmbito do próprio Inep, através da Diretoria de Avaliação da

Educação Básica, sendo que a aplicação do SAEB, a constituição das bases de dados, a análise dos resultados e a produção das escalas de proficiência são de responsabilidade de empresa especializada contratada após licitação internacional.

Finalmente, analisados os dados, construídas as escalas de proficiência e realizados os estudos necessários sobre os fatores contextuais, elabora-se o Relatório Nacional e outros materiais específicos para disseminar os resultados.

Detalhamento das ações para a construção de avaliações de larga escala: o modelo SAEB

O SAEB tem como objetivos (i) monitorar a qualidade, a equidade e a efetividade do sistema de educação básica; (ii) oferecer às administrações públicas de educação, informações que lhes permitam avaliar seus projetos educacionais e formular programas de melhoria da qualidade de ensino; e (iii) proporcionar aos agentes educacionais e à sociedade informes sobre os resultados dos processos de ensino e dos fatores contextuais a eles associados.

Para atingir tais objetivos em 2001, reorganizaram-se as Matrizes de Referência do SAEB, que foram revistas por especialistas nas áreas de Currículo, Psicologia do Conhecimento, Língua Portuguesa e Matemática. Estas foram, então, submetidas à validação por especialistas das secretarias de educação de todos os estados brasileiros e por amostra representativa de professores regentes de turmas de Língua Portuguesa e Matemática.

Paralelamente, especialistas das áreas de Avaliação e Estatística revisaram os questionários contextuais, organizando as questões propostas em alguns constructos que pudessem aperfeiçoar a coleta de dados sobre alunos, professores, escolas, turmas e diretores (*SAEB 2001 - Novas Perspectivas: as Matrizes de Referência de Língua Portuguesa e Matemática*).

Construiu-se, após tais tarefas, um modelo de prova que passou a priorizar determinados tópicos e temas dos conteúdos a serem avaliados, em função das competências e habilidades definidas para as diferentes séries e disciplinas. Foram atribuídas prioridades aos tópicos e temas, o que, na prática, significou alocar um número maior ou menor de itens a cada descritor diante de sua importância na série e disciplina. Dessa forma, foi possível definir o número de itens adequado para avaliar determinados descritores considerados prioritários em determinadas séries e disciplinas.

As provas em Língua Portuguesa centraram-se na Compreensão de Textos, abrangendo os seguintes tópicos: Procedimentos de Leitura; Implicações do Suporte, do Gênero e/ou Enunciado na Compreensão do

Texto; Relação entre Textos; Coerência e Coesão no Processamento dos Textos; Relações entre Recursos Expressivos e Efeitos de Sentido; e Variação lingüística. Na prova de Matemática, abordaram-se os seguintes temas: Espaço e Forma; Grandezas e Medidas; Números e Operações/ Álgebra e Funções; e Tratamento da Informação.

Os 26 cadernos de provas para cada série e disciplina utilizados na testagem do SAEB adotam uma amostragem matricial de conteúdos, conjugada à metodologia de construção de provas denominada Blocos Incompletos Balanceados (BIB). Essa metodologia permite a aplicação de 169 itens de forma a cobrir as Matrizes de Referência. Para isso, divide-se esse conjunto de itens em 13 blocos com 13 itens cada, agrupando, em seguida, os blocos de três em três, formando 26 cadernos diferentes. Cada aluno responde apenas a um caderno de provas, totalizando 39 itens.

Para garantir a comparabilidade com anos anteriores, mantêm-se alguns blocos e itens comuns aplicados nos diferentes ciclos de avaliação. Os resultados são analisados utilizando-se a Teoria da Resposta ao Item, que permite colocar os resultados em uma mesma escala de desempenho, ainda que nem todos os alunos tenham respondido a todos os itens. É de se ressaltar que as provas foram organizadas por especialistas das áreas de conteúdo, psicometristas, estatísticos, especialistas em avaliação e em psicologia cognitiva. (SAEB 2001 – *Novas perspectivas: a organização das provas*).

Em relação aos questionários contextuais, procurou-se obter medidas que permitam captar, além do que os alunos aprendem, como os resultados educacionais se distribuem em face da origem social dos alunos, como evoluem as condições de qualidade da escolarização, como as condições de qualidade da escolarização se distribuem em função da origem social dos alunos e quais fatores escolares promovem eficácia e equidade na educação brasileira. São apresentados questionários aos alunos, professores, diretores havendo ainda um questionário sobre a turma e outro sobre a escola, preenchidos pelo aplicador das provas.

No constructo relacionado aos alunos foram privilegiados a caracterização sociodemográfica, o capital cultural, o capital social, a motivação e auto-estima, as práticas de estudo e a trajetória escolar. No constructo relacionado à sala de aula, o papel desempenhado pelo professor é fundamental. Optou-se, então, por considerar a caracterização sociodemográfica dos professores, a formação, a experiência profissional, as condições de trabalho, o estilo pedagógico e a expectativa destes em relação aos alunos. Quanto à escola, os constructos foram captados pelos questionários do diretor, da escola e do professor.

Aborda-se, nesses, a caracterização sociodemográfica do diretor, formação, experiência, liderança, condições de trabalho do diretor e equipe, trabalho colaborativo, organização do ensino e políticas de promoção, clima acadêmico, clima disciplinar, recursos pedagógicos, situação das instalações e equipamentos e atividades extracurriculares.

Foram incorporados aos questionários do SAEB estudos acerca de fatores associados ao desempenho escolar e à eficácia da escola, a partir de uma síntese dos pressupostos que abordam o tema efeito-escola (SAEB 2001 - *Novas perspectivas: a organização dos questionários*).

Paralelamente às definições e ações anteriormente descritas, organizou-se o Plano Amostral do SAEB 2001, tendo como meta produzir estimativas confiáveis para alguns estratos de interesse. Foram definidos cinco critérios para a estratificação das escolas: o primeiro deles refere-se à série em que o aluno está matriculado, a saber, 4ª e 8ª séries do ensino fundamental e 3ª série do ensino médio. O segundo é o da unidade da Federação (26 estados brasileiros e o Distrito Federal). O terceiro é a dependência administrativa das escolas (estadual, municipal e particular). Em quarto lugar, considerou-se critério a localização da escola (capital ou interior). Finalmente, as escolas foram também agrupadas por tamanho - por um lado, aquelas com até duas turmas da série pesquisada; por outro, as escolas com três ou mais turmas da série. Combinados esses critérios, definem-se os alunos pelos estratos de interesse. No mínimo cada estrato do ensino fundamental possui 2.000 alunos e do ensino médio, 1.100 alunos.

Inicialmente selecionam-se as escolas, depois as turmas dentro das escolas e os alunos destas turmas presentes no dia da avaliação. Após essa seleção, realiza-se a checagem das escolas, para que, em tempo hábil, as escolas possam ser substituídas, se necessário, por escolas-reserva. Procedimentos estatísticos relevantes são utilizados de forma a que o resultado obtido espelhe a proficiência dos alunos brasileiros nas séries e disciplinas avaliadas. (SAEB 2001 - *Novas perspectivas: a construção do Plano Amostral*).

Tendo-se cumprido todas essas etapas e tendo-se feita a divulgação prévia do SAEB, como já descrito, as provas foram aplicadas iniciando-se uma segunda fase: a constituição e análise das bases de dados e o planejamento da disseminação dos resultados.

A constituição das bases de dados do SAEB

Avaliações de larga escala, como as realizadas pelo SAEB, fornecem uma quantidade imensa de dados. Nem sempre, contudo, os sistemas

nacionais de avaliação as utilizam forma a atender aos objetivos propostos. Algumas perguntas, sem dúvida, necessitam de serem feitas se efetivamente se quiser colocar a avaliação a serviço da educação.

No fundo, todas as perguntas têm por objetivo informar se os alunos estão aprendendo o que deveriam aprender e que condições favorecem a aprendizagem, isto é, os enfoques contextualizam as diferentes trajetórias dos alunos ou partem de premissas que terminam por penalizar escolas, alunos e professores? De outra parte, os resultados servem aos professores e diretores, oferecendo-lhes subsídios para melhorar seu trabalho e servem, também, aos gestores para a tomada de decisões e formulação de políticas educacionais ou tangenciam tais aspectos? Os resultados são do conhecimento da população e, especialmente, dos pais dos alunos que freqüentam nossas escolas? A sociedade é envolvida na criação de uma cultura de avaliação?

Normalmente, as bases de dados dos sistemas de avaliação são construídas para responder a essas questões. No caso do Brasil, hoje é possível contar com indicadores confiáveis tanto oriundos do Instituto Brasileiro de Geografia e Estatística (IBGE), que produz diferentes indicadores sobre vários aspectos da população brasileira como um todo, como do Censo escolar (Inep), que organiza dados sobre a população escolar. Inúmeros pesquisadores em várias universidades brasileiras utilizam-se das bases de dados do SAEB para realizarem pesquisas acoplando essas bases às especificidades da população captadas através do Censo Nacional do IBGE e do Censo Escolar do Inep.

É possível, assim, contextualizar resultados em função de taxas de matrícula nos diferentes graus de ensino, taxas de evasão, promoção, repetência, de programas de aceleração, de projetos de escolarização de jovens e adultos, de taxas de conclusão nos diferentes graus de ensino, do tempo médio de conclusão em cada ciclo, do índice de desenvolvimento socioeconômico da região ou estado, da renda *per capita* de cada estado, da composição familiar, da etnia, cor, raça, sexo etc.

O banco de dados do Inep, em geral, é disponibilizado para as instituições de pesquisa que desejem realizar estudos sobre diferentes temas, sendo estas incentivadas a fazê-lo, mediante a assinatura de convênios ou mesmo de termos de compromisso entre o Inep e o pesquisador responsável, com o aval de sua instituição.

É importante assinalar, também, que cada estado participante da avaliação nacional recebe a sua base de dados para aprofundar análises específicas sobre sua realidade.

Programas de avaliação com diferentes universidades (PROAV) foram instituídos para facilitar esse intercâmbio. Todos os dados recolhidos

pelas avaliações realizadas pelo SAEB fazem parte do Banco de Dados do Inep, podendo, assim, serem acessados pelos pesquisadores e pelas instituições desde que cumpridas algumas formalidades legais, entre as quais a definição do campo de pesquisa e o compromisso de não divulgar resultados específicos de escolas ou alunos, já que o SAEB divulga seus resultados apenas em nível nacional, regional e estadual.

Podem ser disponibilizadas informações detalhadas sobre a constituição da amostra, sobre as matrizes de referência, os níveis de proficiência alcançados pelos diferentes estratos, as escalas de proficiência, os itens-âncora dessas escalas, os índices de dificuldades dos itens, os índices de comportamento diferencial dos itens, os índices de discriminação, as bases de dados oriundas das respostas aos questionários contextuais e os parâmetros de todos os itens.

Só não são disponibilizados alguns itens que poderão ser reutilizados em outras provas, mantendo-se, ainda, reserva técnica de itens para atender às demandas dos estados que mantêm acordos de cooperação técnica com o Inep.

Como é normal em sistemas de avaliação de larga escala, o SAEB possui um banco de itens, constantemente atualizado, com itens calibrados, mediante pré-testes em amostras representativas da população. O Banco Nacional de Itens zela pela qualidade dos itens a serem aplicados. Na prática, as habilidades e os parâmetros dos itens são estimados a partir das respostas de um grupo de alunos submetidos a esses itens. Uma vez estabelecida a escala de medida da habilidade, os valores dos parâmetros dos itens são colocados nessa escala única, permitindo calcular as proficiências, também na mesma escala.

Os dados gerados vão sendo gradativamente introduzidos na base de dados, lançando-se as respostas dos indivíduos ao teste e o gabarito das respostas. São calculadas algumas estatísticas descritivas como número de indivíduos submetidos a cada item, número e porcentagem de acertos em cada item e algumas correlações de interesse, como, por exemplo, a correlação bisserial. Depois, estimam-se as habilidades dos respondentes utilizando-se procedimentos estatísticos com base na Teoria da Resposta ao Item.

É importante assinalar que o SAEB adota, além da Teoria Clássica, o modelo estatístico da Teoria da Resposta ao Item (TRI). Nesta, o processo de estimação dos parâmetros dos itens é conhecido como calibração. A TRI permite que os valores das habilidades dos respondentes e dos parâmetros dos itens possam ser colocados em uma única escala. Como o Banco de Itens agrega itens com parâmetros já conhecidos, por terem sido estimados antes da realização das provas, é possível submeter, durante as provas, os

indivíduos a itens já calibrados, sendo necessário apenas estimar as habilidades dos indivíduos e dos novos itens que foram incorporados às provas.

Por outro lado, por possuir itens já calibrados, o Banco Nacional de Itens pode fornecer a estados que desejem realizar suas próprias avaliações itens para compor suas provas que permitem comparabilidade com os resultados do SAEB. Um ponto, no entanto, é necessário observar quando se constrói um banco de itens:

A questão da métrica deve ser muito bem considerada, isto porque quando um banco de itens é construído, uma informação fundamental é a escala em que foram calibrados. Isto porque as habilidades dos indivíduos que serão estimadas futuramente, a partir daqueles itens, estarão nesta mesma métrica, portanto, quaisquer comparações diretas só poderão ser feitas com outros indivíduos que também tenham suas habilidades nesta escala." (Valle, 2000, p.57).

Uma vez que todas as habilidades dos respondentes e todos os parâmetros dos itens sejam conhecidos, estando todos na mesma métrica, isto é, quando todos os parâmetros são comparáveis, pode-se construir escalas de interpretação para os resultados obtidos.

Se não for feita uma interpretação dos dados da escala, pouco saberemos em termos pedagógicos, isto porque a relação pedagógica só se estabelece quando há uma ligação entre os valores obtidos e os conteúdos e habilidades envolvidos na avaliação.

A escala de proficiência do SAEB é definida por níveis-âncora que, por sua vez, são caracterizados por conjuntos de itens chamados de itens-âncora. Para ser considerado âncora, um item deve ser respondido pelo menos por 65% de indivíduos com um determinado nível de habilidade e, no máximo, por 50% de indivíduos com habilidade inferior a este grupo. Um item-âncora é, portanto, típico de um determinado nível, ou seja, é bastante acertado por quem possui um determinado nível de proficiência e pouco acertado por quem não o possui. Painéis de especialistas descrevem, após seleção dos níveis e itens-âncora, as habilidades demonstradas pelos alunos em cada nível.

Para evitar problemas na construção da escala de proficiência, o SAEB 2001 trabalhou com um número grande de itens para cada descritor, de tal forma que se tenha efetivamente elementos para interpretar as habilidades de cada nível da escala. Assim, por exemplo, o descritor Inferir Informação Implícita em um Texto, do tópico Procedimentos de Leitura, apresenta 27 itens, em gradação diferente de dificuldade, para que se possa não só interpretar os resultados, como também observar em que ponto da

construção desta habilidade os alunos apresentam problemas, o que sem dúvida favorecerá o entendimento dos professores quando da disseminação dos resultados.

Problemas na disseminação dos resultados de avaliações de larga escala

Preparada e realizada a avaliação, constituída a base de dados, cabe responder às perguntas feitas anteriormente: o que fazer com os dados coletados, como levar diferentes segmentos a deles se apropriarem e a utilizá-los em prol da melhoria e da equidade em educação?

Segundo Ravela (Preal, 2000) esse é o grande problema dos sistemas nacionais de avaliação. Para esse autor, não só os resultados não chegam a ser difundidos de forma a causar impacto nos diferentes segmentos de interesse, bem como não se informam alguns pontos fundamentais, tais como:

1. em uma amostra censitária, a porcentagem de alunos avaliados em relação ao total de alunos;
2. a estimativa da margem de erros amostrais;
3. os procedimentos usados para equiparar os instrumentos de medição usados em anos anteriores;
4. o modo como, muitas vezes, os fatores associados ao desempenho são apresentados, levando a uma visão quase que causal entre diferentes variáveis e a proficiência dos alunos;
5. a falta de estudos etnográficos que permitam observar o que efetivamente ocorre em escolas com bom ou mau desempenho;
6. a falta de informações claras para os professores sobre os processos de construção de conhecimentos pelos alunos;
7. a falta de informações às famílias, ilustrando-se claramente o que seus filhos estão ou não aprendendo e o que podem fazer para apoiar o trabalho da escola;
8. a necessidade de se ilustrar com mais clareza, nos informes nacionais, o que os alunos aprendem, não apenas publicando um ou outro item relativo a tal ou qual nível de desempenho, mas indo além, buscando explicações mais aprofundadas;
9. a necessidade de se contextualizar os resultados já que a composição sociocultural do alunado e das escolas é bem diferente;
10. a necessidade de análises sobre fatores associados, que incluam perspectivas tais como as de análise hierárquicas, agregando-se valor aos informes;

Em relação a alguns pontos, o SAEB já vem apresentando tais informações, estando todas disponíveis no Banco de Dados além de serem

divulgadas no Relatório Nacional e nos Relatórios Metodológicos. Assim é possível saber a porcentagem da população que participou da amostra, a estimativa do erro amostral e os procedimentos usados para a equalização.

Alguns outros pontos, entre os assinalados, já começaram a partir de 1999 a receber novo tratamento. Não mais os fatores associados ao desempenho são apresentados paralelamente às proficiências alcançadas, mas sim em termos de análises multinível com os fatores hierarquizados, utilizando-se o conceito de valor agregado, o que permite detectar escolas efetivas não obstante, por exemplo, condições sociais adversas. Diversas pesquisas têm sido realizadas sobre esses aspectos, bem como estudos de caso que permitem um olhar para o interior das escolas. Nesse contexto, destacamos: i) Maria Eugénia Ferrão, Kaizô Beltrão e Cristiano Fernandes. "Aprendendo sobre "Escola Eficaz - Evidências do SAEB/99". 2001. (mimeo.); ii) José Francisco Soares (Coord.) "Fatores Explicativos do Desempenho em Língua Portuguesa e Matemática: a Evidência do SAEB/99". 2001 (mimeo.); e iii) Clarilza Prado (Coord.). "Estudo Exploratório SAEB/99". 2001. (mimeo.).

Novos estudos a serem desenvolvidos com base nos dados coletados pelo SAEB 2001 incluem, entre outros tópicos: i) a investigação sobre gestão escolar, buscando evidências sobre o papel da liderança dentro da escola, tendo como pressuposto a literatura sobre escola efetiva; ii) a identificação das diversas práticas pedagógicas dos professores brasileiros e seus efeitos na aprendizagem dos alunos; iii) o efeito da liderança da direção e da colaboração entre docentes nos resultados dos alunos; iv) o impacto de aspectos do clima da escola, como clima disciplinar e clima de colaboração direção-corpo docente, nos resultados escolares, perfil do aluno repetente; v) os fatores sociais associados à probabilidade de repetência; e vi) as condições que determinam uma escola efetiva.

SAEB 2001: disseminação de resultados

Em 2001, pretende-se, a partir da revisão da literatura nacional e internacional, disseminar os resultados de forma mais abrangente, direcionando-a para públicos diferenciados, com informações claras, concisas e que sirvam à tomada de decisões tanto por professores, quanto por diretores e gestores dos sistemas servindo ainda à criação de uma cultura de avaliação na sociedade.

As estratégias de disseminação dos resultados do SAEB 2001 refletem um rigoroso planejamento prévio que levou em conta tanto às diretrizes técnicas elaboradas no âmbito do INEP e a experiência acumulada na realização do SAEB nos últimos anos, como as discussões realizadas com representantes

das secretarias estaduais e municipais de educação, seus professores e diretores, pesquisadores e gestores ligados ao processo de formulação, implementação e gestão das políticas educacionais em seus diferentes níveis(...) As mudanças que estão sendo implementadas nesse sentido são bastante abrangentes, referindo-se tanto à forma e meios utilizados quanto à diversidade de produtos e à própria avaliação dos resultados que agora serão apresentados a partir de um estudo qualitativo que irá enriquecer substancialmente a análise do desempenho dos alunos.(...) Planeja-se para a disseminação dos resultados do SAEB 2001 um trabalho sistemático de difusão da informação pra diferentes públicos, de forma que toda a sociedade possa atribuir significado aos dados que os relatórios e boletins de divulgação de resultados vierem a oferecer (Locatelli e Andrade, 2001).

Os produtos a serem divulgados incluem:

1. *Relatório Nacional e Relatórios Regionais*

Nestes, os resultados serão apresentados de forma contextualizada. Buscar-se-á, assim, explicitar as condições em que os resultados ocorrem, já que o trabalho educacional não ocorre num espaço abstrato, mas em um terreno concreto onde habitam professores e alunos que têm condições absolutamente diversas, especialmente num país com a complexidade do Brasil, com regiões de desenvolvimento socioeconômico absolutamente desiguais.

Fontes de dados como as fornecidas pelo Censo Escolar e pelo IBGE ajudarão a compor o cenário para que sejam relativizadas as desigualdades de resultados, entendendo-se que estas refletem as desigualdades sociais. No entanto, serão conjuntamente apresentados resultados de pesquisa que mostram que a escola pode fazer diferença, apesar das condições adversas vivenciadas por muitas. Para além do determinismo social, é preciso discutir questões intraescolares que ajudem a buscar soluções que minimizem o efeito das condições sócio econômicas sobre a escolarização. Estudos sobre Escola Eficaz e Efeito Escola podem ajudar a buscar respostas para tais questões.

2. *Série de estudos especiais: análise qualitativa das provas do SAEB*

Este estudo analisa todos os 1.062 itens aplicados nas provas do SAEB 2001, tanto em Língua Portuguesa quanto em Matemática e, a partir de um modelo construído pela equipe da Daeb, busca interpretar os momentos de ruptura na construção do conhecimento pelos alunos em função das questões propostas para avaliação.

Por meio desse modelo, pode-se interpretar não itens isolados, mas conjuntos de itens que compõem um determinado descritor, o que

possibilita a construção de gráficos com percentuais de acertos por descritor, em função dos níveis de dificuldades dos itens que o compõem.

Assim, por exemplo, o descritor Selecionar Informações Implícitas em um Texto contém 27 itens que serão analisados de tal forma que será possível identificar, por exemplo, o domínio de recursos lingüístico-discursivos na construção de diferentes gêneros e tipos de textos, a competência de refletir sobre diferentes temas, as competências relacionadas aos elementos estruturais de um texto (forma, função e estrutura), a competência no uso ou não das pistas verbais e não verbais fornecidas pelo texto, o uso ou não das pistas lingüístico discursivas. Todos os textos e as questões a eles relacionadas recebem uma classificação quanto ao nível de dificuldade, além de se explicar o que cada um demanda e que foi ou não atingido pelo aluno. Não se trata apenas de analisar as respostas (gabarito e distratores), mas as operações mentais exigidas para processar o texto. Isto auxiliará os professores a "enxergar os pontos de corte" na construção de conhecimentos por seus alunos.

3. Boletim para diretores e professores e se possível para pais

Trata-se de publicação com os resultados de forma sintetizada apontando-se para os objetivos da avaliação e seus resultados. Pretende-se traduzir em linguagem simples com apoio de gráficos, os resultados, de modo que se entenda o que os alunos efetivamente sabem nas séries avaliadas, o que necessitam saber e os percentuais de acertos nos diferentes domínios avaliados.

4. Série de publicações sobre pesquisas e análises dos resultados do SAEB 2001

Pesquisas sobre Gestão Escolar, Liderança do Diretor, Estilos Pedagógicos e sua distribuição entre os diferentes grupos sociais; análise hierárquica dos fatores associados ao desempenho os alunos; e a análise dos fatores associados à repetência dos alunos serão publicadas em uma série específica de publicações do Inep sobre avaliação.

5. Outras ações envolvendo a participação direta dos professores

Serão realizados, também, Encontros Regionais e um Encontro Nacional para discussão dos resultados. Será, também, produzido um vídeo com os pontos mais importantes a serem discutidos, o qual será reproduzido para as Secretarias Estaduais para que estas discutam os resultados com seus professores. Além disto, por meio de programa interativo produzido pela TV Escola, serão discutidos de forma interativa os resultados alcançados. Todos os materiais serão divulgados, também, via Internet.

Estão previstos, ainda, vários encontros nos diferentes estados para discutir os resultados locais e também, encontros com membros da imprensa para uma explicação mais detalhada do processo de avaliação que não se resume à atribuição de pontos em uma escala e à constituição de *rankings* entre estados ou regiões.

Finalizando, é importante assinalar que diagnósticos são importantes mas se não forem utilizados de nada servirão. É compromisso do MEC, por meio do Inep, devolver a análise dos resultados da avaliação promovida a diretores, professores, gestores, pais, alunos e à sociedade em geral. A difusão de informações em linguagem clara e acessível é a contrapartida natural aos esforços que toda sociedade vem realizando em busca de uma educação de qualidade visando à transformação de nosso sistema educacional através do diagnóstico, da reflexão e da ação.

Referências bibliográficas

ANDRADE, Adler do Couto. *Fatores Associados ao Desempenho: a Experiência do SAEB*. 2001. Texto apresentado na série de seminários "SAEB 2001 – Estratégias para a Ação", realizado entre os meses de novembro e dezembro de 2001 (mimeo.)

_____. A Disseminação dos Resultados de uma Avaliação de Larga Escala: a Experiência do SAEB. In: *Anais do Marco de Aprendizagem Contínua em Avaliação*. Salvador: Agência de Avaliação (UFBA/ISP-Fapex), 2001, pp. 95-104.

ANDRADE, Dalton. Comparando Desempenhos de Grupos de Alunos por Intermédio da Teoria de Resposta ao Item. In: *Estudos em Avaliação Educacional*. Fundação Carlos Chagas, nº 23, janeiro/junho de 2001.

ANDRADE, Melissa.; BOMENY, Helena.; HERNANDEZ, Elizabeth.; MORENO, Flávia. *O Sentido da Disseminação do SAEB*. 2001. Texto apresentado na série de seminários "SAEB 2001 – Estratégias para a Ação", realizado entre os meses de novembro e dezembro de 2001 (mimeo.)

CABRAL, Ana Beatriz. *Construção e Validação de Itens*. (mimeo.). Texto apresentado na série de seminários "SAEB 2001 – Estratégias para a Ação", realizado entre os meses de novembro e dezembro de 2001.

CASTRO, Maria Helena Guimarães. Avanços e Desafios da Educação Brasileira. In: *La experiencia brasileña en la evaluación de la Educación Básica*. Brasília: Inep, janeiro de 2002, pp. 5-15.

CONDÉ, Frederico Neves. *Análise Empírica de Itens*. Texto apresentado na série de seminários "SAEB 2001 - Estratégias para a Ação", realizado entre os meses de novembro e dezembro de 2001 (*mimeo.*).

_____. *As Escalas de Desempenho do SAEB*. Texto apresentado na série de seminários "SAEB 2001 - Estratégias para a Ação", realizado entre os meses de novembro e dezembro de 2001 (*mimeo.*).

LOCATELLI, Iza. O Sistema de Avaliação da Educação Brasileira. In: *La experiencia brasileña en la evaluación de la Educación Básica*. Brasília: Inep, janeiro de 2002, pp. 17-30.

_____. Novas Perspectivas de Avaliação. In: *Ensaio - Avaliação e Políticas Públicas em Educação*. Fundação Cesgranrio, nº 33, Vol. 9, outubro/dezembro de 2001.

_____. Políticas de Avaliação no Brasil. In: *Anais do Marco de Aprendizagem Contínua em Avaliação*. Salvador: Agência de Avaliação (UFBA/ISP-Fapex), 2001. pp. 7-14.

LOCATELLI, Iza; ANDRADE, Adler do Couto. *A Disseminação dos Resultados de Avaliações de Larga Escala: a Experiência do Brasil*. Texto apresentado na X Reunião de Coordenadores Nacionais do Laboratório Latino-Americano de Avaliação da Qualidade da Educação, realizada em Fortaleza, Brasil, de 27 a 29 de agosto de 2001.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS. *SAEB 2001 - Novas Perspectivas*. Brasília, setembro de 2001.

RABELLO, Guilherme Coelho. A Técnica de Equalização: um Estudo Comparativo com os Dados do SAEB. In: *Anais do Marco de Aprendizagem Contínua em Avaliação*. Salvador: Agência de Avaliação. (UFBA/ISP-Fapex), 2001. pp. 29-41

_____. *A Teoria de Resposta ao Item (TRI) e o Sistema de Avaliação da Educação Básica (SAEB)*. Texto apresentado na série de seminários "SAEB 2001 - Estratégias para a Ação", realizado entre os meses de novembro e dezembro de 2001 (*mimeo.*).

_____. *A Metodologia de Equalização e o Sistema Nacional de Avaliação da Educação Básica (SAEB)*. Texto apresentado na série de seminários "SAEB 2001 - Estratégias para a Ação", realizado entre os meses de novembro e dezembro de 2001 (*mimeo.*).

RABELLO, Guilherme Coelho; CONDÉ, Frederico Neves. A Invariância dos Parâmetros na Teoria de Resposta ao Item: um Estudo com os Dados do SAEB. In: *Anais do Marco de Aprendizagem Contínua em Avaliação*. Salvador: Agência de Avaliação. (UFBA/ISP-Fapex), 2001. pp. 43-50.

RAVELLA, Pedro. *Como presentan sus resultados los sistemas de evaluación educativa de América Latina*. Revista Preal, nº 20, agosto de 2001.

VALLE, Raquel. *Teoria da Resposta ao Item*. In: *Estudos em Avaliação Educacional*. Fundação Carlos Chagas, nº 23, janeiro/junho de 2001.