

Avaliação educacional e políticas públicas: a experiência do Sistema Nacional de Avaliação da Educação Básica (Saeb)

Tereza Cristina Cotta

Introdução

A implantação de um sistema de informações educacionais, abrangendo todos os níveis de escolaridade, foi um acontecimento marcante no cenário da educação brasileira.

Há poucos anos, a situação era bem diferente. O principal levantamento estatístico da área, o Censo Escolar, não era realizado desde 1986. Os poucos dados disponíveis provinham do Censo Populacional do Instituto Brasileiro de Geografia e Estatística (IBGE), que, por não ter a finalidade específica de coletar dados sobre educação, continha muitas lacunas. Esta insuficiência na oferta de estatísticas educacionais, que são fundamentais para o gerenciamento do sistema em todos os seus níveis, gerou uma demanda reprimida por dados censitários na área de educação.

O quadro era ainda pior no que se refere à avaliação: não havia nem oferta nem demanda por esse tipo de informação. A demanda por informações de avaliação tem que ser criada, uma vez que depende, por um lado, da importância atribuída à questão da qualidade do ensino, e, por outro, da percepção de que tais informações podem ser úteis para a tomada de decisões pelos gestores educacionais.

No Brasil, durante décadas, a avaliação foi encarada com extrema desconfiança pela comunidade acadêmica. Pedagogos e educadores não aceitavam a legitimidade e a validade da avaliação em larga escala,¹ freqüentemente interpretada como uma tentativa de controle. Além disso, durante muitos anos as faculdades de educação baniram, por questão de princípio, as abordagens de cunho quantitativo, argumentando que elas impunham uma padronização indevida a contextos plurais e multifacetados.

RSP

Revista do
Serviço
Público

Ano 52
Número 4
Out-Dez 2001

Tereza Cristina Cotta é mestre em Ciência Política pela Universidade Federal de Minas Gerais (UFMG) e membro da carreira de Especialista em Políticas Públicas e Gestão Governamental (EPPGG).

Contato:
[tereza.cotta@
enap.gov.br](mailto:tereza.cotta@enap.gov.br)

Também na esfera do governo, o tema da avaliação despertava pouco interesse. Até fins da década de 70, as políticas educacionais estavam voltadas para a expansão do atendimento e a promoção do acesso ao sistema educacional. A avaliação não fazia parte da agenda pública brasileira, porque a qualidade do ensino ainda não havia emergido como uma questão relevante. Esta conjuntura não era propícia à implantação de sistemas de avaliação em larga escala, que, via de regra, se caracterizam por sua extrema complexidade técnica e por requerer grandes investimentos em pessoal e recursos. A partir dos anos 80, porém, a crescente preocupação dos organismos internacionais com a questão da qualidade do ensino e a implantação de sistemas de avaliação em larga escala em diversos países latino-americanos criaram condições para que o tema adquirisse maior visibilidade no Brasil.

A conjuntura econômica e política também favoreceu, direta ou indiretamente, a área de avaliação. A crise fiscal aumentou a pressão social por mais eficiência e transparência no uso dos recursos públicos. Com o retorno à democracia, a sociedade passou a exigir informações sobre **o que e quanto** os alunos estavam realmente aprendendo. Finalmente, a descentralização promovida pelas reformas educacionais, por meio da transferência de competências e recursos para instâncias intermediárias, levou o governo federal a adotar mecanismos mais flexíveis de controle, voltados para os produtos do sistema educacional e não para seus processos.

Mesmo assim, alguns anos ainda se passariam antes da implantação de sistemas de avaliação em todos os níveis de ensino. De fato, até 1995, não havia um sistema de avaliação do ensino superior, pelo menos no que se refere à dimensão da qualidade dos serviços educacionais ofertados por instituições públicas e privadas. Também não se avaliava o perfil de saída dos alunos e egressos do ensino médio, o que só começou a ser feito em 1999.

O Sistema Nacional de Avaliação da Educação Básica (Saeb), implantado em 1990, é uma exceção à regra. Por meio desta iniciativa pioneira do Ministério da Educação, são coletados, a cada dois anos, dados sobre o desempenho dos alunos brasileiros do ensino fundamental e médio, visando fornecer um diagnóstico dos resultados produzidos pelo sistema educacional.

Somente na segunda metade dos anos 90, quando cresce a demanda por informações de avaliação, foram definidos os arranjos institucionais que viabilizaram a implantação de um sistema de informações educacionais no governo federal, abrangendo duas vertentes: a coleta de dados censitários e de dados de avaliação. Cabe indagar, entretanto, se estas informações estão sendo efetivamente utilizadas no processo de formulação de políticas públicas.

O objetivo deste *paper* é investigar em que medida os dados de avaliação e, em particular, os dados coletados pelo Saeb, têm sido transformados em informações úteis para subsidiar decisões conducentes à melhoria da qualidade da educação brasileira. Buscando responder a esta indagação, este artigo analisa os fatores que, de acordo com a literatura especializada, determinam a utilização dos resultados da avaliação.

A avaliação de resultados educacionais: conceitos e experiências

Conforme mencionado anteriormente, um sistema de informações educacionais abrange duas vertentes: a coleta de dados censitários e a coleta de dados de avaliação. Os dados censitários se referem ao contexto social e demográfico da população estudantil, às condições de oferta dos serviços educacionais, aos indicadores de acesso e participação e aos indicadores de eficiência e rendimento escolar.

Os sistemas de avaliação, por sua vez, produzem diversos tipos de dados, dependendo dos objetivos definidos na etapa de desenho da pesquisa. Há várias modalidades de avaliação e cada uma tem características que definem o uso dado às informações produzidas. As tipologias mais utilizadas pela literatura especializada classificam a avaliação em função do seu *timing* (antes, durante ou depois da implementação da política ou programa),² da posição do avaliador em relação ao objeto avaliado (interna, externa ou semi-independente) e da natureza do objeto avaliado (contexto, insumos, processos e resultados). O objetivo deste *paper* é discutir a avaliação dos resultados educacionais e suas possibilidades de utilização no processo decisório.

A avaliação dos resultados educacionais abrange diferentes objetos de investigação. Seu foco pode recair sobre o aluno, as instituições de ensino ou o sistema educacional como um todo. No primeiro caso, a avaliação situa o desempenho do aluno em relação a um padrão previamente estabelecido, fornecendo um diagnóstico dos conhecimentos, competências e habilidades adquiridos por meio da escola.³ No segundo caso, o desempenho de todos os alunos de determinada instituição permite aferir a qualidade dos serviços educacionais por ela oferecidos.⁴ E, finalmente, quando o objetivo é avaliar os resultados **substantivos** produzidos pelo sistema educacional em seus diversos níveis, é necessário utilizar procedimentos standardizados de avaliação, capazes de fornecer informações pertinentes a todo o universo avaliado.

Portanto, para saber **o que e quanto** os alunos brasileiros aprenderam e quais os fatores associados ao rendimento escolar, é preciso recorrer a um tipo de avaliação consideravelmente mais complexo, do

ponto de vista técnico, do que uma avaliação informal realizada em sala de aula, por exemplo. A partir de pesquisas empíricas sofisticadas, baseadas em procedimentos padronizados de construção dos instrumentos de medida, levantamento e processamento de dados, são produzidos indicadores quantitativos e qualitativos que permitem traçar um panorama dos resultados produzidos pelo sistema educacional, por intermédio da análise do desempenho alcançado por seus alunos. Estes indicadores fornecem informações importantes para o gerenciamento da rede e a formulação, reformulação e avaliação de políticas públicas.

Os sistemas de avaliação criados pelo governo federal a partir de 1995 exemplificam as categorias apresentadas acima. Ainda que cada um deles tenha objetivos, finalidades e características distintas, todos podem ser classificados como avaliações em larga escala, tanto em termos de abrangência quanto de rigor técnico e metodológico.

O Exame Nacional de Cursos (ENC), conhecido como “Provão”, visa subsidiar os processos de decisão e de formulação de políticas de melhoria dos cursos de graduação. A intenção do Ministério da Educação é constituir um sistema de avaliação do ensino superior, formado pela avaliação das condições de oferta, pela avaliação da pós-graduação, realizada pela Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), e pelo próprio ENC. As provas, realizadas anualmente, são elaboradas por comissões de cursos, compostas por especialistas de notório saber, atuantes em cada uma das áreas avaliadas.

O Saeb avalia a qualidade do ensino por meio de testes de desempenho aplicados a uma amostra representativa de alunos das 4^a e 8^a séries do ensino fundamental e da 3^a série do ensino médio. O Saeb aplica também questionários socioeconômicos que permitem a investigação sobre os fatores associados ao rendimento escolar. O objetivo é fornecer indicadores que orientem a elaboração e a revisão de políticas federais e estaduais voltadas para a melhoria da qualidade de ensino.

Finalmente, cabe citar o Exame Nacional do Ensino Médio (Enem), instituído em 1999. O Enem avalia as competências e habilidades desenvolvidas pelo aluno ao final da educação básica, que compreende o ensino fundamental e o ensino médio. O objetivo é informar se ele está preparado para enfrentar os desafios da vida moderna como cidadão autônomo, capaz de decidir, propor e fazer, seja na universidade, seja no mercado de trabalho.

Assim como o ENC e o Saeb, o Enem utiliza procedimentos padronizados de construção dos instrumentos de medida, levantamento e processamento de dados, o que o inscreve na categoria da avaliação em larga escala. Todavia, o propósito deste exame é fornecer *feedback* aos alunos e egressos do ensino médio sobre seus conhecimentos e habilidades. Ao contrário do ENC e do Saeb, o Enem não foi concebido para servir

como instrumento de formulação de políticas públicas. Prova disso é a impossibilidade de fazer generalizações sobre o universo de alunos do ensino médio a partir dos dados do Enem. Como o exame é voluntário, não há uma amostra aleatória e representativa da população-alvo e sim uma amostra autosselecionada de examinandos.

O Enem pode ser definido, portanto, como uma prestação de serviços ao cidadão, uma vez que o resultado do exame permite ao indivíduo avaliar o valor agregado pela escolarização à sua bagagem cultural e intelectual e, com base nestas informações, planejar a sua trajetória profissional e a continuidade de seus estudos. Até o momento, porém, o Enem tem sido utilizado principalmente como mecanismo de acesso ao ensino superior, conjugado ao tradicional concurso vestibular.

Dados os limites deste trabalho, não seria possível analisar em profundidade todos os sistemas de avaliação em larga escala, recentemente implantados pelo governo federal na área de educação. Por isso, apenas um desses sistemas, o Saeb, será submetido a um exame mais minucioso. O objetivo principal é estabelecer uma relação entre as principais características deste sistema e suas possibilidades de influenciar a elaboração de políticas educacionais voltadas para a melhoria da qualidade da educação.

Determinantes da utilização dos resultados da avaliação

A questão da utilização dos resultados de avaliação

A avaliação é um dos mais importantes mecanismos para gerar informações que auxiliem os gestores educacionais a tomar decisões.⁵ O problema é que, de maneira geral, parte-se do pressuposto de que tais informações são automaticamente incorporadas ao processo decisório. Nesta linha de raciocínio, a existência de um sistema de avaliação implicaria, por si mesma, a melhoria da qualidade das políticas do setor. Entretanto, a literatura especializada mostra exatamente o contrário: via de regra, há muitos fatores que dificultam a utilização efetiva dos resultados de avaliação:

“Making use of the findings of evaluations is not an easy task. The history of evaluation may be characterized as one of unfulfilled promises. Therefore, it is only natural that doubts are raised about its overall usefulness. Some see evaluation as a new management fad that creates bureaucracy but delivers

few results. It is sometimes seen as inherently too theoretical to work in practice and having an effect on issues of marginal significance rather than on major policy choices. Others fear increased control or do not wish to be held accountable. Those primarily interested in the continuation of programmes may feel threatened by evaluation. Finally, questions can be raised about the value of evaluation in an environment where performance is continuously measured and even contracted for.” (OECD, 1999: 9).

O trecho acima mostra por que o simples fato de realizar uma avaliação não significa, de forma alguma, que as informações produzidas serão utilizadas. Em primeiro lugar, por uma questão de falta de consenso entre os atores envolvidos. Quase sempre eles relutam em reconhecer a legitimidade e a utilidade da avaliação, seja porque se sentem ameaçados por seus resultados, seja porque não foram envolvidos na definição das regras do jogo. Em segundo lugar, por uma questão de probabilidade: os casos em que a avaliação é simplesmente ignorada são muito mais freqüentes do que aqueles em que a avaliação cumpre sua finalidade.

O principal motivo pelo qual a avaliação geralmente não corresponde às expectativas que gera é a presunção de que as condições para o aproveitamento deste tipo de informação estão dadas. Ocorre justamente o contrário: a avaliação é uma atividade que, por suas próprias características, exige que estas condições sejam intencional e meticulosamente criadas, sob pena de comprometer todo o esforço despendido.

A avaliação envolve, necessariamente, um componente de conflito. Avaliar significa formar um juízo de valor com base na comparação entre uma situação empírica e uma situação ideal.⁶ É natural que aqueles que têm interesses afetados por este tipo de investigação ofereçam resistência. Além disso, há uma série de requisitos que determinam o potencial da avaliação como fonte de informações úteis e confiáveis. Estes requisitos devem ser cumpridos em todas as etapas do processo, desde as definições básicas do desenho da investigação (estabelecimento dos objetivos, metodologia, principais participantes), passando pela estratégia de implementação até a interpretação e divulgação dos resultados.

O caráter aplicado da avaliação

“Evaluation is not knowledge for its own sake, but knowledge for action. It is important that evaluation material be used in decision-making. If not, the administration will soon loose interest and evaluation will become an empty exercise.” (OECD, 1999: 9).

A avaliação é, por definição, pesquisa social aplicada. Seu maior desafio é buscar um equilíbrio entre o rigor metodológico e técnico de uma investigação social e o pragmatismo e flexibilidade necessários a um instrumento de apoio ao processo decisório.

Na prática, porém, é difícil alcançar este equilíbrio. Há uma tensão entre a busca de rigor metodológico e a produção de informações prontamente utilizáveis por aqueles que decidem pela implantação, continuidade ou até mesmo abandono de determinada política ou programa. Os avaliadores vivem, freqüentemente, um dilema: por um lado, eles têm a responsabilidade de garantir a qualidade dos dados de avaliação, o que só é possível quando se utiliza um desenho de investigação rigoroso do ponto de vista técnico e metodológico; por outro lado, eles precisam produzir, a partir dos dados coletados, informações úteis e compreensíveis, de maneira tempestiva e a um custo razoável.

O problema é que quanto mais a investigação se aproxima do método científico, mais complexas e herméticas se tornam as informações produzidas e mais difíceis se tornam para o leigo interpretá-las. É fundamental que o avaliador seja capaz de traduzir estas informações para os seus clientes, os gestores educacionais, que geralmente têm pouca ou nenhuma familiaridade com o jargão técnico.

Saber comunicar os resultados e conclusões da avaliação para o seu público-alvo é imprescindível. Ruídos e falhas de comunicação podem fazer que uma investigação rigorosa do ponto de vista técnico e metodológico seja solenemente ignorada. Esta situação representa o pior dos mundos: a avaliação não agrega valor ao processo decisório, fazendo que as decisões continuem a ser tomadas com base em critérios históricos, conjunturais, subjetivos e até mesmo aleatórios; os recursos públicos investidos na avaliação não oferecem o retorno desejado; e a própria avaliação cai em descrédito, passando a ser vista como uma atividade muito exigente em termos de recursos, mas que oferece poucos benefícios concretos.

Vários são os requisitos para que os resultados de avaliação sejam efetivamente utilizados. Deve haver uma oferta regular e de boa qualidade, de forma a garantir a validade e acuidade das informações produzidas; deve haver demanda por estas informações, o que depende da existência de uma cultura de avaliação e do grau de transparência do processo de formulação de políticas públicas; e, finalmente, devem haver, no mercado, profissionais capazes de “vender” as informações produzidas não apenas para seus clientes diretos como também para o público em geral.

O objetivo das próximas seções é analisar o caso do Saeb à luz das categorias acima, procurando identificar que aspectos do lado da oferta e da demanda e do funcionamento do “mercado” de avaliação educacional facilitam ou dificultam a utilização das informações geradas por esse sistema na formulação e reformulação de políticas públicas.

O Sistema Nacional de Avaliação da Educação Básica (Saeb)

Histórico

No final da década de 80, a melhoria da qualidade do ensino e o aumento da eficiência do sistema educacional se tornaram objetivos prioritários, suplantando a questão do acesso à escola, principal preocupação do período anterior. É neste contexto que surge o Saeb, um projeto ambicioso em termos da sua concepção e de seus objetivos.

O Saeb foi delineado a partir de dois pressupostos básicos: a) o rendimento dos alunos reflete a qualidade do ensino ministrado; e b) nenhum fator determina, isoladamente, a qualidade do ensino. Presume-se que a escolarização formal deva promover a aquisição de um certo conjunto de conhecimentos e habilidades pelos alunos, e que este aprendizado ocorre devido à interação entre diversos fatores. Nesta linha de raciocínio, fica claro que a qualidade da educação tem a ver com o contexto, os insumos e os processos do ensino e aprendizagem. As iniciativas que não atacam simultaneamente estes flancos estariam, portanto, fadadas ao fracasso.

Esta visão influenciou profundamente o desenho do Saeb. Desde o primeiro ciclo de avaliação, em 1990, já havia a preocupação de conjugar os testes de desempenho com questionários socioeconômicos e contextuais. Isto faz sentido quando se considera que os principais usuários das informações coletadas são as unidades gestoras do sistema educacional (Secretarias de Estado do MEC, Secretarias Estaduais e Municipais de Educação). As políticas, programas e projetos governamentais não incidem diretamente sobre o desempenho do aluno, mas atuam sobre as variáveis de contexto, os insumos e os processos. Portanto, é fundamental vincular estas variáveis aos resultados escolares. Ainda que o simples cruzamento destas informações não seja suficiente para que se estabeleçam relações de causalidade, a descoberta dos fatores mais fortemente correlacionados com o desempenho sinaliza as áreas nas quais os gestores educacionais devem intervir para promover melhorias na qualidade do ensino.

Outro aspecto importante é a concepção de sistema de avaliação. Já em 1990, quando o Saeb ainda se chamava Sistema de Avaliação do Ensino Público de 1^a Grau (Saep), a idéia era desencadear um processo de avaliação por meio de levantamentos periódicos de dados, visando captar a evolução do quadro educacional brasileiro. Este conceito está em sintonia com as tendências mais recentes em avaliação, que reconhecem que os processos de mudança social têm um período de maturação relativamente longo, o mesmo se aplicando às intervenções na área de educação.

O Saeb se caracteriza também pela sua abrangência e alcance. No primeiro e no segundo ciclos, registrou-se a participação de 25 e de 26

unidades da Federação, respectivamente. A partir de 1995, o Saeb passou a contar com a adesão de todos os Estados brasileiros. Isto demonstra a preocupação em estabelecer uma relação de parceria entre o governo federal, os Estados e os municípios, os quais, junto aos demais atores do processo educativo (administradores, diretores, professores, alunos, pais, comunidade etc.), partilham a responsabilidade pela qualidade da educação brasileira.

Ao longo de seus dez anos de existência, o Saeb passou por diversas modificações. As maiores mudanças ocorreram a partir de 1995, com a elaboração das **Matrizes Curriculares de Referência para o Saeb**⁷ e a utilização de novas técnicas para a aferição do desempenho dos alunos, como a amostragem matricial de conteúdos e a Teoria da Resposta ao Item (TRI). A seção seguinte discute as principais características do Saeb a partir deste período.

As principais características do Saeb

O objetivo geral do Saeb é monitorar as políticas voltadas para a melhoria da qualidade, equidade e eficiência do sistema educacional brasileiro no âmbito da educação básica. A cada dois anos, são aplicados testes de rendimento a alunos da 4^a e 8^a séries do ensino fundamental e da 3^a série do ensino médio, nas várias disciplinas que integram as propostas curriculares das 27 unidades da Federação. Tais testes visam aferir “o que o aluno sabe e é capaz de fazer”, isto é, o conhecimento, o nível de desenvolvimento cognitivo e as habilidades instrumentais adquiridas pelo aluno na sua passagem pela escola. O foco do Saeb é o sistema educacional, e não o aluno individualmente considerado.

O Saeb conta com dois instrumentos de avaliação, os testes de rendimento, com base nos quais se estima a proficiência dos examinandos,⁸ e os questionários socioeconômicos, que investigam fatores contextuais que incidem sobre o processo de ensino e aprendizagem. Estes fatores aparecem agrupados em quatro áreas de observação: escola, gestão escolar, professor e aluno. O questionário da escola indaga sobre os equipamentos disponíveis e as características físicas das escolas; o questionário do diretor identifica o modelo de gestão escolar; o questionário do professor descreve o perfil do corpo docente e as práticas pedagógicas e o questionário do aluno enfoca as características socioeconômicas e culturais do aluno, bem como seus hábitos de estudo.

Desde o primeiro levantamento de dados, o Saeb avalia alunos em Matemática e Língua Portuguesa. No Saeb 97, a área de Ciências foi incorporada — desmembrada, na 3^a série do Ensino Médio, em Física, Química e Biologia. No Saeb 99, foram aplicadas também provas de História e Geografia. A divulgação do Saeb 99 privilegiou os resultados

em Matemática e Língua Portuguesa, disciplinas que constituem uma série histórica que mostra a evolução do sistema ao longo dos últimos anos. O Saeb 2001 testou o rendimento dos alunos apenas em Língua Portuguesa e Matemática, consolidando a tendência de focar o exame apenas nessas duas disciplinas.

A população de referência do Saeb é o universo de alunos brasileiros matriculados na 4ª e 8ª séries do ensino fundamental e na 3ª série do ensino médio. Devido à magnitude deste contingente de alunos, a cada levantamento de dados avalia-se uma amostra **nacionalmente representativa** em todas as unidades da Federação. O plano amostral define os níveis de agregação dos dados, de modo a garantir a precisão desejada para os resultados, que pode ser calculada em função de uma ou mais variáveis. Para cada valor da média de proficiência calcula-se uma medida da precisão desta estimativa, que é o **erro padrão**.

A metodologia de amostragem do Saeb tem variado sensivelmente de um ciclo de levantamento de dados para outro. Em 1997, por exemplo, a amostra foi estratificada por Zona (rural ou urbana), Localização (capital ou interior) e Dependência Administrativa (estadual, municipal e particular). Não fizeram parte da população pesquisada a zona rural da Região Norte, as escolas federais, os alunos de cursos profissionalizantes do ensino médio e os alunos de turmas multisseriadas no ensino fundamental.

Até 1997, o plano amostral impunha uma importante limitação à interpretação dos resultados do Saeb. Como a amostra era representativa apenas da população de alunos, as inferências sobre as escolas e professores sempre se referiam a esta unidade de análise. Por exemplo, os dados sobre as condições das bibliotecas das escolas incluídas na amostra tinham que ser interpretados da seguinte maneira: “38,71% dos alunos da 4ª série do Ensino Fundamental estudam em escolas com biblioteca em bom estado de conservação”. Portanto, na prática, não havia como utilizar adequadamente os dados coletados por meio dos questionários contextuais.

Para solucionar este tipo de problema, o plano amostral do Saeb 1999 foi aperfeiçoado em três sentidos. Em primeiro lugar, houve uma expressiva ampliação do número de alunos e escolas participantes da avaliação, garantindo maior precisão dos resultados. Em segundo lugar, passou a ser possível a análise dos dados em três níveis — alunos, diretores e escolas, ainda que, devido ao desenho da pesquisa, não seja possível fazer inferências individuais sobre os elementos do universo avaliado. Isto representou um grande avanço, tendo em vista que possibilitou uma melhoria na caracterização do sistema educacional e na elaboração de perfis dos alunos avaliados e das escolas que freqüentam. Finalmente, na maioria das unidades da Federação, os dados de desempenho foram divulgados por estratos de interesse — Dependência Administrativa (estadual, municipal, particular) e Localização (urbana e rural).

O desenho amostral do Saeb 1999 prevê a exclusão dos alunos das escolas rurais da 8ª série do ensino fundamental e da 3ª série do ensino médio, tendo em vista o tamanho reduzido do universo excluído e o pequeno impacto de tal exclusão nas médias de desempenho das populações de interesse. Apenas no caso da 4ª série do ensino fundamental, nos Estados onde a matrícula nas escolas da zona rural representa uma parcela significativa da matrícula total, estes alunos foram incluídos na amostra.⁹ O plano amostral do Saeb 2001 manteve as mesmas exclusões dos levantamentos anteriores.

A divulgação do Saeb

A divulgação do Saeb 1999 foi cercada de grande expectativa. Pela primeira vez, dispor-se-ia de uma série histórica de resultados, permitindo a identificação das principais tendências de evolução da Educação Básica no período 1995/1999. A análise de tais tendências possibilitaria aos gestores educacionais, em suas diferentes esferas de atuação, definir ações voltadas para a correção das distorções identificadas em seus sistemas de ensino.

Em virtude destas expectativas e da experiência com os ciclos anteriores de avaliação, buscou-se em 1999 aprimorar a estratégia de divulgação em dois aspectos: a adequação dos produtos às necessidades de diversas audiências e a busca de um maior envolvimento das instâncias e atores relevantes na definição da estratégia de divulgação. Em vez de concentrar em um único documento toda a informação pertinente aos aspectos descritivos, analíticos, técnicos e metodológicos da pesquisa, optou-se pela elaboração de três tipos de produtos: os informes de resultados, a Sinopse Estatística do Saeb 1999 e os relatórios metodológicos.

O envolvimento dos diversos atores do sistema educacional afetados pelos resultados do Saeb na definição da estratégia de divulgação é condição fundamental para a legitimação do sistema e a utilização das informações produzidas. O reconhecimento deste fato fez que se buscasse estabelecer parcerias com os principais interlocutores do governo federal, por meio da criação de fóruns de discussão,¹⁰ mecanismos e estratégias que permitissem a esses atores expressar suas preocupações e receber um tratamento diferenciado em termos do acesso e do nível de detalhamento das informações fornecidas.

O potencial de utilização dos resultados do Saeb

A utilização dos resultados da avaliação não acontece de maneira automática. Ao contrário, ela depende de uma conjunção favorável de fatores relacionados com o lado da oferta e da demanda e com a própria dinâmica do mercado de avaliação.

A análise da experiência do Saeb corrobora estas afirmações. Em seus dez anos de existência, o sistema avançou sob diversos aspectos. Após seis levantamentos de dados, é uma iniciativa consolidada, que mobiliza todo o sistema, contribuindo para a criação de uma cultura de avaliação no país. Além disso, a partir de 1995, o sistema passou a utilizar técnicas sofisticadas de mensuração do rendimento dos alunos, equiparando-se aos melhores sistemas de avaliação internacionais.

Porém, os desafios a serem enfrentados pelo Saeb talvez sejam tão grandes quanto suas realizações. O primeiro é garantir a efetividade no atendimento às demandas de informação do sistema educacional. Para superar este desafio, o Saeb precisa identificar quem são seus usuários preferenciais e se manter permanentemente focado em seus interesses e necessidades. Isto nos leva ao segundo desafio, que é o de construir e institucionalizar relações de parceria entre os governos federal, estadual e municipal, envolvendo ativamente estes atores em todas as etapas da avaliação, desde as decisões estratégicas sobre os rumos do sistema até a sua implementação.

O Saeb só conseguirá dar um salto de qualidade, entretanto, quando as razões que levaram a este estado de coisas forem melhor investigadas. Por que um sistema de avaliação que foi objeto de tamanhos investimentos em recursos materiais e humanos ainda não foi institucionalmente incorporado ao processo decisório da área? O componente político e institucional do sistema está sendo adequadamente equacionado?

A metáfora do “mercado de avaliação educacional” pode ajudar a encontrar respostas para algumas destas indagações. Quais são os principais problemas apresentados pelo Saeb sob a ótica da oferta e da demanda de informações de avaliação?

No lado da oferta, destacam-se as seguintes questões: arranjos institucionais favoráveis às atividades de avaliação; envolvimento dos *stakeholders* no processo; definição dos objetivos de avaliação; adequação do desenho da avaliação; implementação da avaliação e estratégia de divulgação dos resultados. No lado da demanda, há de se considerar a cultura de avaliação e a qualidade da demanda.

O lado da oferta

Arranjos institucionais

Como enfatiza a literatura especializada, a forma de organização das atividades de avaliação influencia o seu papel no processo decisório:

“It is often claimed that a certain level of institutionalization has to occur before evaluations can play a genuine role in the management of government organizations. Institutionalization

aims at creating conditions for a high-quality supply of evaluations, the utilization of the evaluation findings, and organizational learning.” (OECD, 1999: 26)

Neste sentido, a avaliação educacional brasileira se encontra em uma situação privilegiada. A reestruturação do Instituto Nacional de Estudos e Pesquisas Educacionais (Inep), em 1997, foi decisiva para a implantação de um sistema de informações educacionais que abrangesse a coleta de dados censitários e de avaliação. Na qualidade de autarquia vinculada ao Ministério da Educação, o Inep goza de autonomia administrativa e financeira, condição fundamental para a mobilização dos recursos materiais e humanos necessários à implantação de sistemas de avaliação de ponta.

O Saeb foi instituído no final da década de 80, bem antes, portanto, da reformulação do Inep. Não obstante, o arranjo institucional acima mencionado foi decisivo para a consolidação do sistema. Foram as reformulações técnicas e metodológicas feitas a partir de 1995 que imprimiriam, ao sistema, a feição que ele hoje ostenta.

A literatura aponta, também, a relação entre o tipo de arranjo institucional e o grau de sistematização dos mecanismos de monitoramento de desempenho. No caso em questão, a existência de uma instituição especificamente voltada para a produção e disseminação de informações educacionais moldou um sistema de avaliação com as seguintes características: a) formal, isto é, de caráter compulsório¹¹ e periodicidade regular; b) centralizado, do ponto de vista da responsabilidade pela sua concepção e implementação; e c) externo, sob a perspectiva da posição do avaliador em relação ao objeto investigado.¹²

Porém, esta forma de organização das atividades de avaliação também apresenta desvantagens. Por um lado, o relativo insulamento da instituição responsável pelo sistema diante dos demais atores do sistema educacional garante o acesso aos recursos — materiais e de poder — necessários à implementação de um projeto desta envergadura. Por outro lado, este mesmo insulamento dificulta a articulação com atores relevantes, o que afeta a legitimidade da avaliação e as chances de uso efetivo das informações produzidas.

Envolvimento dos atores relevantes

É lugar comum dizer que os atores que têm seus interesses afetados pelo sistema de avaliação devem participar ativamente de todas as etapas do processo. Entretanto, este preceito não é facilmente transportado para a prática. Um dos maiores desafios do Saeb ainda é criar mecanismos que efetivamente garantam a representação dos atores relevantes na definição das principais características da investigação.

O primeiro passo é identificar **quem** são os atores relevantes. Pode-se dizer, utilizando o conceito de forma abrangente, que estes atores não são apenas aqueles diretamente atingidos pela divulgação dos resultados da avaliação, os gestores educacionais. Os atores mobilizados durante a implementação do sistema, como, por exemplo, a equipe técnica das Secretarias Estaduais de Educação, também deveriam ser alçados à condição de interlocutores privilegiados, mesmo porque o apoio desta equipe é fundamental nas etapas de amostragem e de levantamento de campo.

O caso dos municípios e das escolas também mereceria atenção especial. Como o desenho da amostra não permite que se façam inferências individuais nestes níveis de agregação, cria-se uma situação muito peculiar. A cada dois anos municípios, escolas e alunos incluídos na amostra são mobilizados para aplicar as provas do Saeb. Porém, em função das características da avaliação, focada no sistema educacional, eles nunca se reconhecem nos resultados divulgados, que se referem aos Estados e ao Brasil.

Mesmo no que diz respeito aos atores tradicionalmente reconhecidos, os Secretários Estaduais de Educação, ainda não se alcançou a situação desejada. O papel de instâncias como o Comitê Consultivo do Saeb (Comsaeb) poderia ser ampliado, de forma a dar mais transparência e legitimidade às decisões sobre características estratégicas do sistema. Deve-se considerar, inclusive, a possibilidade de criar foros deliberativos, instituindo um modelo de gestão compartilhada do sistema.

A despeito da dificuldade de equilibrar interesses díspares, e, muitas vezes contraditórios, é fundamental adotar uma abordagem participativa no gerenciamento de sistemas de avaliação desta envergadura. O estabelecimento de parcerias com os atores-chave é essencial para a efetiva utilização das informações produzidas. Caso contrário, o sistema fica reduzido à sua dimensão formal, e a tendência é que seja encarado com indiferença ou até mesmo com hostilidade. A consulta e o envolvimento dos atores relevantes é condição *sine qua non* para a legitimidade e a efetividade dos sistemas de avaliação.

Definição dos objetivos de avaliação

A definição dos objetivos da avaliação é uma etapa de fundamental importância. Dela depende o desenho da população-alvo, a seleção das metodologias a serem utilizadas, a elaboração do plano de análise dos dados e a estratégia de disseminação dos resultados. Problemas nesta etapa geram um “vício de origem” que dificilmente pode ser contornado nas etapas posteriores do processo avaliativo.

Definir o que se quer avaliar é uma tarefa difícil, particularmente na ausência de uma cultura de avaliação consolidada. De modo geral, não

se sabe que tipo de informação a avaliação fornece nem como ela pode ser utilizada na melhoria do sistema educacional. Mesmo nestas condições desfavoráveis, é imprescindível ouvir os usuários diretos e potenciais das informações quando da definição do objeto e do propósito da avaliação. Os interesses e necessidades desses atores devem ser levados em conta, caso se queira aumentar a efetividade do sistema.

Após dez anos de existência, é possível que o Saeb não esteja mais atendendo às demandas de informação de seus principais usuários. As profundas mudanças que transformaram o cenário educacional brasileiro nos últimos anos certamente alteraram as prioridades dos gestores do sistema, bem como o seu entendimento da questão da qualidade da educação. Neste contexto, seria necessário estabelecer um novo pacto com os atores relevantes, por meio da instituição de mecanismos de consulta e da criação de foros permanentes de participação. Desta forma, se garantiria a ligação do Saeb com o processo decisório, permitindo que o sistema cumprisse o seu objetivo primordial.

Adequação do desenho da avaliação

Definidos o objeto e o propósito da avaliação, elabora-se um plano de investigação que viabilize a coleta de dados de boa qualidade, que servirão como base para a produção de informações relevantes, acuradas, válidas e confiáveis. O desenho da avaliação se refere às características técnicas e metodológicas da investigação, materializadas no plano amostral, na elaboração dos instrumentos de coleta de dados, nos procedimentos para construção e interpretação das escalas de desempenho e na análise das variáveis incluídas nos questionários contextuais.

A metodologia de amostragem do Saeb tem variado sensivelmente de um ciclo de levantamento de dados para outro. Ainda que os planos amostrais tenham sido sucessivamente aprimorados, essas variações limitam as possibilidades de análise da série histórica de resultados, fazendo que as expectativas dos usuários não possam ser plenamente atendidas.

Entretanto, o principal problema do desenho do Saeb é a ausência de um referencial teórico para a interpretação das estimativas de desempenho. Avaliar é, por definição, comparar a situação real com a situação ideal, o que só pode ser feito se houver um padrão mínimo de desempenho considerado desejável. Segundo a literatura especializada, uma boa avaliação deve ter critérios relevantes, apresentar evidências adequadas e produzir conclusões confiáveis e claras.

A experiência internacional mostra que a definição de *standards* de desempenho não é, de forma alguma, uma questão trivial. Trata-se de um tema politicamente sensível, capaz de gerar controvérsias que ultrapassam as fronteiras do sistema educacional. No Brasil, esta discussão é

ainda mais difícil, devido à inexistência de um currículo nacionalmente válido. Com a Lei de Diretrizes e Bases da Educação Nacional (LDB), de 1996, Estados, municípios e até mesmo estabelecimentos de ensino passaram a ter autonomia para definir suas propostas curriculares. Neste contexto, é natural que haja dificuldade em criar consenso sobre os conhecimentos e habilidades que os alunos brasileiros devem demonstrar em determinados estágios do ensino fundamental e médio.

A solução encontrada pelo Saeb foi buscar o que havia de comum nas propostas curriculares estaduais e, com base neste referencial, definir o que seria avaliado nos testes de desempenho. Surgiram assim as **Matrizes Curriculares de Referência**, documento cuja elaboração iniciou-se em 1996, por meio de uma “ampla consulta nacional sobre conteúdos praticados nas escolas brasileiras de ensino fundamental e médio e a reflexão de professores, pesquisadores e especialistas sobre a produção científica em cada área que se torna objeto de conhecimento escolar.” (Pestana et al., Inep, 1999)

As matrizes, entretanto, não conseguem suprir a falta de um modelo de ensino e aprendizagem e de padrões de desempenho claramente estabelecidos. Devido à grande distância entre o que consta das propostas curriculares estaduais, o que é efetivamente ensinado em sala de aula e o que os alunos realmente aprendem fica difícil de dizer “com que régua” se deve medir o desempenho dos alunos, ainda mais considerando a grande heterogeneidade que caracteriza o sistema educacional brasileiro.

Pode-se dizer que o Saeb ainda carece de critérios de avaliação propriamente ditos, o que compromete seriamente as perspectivas de efetiva utilização dos resultados. Sem padrões de avaliação, a interpretação das estimativas de desempenho acaba se restringindo a comparações entre diferentes segmentos da população de alunos, como os *rankings* que ordenam os Estados por ordem decrescente de desempenho. Na falta de um referencial teórico, uma avaliação referida a critério, que tem função diagnóstica, é indevidamente transformada em uma avaliação referida à norma, que tem função de discriminar os grupos de alto e baixo desempenho.

A questão dos *standards* de avaliação só será adequadamente enfrentada por meio de um amplo processo de discussão. É tarefa de toda a sociedade definir os conhecimentos, habilidades e competências que o sistema educacional deve desenvolver nos alunos, para promover sua inserção no mercado de trabalho e criar condições para o exercício pleno da cidadania.

A questão da implementação

A avaliação educacional em larga escala supõe a padronização dos procedimentos de coleta e processamento dos dados. Portanto, a forma

como o sistema de avaliação é implementado influencia a qualidade e a disponibilidade de informações. Falhas operacionais podem ocasionar problemas sérios como um número muito alto de respostas faltantes (*missing*), inviabilizando inferências sobre determinados estratos da amostra. Equívocos na administração dos testes levam ao expurgo dos dados incorretos e podem chegar a comprometer completamente as estimativas de desempenho.

Por estes motivos, a logística das operações de campo merece especial atenção. Em um sistema de abrangência nacional como o Saeb, a capacidade de planejamento é um fator crítico para o sucesso da avaliação. A participação das equipes técnicas das Secretarias Estaduais de Educação, nesta etapa, é fundamental, não apenas devido à sua maior proximidade das escolas amostradas, mas também como uma forma de envolvê-las ativamente no processo, o que não acontece quando elas se limitam a cumprir orientações emitidas pelo órgão central.

Além disso, como via de regra as operações de campo são terceirizadas, é importante definir pormenorizadamente especificações para todas as etapas operacionais, com especial ênfase no treinamento das equipes responsáveis pela preparação e aplicação dos testes. E, para evitar problemas futuros, deve-se acompanhar passo a passo a **execução** dessas etapas. A adoção de mecanismos independentes de controle de qualidade é sempre uma medida recomendável.

A divulgação dos resultados

O calcanhar-de-aquiles do Saeb é, sem dúvida, a estratégia de divulgação dos resultados. Neste campo, há problemas de todos os tipos: a) dificuldade em identificar os principais usuários; b) inadequação dos produtos de divulgação para as audiências visadas; c) problemas no *timing* da divulgação; d) problemas de articulação com os atores do sistema educacional; e e) dificuldade de adequar as informações para a linguagem da mídia.

De certa forma, a falta de efetividade na comunicação das evidências do Saeb tem a ver com o tipo de informação produzida. A divulgação das estatísticas do Censo Escolar, por exemplo, é uma tarefa muito mais simples. Os dados censitários, como o próprio termo indica, referem-se a todos os elementos do universo considerado (escolas, professores, população estudantil). Por isso, o que se divulga são estatísticas descritivas, diretamente interpretáveis, não requerendo do leitor habilidades além da leitura de quadros e tabelas. Os dados gerados por sistemas de avaliação, por sua vez, geralmente são pouco amigáveis, exigindo o manejo de conceitos sofisticados relacionados com a amostragem (amostra aleatória e representativa, estratos, erro amostral) e com o entendimento das escalas de desempenho (noção de escala, análise de tendências etc.).

Por outro lado, os problemas de comunicação se devem à dificuldade de interpretação dos resultados. As razões para isso vão desde a já mencionada falta de critérios claros e consensuais para contrastar o empírico com o ideal até contradições mais profundas do modelo de avaliação.

O objetivo do Saeb é fornecer um diagnóstico da qualidade do ensino ministrado em todo sistema educacional. Porém, a forma como o sistema está estruturado, tendo em vista a autonomia política e administrativa de Estados e municípios e a superposição de competências entre as diferentes esferas de governo, dificulta a interpretação dos resultados nesse nível de agregação. Por exemplo: ao analisar o desempenho dos alunos de determinado Estado, como separar a influência das políticas estaduais, as iniciativas dos gestores municipais e as ações no âmbito da escola? É evidente que o gestor estadual não tem governabilidade sobre todas as ações que incidem sobre a qualidade do ensino ministrado. Esta situação paradoxal faz que os usuários das informações de avaliação não se reconheçam nos resultados obtidos, além de dificultar a responsabilização dos gestores públicos perante a sociedade.

Até o ciclo de 1997, por limitações da amostra do Saeb, os secretários estaduais não tinham elementos para estabelecer umnexo entre sua atuação diante das Secretarias de Educação e os resultados alcançados no exame. A partir de 1999, o desenho amostral passou a permitir estimativas de desempenho por rede (estadual e municipal) em quase todas as unidades da Federação, apesar de continuarem não fornecendo resultados por escola e por turma. Este foi um importante avanço, o qual deveria ser mantido nos próximos ciclos de avaliação.

Mesmo sabendo que a educação é, por definição, um fenômeno social influenciado por múltiplos fatores, internos ou externos à escola e ao sistema educacional, deve-se admitir que, em níveis mais altos de agregação, é quase impossível explicar o desempenho sem lançar mão de procedimentos estatísticos sofisticados. Este fato torna a interpretação dos resultados de avaliações como o Saeb um exercício analítico complexo, dado que não há uma explicação única para as estimativas e as tendências de evolução do desempenho.

Por tudo o que foi exposto, é fácil compreender por que, após dez anos de existência, o Saeb ainda enfrenta dificuldades em produzir informações oportunas, pertinentes, significativas e acessíveis ao grande público. Como este é um desafio comum a todas as iniciativas de avaliação educacional, é fundamental buscar nas experiências internacionais subsídios para aprimorar continuamente nossa capacidade de comunicar os resultados da avaliação, condição fundamental para a sua plena utilização.

O lado da demanda

A cultura de avaliação

Por mais que se tenha avançado, nos últimos anos, na área de avaliação educacional, ainda não se pode dizer que há uma cultura de avaliação no Brasil. De fato, mesmo em países que têm tradição na área, são necessários muitos anos antes para que a cultura de avaliação esteja verdadeiramente consolidada. Pode-se dizer que isto aconteceu com: a) apoio à avaliação por parte de atores que ocupam posições-chave no governo ou em outras esferas do sistema educacional; b) demanda regular pelas informações de avaliação em todos os níveis do sistema; c) registro de experiências bem-sucedidas de utilização das informações de avaliação para subsidiar o processo decisório; d) profissionalização das atividades de avaliação nas esferas pública e privada; e e) consolidação da avaliação educacional como campo de investigação acadêmica.

A criação de uma cultura de avaliação depende, portanto, de um processo de aprendizado coletivo que, no Brasil, foi recentemente iniciado. Para garantir a sua continuidade, é essencial continuar investindo no aprimoramento dos sistemas existentes no tocante à interpretação e divulgação de seus resultados. Só assim ficará patente para a sociedade o valor da avaliação como mecanismo de apoio ao processo decisório.

A qualidade da demanda

É importante incrementar a demanda por avaliação não apenas do ponto de vista quantitativo, mas também qualitativo. A qualidade da demanda depende da disseminação do conhecimento sobre o que é e para que serve a avaliação. Este conhecimento permite a definição, de forma participativa, de objetivos claros para o sistema, criando condições favoráveis à utilização das informações produzidas. Cria-se assim um círculo virtuoso que impulsiona a expansão das atividades de avaliação.

Devido à falta de interlocutores na academia, no Brasil a qualificação da demanda tem frequentemente ocorrido na prática por meio de cooperação técnica vertical e horizontal, ou por meio de eventos nos quais as experiências de avaliação educacional são compartilhadas e submetidas ao escrutínio de especialistas e gestores de projetos na área. Entretanto, apenas a criação de centros de pesquisa em avaliação educacional e a capacitação contínua de recursos humanos em avaliação propiciará a formação da massa crítica necessária ao aprimoramento dos sistemas de avaliação existentes.

As expectativas dos *stakeholders*

As expectativas dos principais usuários são determinantes para o aproveitamento das informações de avaliação no processo decisório. É preciso que eles saibam exatamente para que serve um estudo desta natureza e que estejam conscientes das limitações impostas pelo desenho da investigação.

No caso do Saeb, por exemplo, o desenho da amostra impede que sejam feitas inferências sobre municípios, escolas e turmas. Também há restrições na interpretação da série histórica de resultados. E, principalmente, deve-se ter clareza de que o propósito do SAEB é fornecer um diagnóstico do desempenho dos alunos brasileiros, e não propriamente **explicar** os resultados. Como foi mencionado anteriormente, o estabelecimento de nexos de causalidade requer o uso de metodologias estatísticas sofisticadas, como os modelos hierárquicos. O que pode e deve ser feito é planejar a inclusão deste tipo de estudo no plano de disseminação dos resultados.

Considerações finais

O impacto da avaliação no processo decisório depende da capacidade de avaliar e da disposição em utilizar os resultados da avaliação. A análise da experiência do Saeb mostra que não basta cuidar da qualidade técnica e metodológica do estudo; é necessário um esforço deliberado para promover a incorporação das informações de avaliação no ciclo de políticas públicas.

Isto é especialmente verdadeiro na ausência de uma cultura de avaliação consolidada. Os gestores educacionais geralmente desconhecem o potencial da avaliação como instrumento para o desenho e a revisão de políticas. Além disso, a lógica do processo decisório na administração pública brasileira torna-o, até certo ponto, impermeável a informações “objetivas”. A formulação de políticas tende a acontecer em um ambiente de baixa racionalidade, em que critérios históricos de alocação de recursos e cálculos políticos ainda têm peso considerável.

Contudo, é importante saber reconhecer os tipos de decisão nos quais as informações de avaliação podem ser úteis. Variáveis como o *timing* do processo decisório, a existência de cursos alternativos de ação disponíveis e o grau de transparência das decisões devem ser consideradas. De maneira geral, o impacto das informações de avaliação é cumulativo, ou seja, tais informações nem sempre provocam mudanças imediatas em políticas e programas, mas elas servem, via de regra, para aumentar o conhecimento acumulado sobre uma dada situação ou forma de intervenção na realidade social.

Para que a avaliação efetivamente direcione o processo decisório, portanto, é necessário que três condições sejam cumpridas: disponibilidade de informações de boa qualidade (relevantes, acuradas, válidas e confiáveis); estabelecimento de canais de comunicação permanente com os principais usuários; e capacidade de adaptação ao ambiente político e burocrático.

Notas

- ¹ A avaliação em larga escala se caracteriza pela estandardização dos procedimentos de coleta e de processamento de dados.
- ² A literatura distingue também a avaliação formativa, utilizada para monitorar o processo de aprendizagem durante a instrução, e a avaliação somativa, realizada ao final de um curso ou unidade de instrução, para determinar a extensão em que os objetivos foram alcançados ou para verificar se os alunos dominaram aspectos específicos do aprendizado.
- ³ Esta definição é de uma avaliação referida a critério, em que os resultados são interpretados em função do domínio delimitado de competências a que se refere a avaliação e da existência de um nível de desempenho previamente definido como satisfatório ou insatisfatório. Em contrapartida, a avaliação referida à norma interpreta os resultados com base em normas obtidas em amostras representativas de indivíduos (Pasquali, 1999).
- ⁴ A avaliação institucional abrange outras dimensões como a infra-estrutura das instituições escolares, os processos de gestão, formação, qualificação e produtividade dos recursos humanos, entre outros.
- ⁵ Há vários outros mecanismos para obtenção de *feedback* sobre políticas e programas públicos, como os sistema de auditoria, monitoramento do desempenho e análise política (OECD, 1999).
- ⁶ Avaliação e mensuração são conceitos distintos, ainda que intimamente relacionados. Avaliação é um termo mais abrangente e inclusivo do que mensuração, que é estritamente quantitativo e não envolve julgamentos de valor.
- ⁷ Ver seção. “Definição dos objetivos de avaliação”.
- ⁸ O Saeb se baseia em um modelo analítico que tem a proficiência como variável dependente (“conjunto de habilidades, apresentado em uma escala contínua, evidenciadas pelo desempenho dos alunos nas disciplinas específicas: o que sabem, compreendem e são capazes de fazer.”). Ver: Saeb/95: Relatório Final. Brasília, Inep, 1998.
- ⁹ Informações detalhadas sobre o desenho e o comportamento da amostra poderão ser obtidas no Relatório Metodológico da Amostra do Saeb, 1999.
- ¹⁰ Como o Comitê Consultivo do Saeb — Comsaeb, formado por representantes do Instituto Nacional de Estudos e Pesquisas Educacionais (Inep), por Secretários Estaduais indicados pelo Conselho Nacional de Secretários de Educação (Consed), pelos titulares das secretarias finalísticas do MEC e por membros da União Nacional dos Dirigentes Municipais de Educação (Undime).
- ¹¹ O termo compulsório deve ser utilizado com cautela neste contexto, uma vez que a adesão das unidades da Federação ao Saeb é voluntária.

¹² Em contrapartida, o Saeb apresenta uma característica típica de sistemas informais de avaliação, o foco no processo decisório do poder executivo e não na prestação de contas ao Poder Legislativo (OECD, 1999: 26).

Referências bibliográficas

- AGUILLAR, M.J. e ANDER-EGG, E. (1994), *Avaliação de serviços e programas sociais*. Petrópolis, RJ: Vozes.
- CHEIBUB FIGUEIREDO, A. e FIGUEIREDO, M. (1986), “Avaliação política e avaliação de políticas: um quadro de referência teórica”. *Análise e Conjuntura*, Belo Horizonte: Fundação João Pinheiro, v. 1, nº 3, set/dez., p. 107-127.
- COHEN, E. e FRANCO, R. (1993), *Avaliação de projetos sociais*. Petrópolis, RJ: Vozes.
- EBAP/FGV. (1992), “Projeto: avaliação de programas públicos” (Síntese do documento produzido pelo *General Accounting Office* — GAO/PMED, 10.1.2, march 1992. “The Evaluation Synthesis”, Rio de Janeiro).
- GODFREY, M. (1987), “Monitoring the utilization of educated manpower”, in *Monitoring Labour Markets*, International Labour Organization.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS. (1999), *Estudo Comparativo dos Resultados do SAEB 1995 e 1997*.
 _____ . (1998), *SAEB 95: Relatório final*. Brasília.
 _____ . (1998), *SAEB 97: Primeiros resultados*. Brasília.
- OECD. (1999), “Improving evaluation practices: best practices guideline for evaluation and background paper”. *PUMA/PAC (99)1*, Paris.
- ONU, Grupo do CAC. (1984), *Seguimiento y evaluación. Pautas básicas para el desarrollo rural*. Roma: FAO.
- PASQUALI, L. (ed.) (1999), *Instrumentos psicológicos: manual prático de elaboração*. LabPAM/IBAPP, Brasília.
- PESTANA, M. I. G. et. al. (1999), *Matrizes curriculares de referência para o SAEB*. 2ª ed. rev. ampl. — Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais.
- POSAVAC, E.J. e CAREY, R.G. (1992), *Program evaluation: methods and case studies*. Englewood Cliffs, New Jersey: Prentice Hall, 4ª ed.
- ROSSI, P. e FREEMAN, H. (1989), *Evaluación: un enfoque sistematico*. Sage Publications Inc., caps. IV e V.
- SECRETARIA DE DESENVOLVIMENTO, INOVAÇÃO E AVALIAÇÃO EDUCACIONAL/MINISTÉRIO DA EDUCAÇÃO E DO DESPORTO. (1995), *Relatório do 2º ciclo do Sistema Nacional de Avaliação da Educação Básica (SAEB) 1993*. Brasília.
- SECRETARIA NACIONAL DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA/MINISTÉRIO DA EDUCAÇÃO. (1991), *Relatório preliminar dos Sistema de Avaliação do Ensino Público de Primeiro Grau — Ciclo de 1990*. Brasília.
- SULBRANDT, J. (1993), *La evaluación de los programas sociales: una perspectiva crítica de los modelos usuales*. CLAD.
- VIANA, A.L. (1996), “Abordagens metodológicas em políticas públicas”. *RAP*, v.3, nº 2, mar/abr., p. 5-43.

Resumo
Resumen
Abstract

RSP

Revista do
Serviço
Público

Ano 52
Número 4
Out-Dez 2001

Avaliação educacional e políticas públicas: a experiência do Sistema Nacional de Avaliação da Educação Básica (Saeb)

Tereza Cristina Cotta

O objetivo do *paper* é investigar em que medida os dados de avaliação e, em particular, os dados coletados pelo Sistema Nacional de Avaliação da Educação Básica (Saeb), têm sido transformados em informações úteis para subsidiar decisões conducentes à melhoria da qualidade da educação brasileira. Buscando responder a esta indagação, este artigo analisa os fatores que, de acordo com a literatura especializada, determinam a utilização dos resultados da avaliação.

Evaluación educacional y políticas públicas: la experiencia del Sistema Nacional de Evaluación de la Educación Básica — Saeb

Tereza Cristina Cotta

El objetivo del *paper* es investigar en qué medida los datos de evaluación y, en particular, los datos recolectados por el Sistema Nacional de Evaluación de la Educación Básica (Saeb), han sido transformados en informaciones útiles para subsidiar decisiones que conduzcan a la mejoría de la calidad de la educación brasileña. Buscando responder a esta indagación, este artículo analiza los factores que, de acuerdo con la literatura especializada, determinan la utilización de los resultados de la evaluación.

Educational evaluation and public policies: the National System of Basic Education Evaluation experience

Tereza Cristina Cotta

The objective of this paper is to investigate to what extent the evaluative data and, particularly, the data collected by the National System of Basic Education Evaluation have been used to inform decisions aiming to improve the quality of Brazilian education. Trying to answer this question, this paper analyses the factors that, according to the specialized literature, determine the utilization of the evaluation results in the decision making process.

Tereza Cristina Cotta é mestre em Ciência Política pela Universidade Federal de Minas Gerais (UFMG) e membro da carreira de Especialista em Políticas Públicas e Gestão Governamental (EPPGG).

Contato:
tereza.cotta@enap.gov.br