

Avaliação de sistemas educacionais no Brasil

BERNARDETE A. GATTI

gatti@fcc.org.br

Fundação Carlos Chagas, São Paulo, Brasil

RESUMO:

Este artigo apresenta em seqüência temporal algumas iniciativas e modelos de avaliação de sistemas educacionais no Brasil. Lembra a estrutura federativa do Estado brasileiro e o desdobramento das avaliações nos sistemas federal, estaduais e municipais de ensino. Aborda a questão da formação de pessoal em avaliação educacional, as características assumidas pelas avaliações implementadas, tanto na educação básica, como no ensino superior. Apontam-se as mudanças de modelo e as controvérsias advindas com o desenvolvimento dessas avaliações. Comentam-se alguns de seus efeitos.

PALAVRAS-CHAVE:

Avaliação educacional, Sistemas de ensino, Rendimento escolar, Avaliação institucional.

A Avaliação Educacional, hoje, é um campo de estudos com teorias, processos e métodos específicos, mas também, um campo abrangente que comporta sub-áreas, com características diferentes, por exemplo, avaliação de sistemas educacionais, avaliação de desempenho escolar em nível de sala de aula, avaliação institucional, avaliação de programas, auto-avaliação. Comporta, também, diferentes abordagens teóricas como a sistêmica, a iluminativa ou compreensiva, a avaliação participativa, etc. No Brasil este campo de conhecimento só veio a merecer maior atenção, estudos e análises críticas mais fundamentadas, há relativamente pouco tempo. As questões ligadas à avaliação educacional tendo adentrado muito tardiamente nas discussões no campo da educação, também sofreu aqui, em sua valorização e desenvolvimento como campo teórico, dos efeitos relativos à quase ausência de formação nos cursos de educação de profissionais especializados. Esta formação vem se dando por outros caminhos. Campo sujeito ainda a fortes críticas ideológicas, conta com massa crítica intelectual relativamente pequena, a qual está em crescimento em função de programas avaliativos postos em prática em vários níveis do sistema educacional brasileiro, com abrangência nacional ou regional. Isto implicou em preparação especial de funcionários nos diferentes níveis de administração do sistema escolar brasileiro, e na formação de pessoal universitário no exterior. Alguns grupos de especialistas foram se constituindo em universidades e centros de pesquisa, e, com os estímulos de formação no exterior pelas

políticas nacionais de capacitação de pessoal, vem aumentando o número de pesquisadores na área. Com isso, tem se tornado possível melhor qualificar os processos avaliativos de amplo espectro, relativos ao desempenho de alunos da educação básica (ensino fundamental e médio), ou do nível superior, e, de instituições de ensino. Neste artigo pretendemos tratar desses processos avaliativos, os de maior porte, no Brasil, recuperando sua trajetória no tempo. Sendo o Brasil uma federação, lembramos que o sistema educacional comporta a rede pública de ensino — com gestão federal, estadual ou municipal — e, a rede de escolas privadas, autorizadas e supervisionadas por um dos três níveis federativos. Cada nível federativo tem seu grau de responsabilidade definido pela Constituição de 1988, e, no referente às redes de ensino, pela Lei de Diretrizes e Bases da Educação Nacional, aprovada no ano de 1996.

AVALIAÇÕES EM LARGA ESCALA NA EDUCAÇÃO BÁSICA: ENSINO FUNDAMENTAL E MÉDIO

Dada a constituição tríplice das redes de ensino no Brasil estaremos nos referindo ora a avaliações nacionais, ora a estaduais, ora a municipais.

Para traçar a trajetória das avaliações de desempenho de redes de ensino tomamos como ponto de partida os anos mil novecentos e sessenta, porque foi nessa década que surgiu fortemente a preocupação específica com processos avaliativos escolares,

baseados em critérios mais claramente enunciados e instrumentos que poderiam garantir, até certo ponto, que a avaliação do nível de realização obtido estivesse mais objetivamente garantida. Nessa década e na subsequente, profissionais receberam formação mais aprofundada na área de avaliação de rendimento escolar, alguns no exterior. Na Fundação Getúlio Vargas do Rio de Janeiro criou-se em 1966 o CETPP, Centro de Estudos de Testes e Pesquisas Psicológicas, onde testes educacionais passam a ser desenvolvidos e estudados (Fundação Getúlio Vargas, 1970). A equipe do CETPP elaborou um conjunto de provas objetivas para as últimas séries do ensino médio, nas áreas de Linguagem, Matemática, Ciências Físicas e Naturais e Estudos Sociais. Realizou-se, então, pesquisa com conjuntos de alunos do ensino médio, a qual incluía um questionário sobre características sócio-econômicas dos alunos e suas aspirações. Esta pode ser considerada a primeira iniciativa relativamente ampla, no Brasil, para verificação da aquisição de conhecimentos e sua relação com diferentes variáveis, como sexo, nível socioeconômico, e outras. Nesse Centro desenvolveram-se, ainda, cursos sobre elaboração de provas objetivas, com especialistas estrangeiros, e fizeram-se várias publicações sobre temas ligados à avaliação educacional. Simultaneamente, na Fundação Carlos Chagas especialistas em testes e medidas também estavam sendo formados. No entanto, a expertise conseguida não teve nesse momento aplicação em avaliações de redes de ensino, tendo sido utilizada em processos seletivos para universidades, cursos superiores e cargos públicos.

Em meados dos anos 1970 desenvolveu-se, pela iniciativa do ECIEL (Programa de Estudos Conjuntos de Integração Econômica Latinoamericana), um estudo avaliativo de porte, no Brasil e outros países da América Latina, sobre os determinantes dos níveis de escolaridade e do rendimento escolar obtidos por alunos com diferentes características pessoais e sócio-econômicas. Aos alunos foi aplicado questionário para levantamento de dados sobre situação sócio-econômica, atitudes com relação ao processo escolar e aspirações, e, um exame de compreensão de leitura e de ciências. Foram coletados dados de diretores, professores e escolas (Castro & Sanguinetti, 1977). Realizou-se estudo, nesse mesmo período, com a finalidade de chegar a um instrumento

de medida que permitisse verificar a situação do desempenho dos alunos nas 1^{as} séries do ensino fundamental. Amostras foram utilizadas para testar o instrumento, em duas formas paralelas, com crianças de todas as regiões geográficas do país, abrangendo Leitura, Escrita e Matemática. Este projeto derivou da experiência com outros estudos durante a década de setenta no então estado da Guanabara (hoje cidade do Rio de Janeiro), e, outros pequenos estudos, feitos em convênio com o Centro Brasileiro de Pesquisas Educacionais/INEP/Ministério da Educação, desde a década de 1960.

Não houve, no entanto, nos anos imediatamente subsequentes outras iniciativas com a intenção de avaliações mais abrangentes, embora alguns ensaios localizados tenham sido desenvolvidos. Não se observa, nesse período, preocupação de administrações públicas com uma avaliação sistemática do rendimento escolar dos alunos das redes de ensino. Em nível nacional essa preocupação vai aparecer a partir de 1988, com alguns estudos exploratórios, ocorrendo a implementação de um sistema nacional de avaliação da educação básica apenas em 1990. Houve apenas uma municipalidade — a da cidade de S. Paulo — que se constituiu em exceção e preocupou-se no início dos anos 1980 com essa questão, tendo realizado o primeiro estudo avaliativo de uma rede de ensino. A avaliação foi conduzida pelo Departamento de Planejamento da Secretaria Municipal de Educação da Cidade de São Paulo, com o objetivo de verificação do nível de escolaridade dos alunos da Rede Municipal de Ensino, abrangendo, além das séries iniciais do ensino fundamental, também o terceiro estágio da educação infantil. Foram avaliadas todas as crianças do terceiro estágio de educação infantil, as de 1^a, 3^a, 5^a, 7^a séries do ensino fundamental e, os alunos do primeiro ano do ensino médio. As provas abrangeram Língua Portuguesa, Matemática e Ciências. Foi construída uma bateria de testes específicos, baseada no currículo escolar definido pela Secretaria Municipal de Educação para suas escolas, à época. A idéia era poder utilizar esses dados para planejamento das atividades escolares e extra-escolares. Porém, tendo havido mudança de administração em 1982, o estudo ficou com seus resultados sem utilização, uma vez que os novos dirigentes não se interessaram em utilizar os resultados obtidos e levar adiante o processo iniciado dois anos antes.

Numa outra vertente, a da avaliação de políticas educacionais e avaliação de programas, entre 1978 e 1982, encontramos alguns estudos que se propõem avaliar a política nacional de educação, enfocando a seletividade da educação, utilizando-se desde dados demográficos até estudos de caso (Fundação Carlos Chagas, 1981, 4v.). Estes estudos abrem um outro leque de possibilidades em avaliação numa direção diferente da avaliação de desempenho por testes.

Como se observa, lenta e esparsamente, algumas competências vinham se formando na área de avaliação, sobretudo envolvendo a questão da medida de rendimento escolar. Algumas administrações públicas da educação mostraram interesse nesses estudos mas, sem continuidade, nem em relação a equipes que se formavam, nem em relação aos estudos. A descontinuidade em políticas tem sido uma marca nas gestões públicas no Brasil, aliás, não só em relação aos estudos avaliativos, como a todos os aspectos de iniciativas de administrações anteriores. Também, nesse mesmo período — inícios dos oitenta — o pensamento educacional volta-se à crítica dos empirismos, dos tecnicismos, das operacionalizações burocráticas e a área de avaliação, que mal começara a formar quadros, sofre uma retração grande no seio das universidades.

Durante a década de 1980, uma experiência que pode ser colocada como um marco na história de desenvolvimento de estudos de avaliação de políticas e programas envolvendo não só estudos de rendimento escolar mas de variados fatores, foi a da avaliação do Projeto EDURURAL — um projeto de educação desenvolvido em todos os estados do Nordeste brasileiro. A avaliação acompanhou a implementação e desenvolvimento do projeto, sob vários aspectos, de 1982 a 1986. Avaliaram-se as formas de gerenciamento geral do projeto e, por amostra, o gerenciamento local, analisando-se o sistema de monitoria, os professores, as Organizações Municipais de Ensino, os alunos, as famílias. Desenvolveram-se testes para avaliar crianças em nível de segundas e quartas séries do ensino fundamental, considerando sua pertinência à zona rural em classes, em geral, multisseriadas. Estas provas foram desenvolvidas a partir de amostras de exercícios e trabalhos colhidos nas escolas dos três estados onde a avaliação se desenvolvia, Piauí, Ceará e Pernambuco. Coletou-se materiais de alunos, trabalhou-se com o pessoal

da região, e, depois é que se deu a forma final das provas. O que se tentava era construir um conjunto de provas tanto quanto possível adequado àquela realidade, buscando maior validade para os dados de rendimento escolar. As crianças foram avaliadas em 1982, 1984 e 1986. Nesses mesmos anos, os demais dados sobre o programa, seu gerenciamento, as características das escolas, o pessoal docente, as famílias, etc., também eram coletados e análises integradas e multivariadas foram feitas procurando encontrar traços característicos relevantes para verificar-se a contribuição sócio-educacional do programa, como também os impasses. Análises qualitativas dos dados levantados foram desenvolvidas, tendo sido agregados estudos de caso ao modelo avaliativo, para análise e compreensão em maior profundidade de situações mais específicas (Universidade Federal do Ceará, 1988, 7v.). Seminários foram realizados sobre seus resultados, durante o processo e depois. Esses resultados tiveram várias conseqüências sobre as políticas vigentes. Além disso, muitas outras análises derivaram destes dados, desde análises com enfoque em políticas públicas, até análises do que acontece com a escola que se situava na própria casa da professora; ou, qual o problema das escolas que estavam junto a assentamentos de pessoal sem terra; etc. Com metodologia clara e bem definida, com os cuidados de coleta e análise, representou um exemplo do que se poderia fazer com estudos dessa natureza na direção de se propugnar por uma escola mais condizente com as necessidades das populações menos favorecidas socialmente. O cuidado com as interpretações, com a clareza dos limites de significação dos dados, tendo presente os pressupostos sobre os quais se assentava o modelo avaliativo, ofereceu uma oportunidade para se formar pessoas e pensar a área criticamente a partir de uma ação direta.

ESTUDOS QUE DERAM SUBSÍDIOS À IMPLANTAÇÃO DO SISTEMA NACIONAL DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA

Ao final dos anos 1980 discussões sobre vários problemas que ocorriam nos sistemas educacionais, apontados pelos pesquisadores da área de educação, chega ao auge com o debate público sobre os indicadores que mostravam o alto índice de fracasso escolar (repetência e evasão escolar) na escola básica, no

país. Uma das questões que se colocava é que não se possuíam dados sobre o rendimento escolar dos alunos em nível de sistema e os fatores a ele associados. Nesse momento, promoveu-se no MEC [Ministério da Educação] uma oficina de trabalho com um grupo de educadores para discutir a questão. Colocava-se então que, com a mudança da nossa Constituição, poderia haver uma nova estrutura política e que isto sinalizava a necessidade de uma mudança na atuação do Ministério da Educação em relação ao ensino fundamental e médio, levando-o a ter um papel orientador e um papel avaliador como referência para políticas e avanços na área da educação básica. Ao final de 1987, foi proposto que se fizesse uma avaliação de rendimento escolar em 10 capitais de estados do país, para se aquilatar se um processo de avaliação mais amplo por parte do Ministério seria viável e traria resultados relevantes. A avaliação foi feita nas 1^{as}, 3^{as}, 5^{as} e 7^{as} séries de escolas públicas em 10 capitais de Estados, com provas em Língua Portuguesa (com redação), matemática e ciências. Foi um estudo piloto para verificar a viabilidade do processo, como as administrações e as escolas receberiam esse tipo de avaliação, se as provas seriam adequadas, etc. Utilizou-se a teoria clássica em avaliação e o grande desafio foi mesmo a construção de provas que tivessem validade para as realidades dos diferentes estados envolvidos (não se dispunha de um currículo nacional). Isto tornou-se possível a partir de parcerias locais. Com esta etapa tendo sido relativamente bem sucedida, com as discussões sobre os resultados das provas e fatores a eles associados havidas em alguns estados, e discussões em seminários nacionais ou locais, expandiu-se o estudo avaliativo para mais 20 capitais e, depois mais 39 cidades distribuídas em 14 estados e, à época, um território (Vianna, 1988, 1989a, 1989b). Paralelamente, com apoio da Secretaria de Ensino de Segundo Grau do Ministério da Educação, realizou-se uma avaliação de rendimento escolar de alunos do 3º ano do ensino médio, abrangendo todas as modalidades existentes (geral, normal, técnico industrial, comercial, etc.).

Também no ano de 1988, como extensão à iniciativa do MEC, a Secretaria de Estado da Educação do Paraná realizou uma avaliação dos alunos de 2ª e 4ª séries nas escolas desse estado. Provas específicas de Língua Portuguesa, Matemática, Ciências e Estudos Sociais foram elaboradas, a partir de itens

produzidos por professores locais, com base nos guias curriculares vigentes no Paraná (Vianna & Gatti, 1988). Em 1991 completando o ciclo de avaliações exploradoras proposto pelo MEC realizou-se o mesmo tipo de avaliação feito nas escolas públicas, em escolas privadas, em 11 Estados e no Distrito Federal.

As avaliações de rendimento escolar realizadas entre 1988-91 trouxeram o impacto dos baixos resultados médios, muito aquém do esperado, que repercutiram tanto no Ministério, como nas Secretarias de Educação e na mídia, criando nas administrações públicas interesse pelos processos avaliativos. Com os resultados obtidos e dados das escolas, dos professores e dos alunos, dispunha-se de muitas informações, sobre as quais se poderia debruçar, refletir e tirar inferências tanto para políticas dirigidas às redes de ensino como um todo, como para questões da aprendizagem de alunos nas salas de aula. Estes primeiros estudos mais abrangentes de alunos e escolas serviram de base para a implantação do SAEB — Sistema de Avaliação da Educação Brasileira.

Para isso contribuiu, também, a participação do Brasil, logo no início dos anos 1990, no segundo Programa Internacional de Avaliação de Proficiência Educacional. Esse programa internacional, que à época envolvia 27 países, pretendia ser comparativo. Portanto, a metodologia de aplicação e os testes foram elaborados por uma equipe internacional de especialistas: chineses, árabes, americanos, suecos, holandeses, etc., num processo de validação recíproca. Mas, no Brasil o estudo ficou restrito a apenas duas capitais (Fortaleza e S. Paulo) tornando a comparabilidade pretendida comprometida. Foram avaliadas crianças de 13 anos de idade, não importando a série em que o aluno se encontrava (a correspondência idade-série não era garantida). A idéia desse estudo internacional era procurar saber qual domínio de conhecimentos teriam as crianças dessa idade em cada um dos países participantes. No estudo aqui realizado não foram incluídas crianças que estavam fora da escola e crianças com excessiva defasagem idade-série, problema que na maior parte dos países participantes não se colocava. A avaliação, embora realizada apenas em duas capitais, mostrava o precário desempenho dos alunos de 13 anos no Brasil, na comparação com outros países: penúltima colocação.

A contundência dos dados obtidos até aqui leva o Ministério da Educação e alguns dos Estados da

federação a desenvolver uma política de avaliação de redes de ensino de forma a trazer a público os resultados, na intenção de se buscar formas de superação da precária situação das aprendizagens escolares. Os resultados médios eram muito preocupantes. Estímulos vindos de órgãos internacionais tiveram papel relevante na consolidação e ampliação dessas avaliações.

O SISTEMA NACIONAL DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA — SAEB

A partir de 1993, o Ministério da Educação, em articulação com as Secretarias Estaduais de Educação, implantou o Sistema de Avaliação da Educação Básica (SAEB). Realizado por amostra nacional de alunos, inicialmente com provas objetivas no modelo clássico, essa avaliação passou a ser realizada anualmente, abrangendo séries do ensino fundamental e médio, tendo aperfeiçoado sua metodologia e, finalmente, em 1995, adotado os procedimentos preconizados pela Teoria da Resposta ao Item. Pesquisadores da Fundação Carlos Chagas capitanearam a introdução dessa metodologia no SAEB (Andrade & Valle, 1998; Fletcher, 1991, 1995). No entender dos especialistas poderia oferecer informações mais seguras sobre o repertório escolar das crianças e jovens, bem como oferecer condições de comparabilidade em escala, o que não era possível no modelo anterior. Pelos documentos disponíveis verifica-se que a intenção associada a essa avaliação era a de prover informações para tomadas de decisão quanto a diversos aspectos das políticas educacionais, bem como para pesquisas e discussões, a partir da geração e organização de informações sobre o desempenho acadêmico dos alunos no sistema e fatores a ele associados. Hoje estão disponíveis várias bases de dados sobre as sucessivas avaliações feitas em nível nacional, e relatórios que disponibilizam as análises realizadas. O SAEB compõe-se de dois grandes eixos: o primeiro voltado ao acesso ao ensino básico no qual se verifica o atendimento à demanda (taxas de acesso e taxas de escolarização) e a eficiência (taxas de produtividade, taxas de transição e taxas de eficiência interna); o segundo eixo é relativo à qualidade, implicando no estudo de quatro dimensões relativas: 1. ao produto — desempenho do aluno quanto a aprendizagem de conteúdos e desenvolvimento de habilidades e competências; 2. ao contexto —

nível sócio-econômico dos alunos, hábitos de estudo, perfil e condições de trabalho dos docentes e diretores, tipo de escola, grau de autonomia da escola, matriz organizacional da escola; 3. ao processo — planejamento do ensino e da escola, projeto pedagógico, utilização do tempo escolar; estratégias de ensino; 4. aos insumos — infra-estrutura, espaço físico e instalações, equipamentos, recursos e materiais didáticos. Os instrumentos de coleta são as provas aos alunos, questionários aplicados a alunos, professores e diretores, questionário sobre as condições da escola. Nos dois últimos anos agregou-se ao modelo geral de avaliação, estudos de caso de situações específicas consideradas relevantes para análise mais aprofundada de fatores contextuais intervenientes nos resultados (Brasil/Ministério da Educação, 2002; Fontanive & Klein, 2000; Pestana, 1999).

O SAEB veio sendo objeto de vários estudos e discussões, inclusive de comissões em nível de Ministério na busca de seu aperfeiçoamento. Problemas técnicos têm sido superados, por exemplo, quanto à modelagem das provas e o teor dos itens e sua validade, quanto ao processo de amostragem, que vem sendo aperfeiçoado, etc., levantando-se, também, problemas quanto à divulgação, disseminação e apropriação dos resultados nos diferentes níveis de gestão do sistema e aos professores. Franco (2001) nos oferece observações interessantes quanto às potencialidades e aos problemas relativos a essa avaliação, pontuando sugestões sobre as formas de tornar o SAEB mais relevante para gestores e professores, por exemplo, ilustrando com exemplos claros dificuldades dos alunos, propondo alternativas pedagógicas, em linguagem adequada, construir melhor explicitação dos fatores que contam em educação (efeito escola, efeito sala de aula), apresentação das escalas de proficiência em linguagem e esquemas gráficos facilmente inteligíveis para não especialistas, estabelecimento de relacionamentos com secretarias estaduais e municipais mais efetivos para apropriação dos resultados, etc.

Mais recentemente o Ministério de Educação expandiu a avaliação de desempenho do sistema educacional, com a *Prova Brasil*, aplicada a todos os alunos dos primeiros anos do ensino fundamental, integrando essa prova ao SAEB. Com isso avançou-se na construção de um indicador — o Índice de Desenvolvimento da Educação Básica (IDEB), a

partir de 2007, que integra o fluxo escolar e as médias de desempenho nas avaliações nacionais. Essas médias referem-se ao desempenho no SAEB para os Estados e na *Prova Brasil* para os Municípios. É possível dispor de um valor do Indicador para o país como um todo, cada Estado, Município e Escola.

INICIATIVAS REGIONAIS

Foi durante os anos mil novecentos e noventa que diversas avaliações de sistemas escolares vão tomar corpo também em várias regiões do país, além do âmbito nacional já exposto. Abaixo relatamos três iniciativas em Estados da Federação, as mais consolidadas.

Estado de S. Paulo

Assim, em 1992 foi realizada no Estado de S. Paulo avaliação dos alunos das Escolas Padrão, projeto de inovação curricular desenvolvido pela Secretaria Estadual de Educação, inicialmente com 306 escolas no Estado, depois ampliando-se para mais de mil, com vistas a abranger aos poucos todas as escolas do sistema, o que não chegou a acontecer. O objetivo da avaliação era verificar se com os insumos educacionais diferenciados que recebiam as escolas e professores se obtinham melhorias no nível de aprendizagem escolar por parte dos alunos. A avaliação então realizada abrangeu Língua Portuguesa, Redação, Matemática, Ciências, História e Geografia, tendo sido aplicado um questionário para sondagem de características dos alunos. Foram submetidos a ela todos os alunos de 8ª série. Essa avaliação seria o marco zero, com a qual sucessivamente se procuraria comparar o futuro desenvolvimento educacional das crianças. Esperava-se, no início dos trabalhos, poder realizar avaliações sucessivas, durante vários anos, com dados da progressão dessas crianças que estavam vinculadas ao Projeto da Escola Padrão. Os primeiros resultados foram interpretados psico-pedagogicamente e recomendações curriculares foram feitas e consolidadas em documentos distribuídos às escolas. A partir dos dados obtidos vários estudos com foco específico foram realizados para subsidiar ações específicas junto às escolas (Gatti, 1996). Essa avaliação, que pretendia ser longitudinal, foi descontinuada por mudança na administração da Secretaria de Educação do Estado. A nova gestão desenvolveu outros estudos visando comparar Escolas Padrão e Não Padrão, considerando a região

da escola, os períodos, diurno e noturno dos cursos. Os dados foram divulgados para a rede. Enfatizava-se que se deveria enfrentar com critério e determinação os pontos críticos revelados em relação ao processo de ensino. As provas foram analisadas e implicações pedagógicas foram apontadas, com recomendações quanto aos pontos críticos em cada disciplina avaliada (Mondel & Maluf, 1994).

Ao mesmo tempo, entre 1992 e 1994, outro projeto de avaliação vinha sendo desenvolvido, amostralmente, nas escolas estaduais do Estado de S. Paulo, com caráter de uma verdadeira pesquisa avaliativa: O Projeto de Avaliação de Impacto do Ciclo Básico e da Jornada Única na Área Metropolitana de São Paulo. A equipe de pesquisa fez um estudo longitudinal de acompanhamento de um mesmo segmento de crianças, durante 3 anos, com o objetivo complexo de analisar mudanças de aprendizagem e de características cognitivas ao longo do tempo. Esse foi um tipo de pesquisa em avaliação nunca antes desenvolvido no país, inclusive com um estudo paralelo de caráter antropológico. Este permitiu que se tivesse uma compreensão profunda do que se passava no cotidiano das escolas e no desenvolvimento cognitivo das crianças (Neubauer *et al.*, 1996). A partir de 1995 instituiu-se no Estado de S. Paulo, como parte de um programa para a educação pública do estado, o SARESP (Sistema de Avaliação do Rendimento Escolar no Estado de S. Paulo), que vem se realizando até hoje, abrangendo séries sucessivas, com questionários informativos, e fazendo-se as equalizações necessárias ano a ano para se obter comparabilidade. Neste modelo todos os alunos matriculados nas séries envolvidas são avaliados (Espósito, 2000; SEESP, 1996, 1998). O sistema SARESP tem mantido as avaliações periódicas e em 2008 fundamentou a criação de um indicador de desenvolvimento educacional específico para o Estado de S. Paulo — IDESP —, com características de construção um pouco mais nuanceadas do que o índice do Ministério da Educação.

Estado de Minas Gerais

Também o Estado de Minas Gerais começa a partir de 1992 a desenvolver um programa de avaliação das suas escolas públicas. Como as demais iniciativas, o propósito da avaliação era dar base para a melhoria da qualidade do ensino e fazia parte de

uma proposta mais ampla do governo do estado para a educação. As avaliações foram planejadas e realizadas em ciclos, a cada dois anos, eram avaliações censitárias, incorporando outras informações através de questionários. Em 1998, com a implantação da progressão continuada nas escolas, passou-se a fazer a avaliação todos os anos. Desenvolveu-se com este programa de avaliação uma proposta formativa de quadros na própria Secretaria, através de cursos específicos oferecidos aos técnicos, aqui e no exterior (Antunes *et al.*, 1992; Souza, 1999). Hoje, o Estado de Minas Gerais possui um complexo modelo de avaliação de sua rede de ensino — o SIMAVE: Sistema de Avaliação da Qualidade do Ensino e das Escolas. Esse Sistema é composto por três programas de avaliação: PROALFA — voltado à avaliação dos níveis de alfabetização; PROEB — verifica a eficiência e a qualidade do ensino com base no desempenho nas séries finais dos blocos de ensino; PAAE — realiza diagnósticos progressivos da aprendizagem para subsidiar intervenções pedagógicas.

Estado do Paraná

O Estado do Paraná realizou em 1995 sua primeira avaliação em larga escala como expansão do SAEB, oferecendo aos seus municípios e às escolas resultados particularizados. Deu-se continuidade ao processo nos anos subsequentes e, em 2000, introduziu no processo avaliativo do ensino fundamental a metodologia da teoria da resposta ao item, tomando como referência os conteúdos do Currículo Básico para as Escolas Públicas do Paraná. Desenvolveram-se estudos analíticos detalhados sobre o desempenho dos alunos de 4^a e 8^a séries. As matrizes de referência das provas trazem inovações em sua concepção em relação a outras abordagens, por exemplo, em Língua Portuguesa, contempla cinco áreas: procedimentos de leitura, recursos expressivos e efeitos de sentido, variação linguística, coerência e coesão, e, conhecimento lingüístico operacional. Análises diversificadas encontram-se consolidadas em relatórios publicados e distribuído à rede. A disseminação desses resultados foi realizada por três tipos diferentes de relatórios: para diretores de escola e secretários municipais de educação, com uma visão mais global e comparativa; para professores e equipes pedagógicas contendo análises detalhadas dos itens, dos tipos de erros, com

recomendações e sugestões pedagógicas; e, para presidentes de associações de pais e mestres e pais de alunos foi feito no Boletim da Escola uma síntese específica da situação da escola e uma descrição dos níveis de desempenho alcançados (SEEPR, 2001). Mais recentemente implementou-se nesse Estado um processo de avaliação institucional envolvendo escolas, regionais de ensino e Secretaria.

Estado do Ceará

O Estado do Ceará desenvolveu um sistema avaliativo de suas escolas a partir de 1996, com avaliação, por amostras, de desempenhos de alunos do ensino fundamental. Realizou, também, um processo de avaliação institucional de escolas, com envolvimento da comunidade, o qual se mostra como um modelo diferenciado em relação ao que se fez em outros estados. Neste Estado foi institucionalizado o SPAECE — Sistema Permanente de Avaliação da Educação Básica. Conforme documentado pela Secretaria Estadual de Educação (Ceará, 2009), compreende três vertentes: Avaliação de Desempenho Acadêmico, Avaliação Institucional e Estudos e Pesquisas Educacionais. Na primeira, de natureza externa, o Sistema avalia as competências e habilidades dos alunos do Ensino Fundamental e do Ensino Médio, nas disciplinas de Língua Portuguesa e Matemática. Identifica-se o nível de proficiência e a evolução do desempenho dos alunos. Abrange todos os alunos das escolas estaduais e municipais. São aplicados questionários investigando dados sócio-econômicos e hábitos de estudo dos alunos, perfil e prática dos professores e diretores. Na vertente da Avaliação Institucional, possibilita à escola, através da Auto-Avaliação e da Avaliação de Desempenho do Núcleo Gestor, conhecer e aperfeiçoar as inter-relações, os serviços prestados, o desempenho do corpo docente e discente, dos funcionários e gestores. Objetiva a implementação de mudanças no cotidiano escolar. A terceira vertente compreende Estudos e Pesquisas Educacionais, bem como Avaliações de Programas, buscando aprofundar o conhecimento das situações-problema e das tendências detectadas nas avaliações.

CONTRIBUIÇÕES E DESAFIOS

Essas experiências de mais amplo espectro contribuíram para a formação e o aperfeiçoamento de pesquisadores e técnicos em avaliação de desempenho escolar e institucional em sistemas de ensino, tanto em nível do Ministério da Educação, nas Secretarias e em diversas instituições que passaram a se dedicar ao campo. Depreende-se pelos documentos consultados que se buscou ao longo do tempo criar um ambiente propício à receptividade de avaliações, incrementar competências no campo, e estimular o emprego das avaliações para mudanças, intervenções, melhorias nas escolas e no ensino. Também, despertar estudos críticos. Ou seja, pensava-se em provocar mudanças na representação dos processos avaliativos que tinham conotação punitiva e de depreciação dando-lhes novo significado. Todos sabemos das dificuldades de mudanças dessa natureza — mudanças sócio-culturais — mas, tentava-se no nível das gestões estimular o início desse processo. Fatores internos e externos interuseram-se e as contradições implicadas nesses procedimentos tomaram corpo. De um lado, uma ênfase muito forte que acaba sendo dada aos melhores e aos piores desempenhos, a mídia valorizando apenas a média das notas obtidas e não outros resultados até mais importantes. De outro, observou-se a pequena utilização dos dados por parte das redes, seja por processos inadequados de disseminação e de dar acesso e visibilidade aos dados, seja por dificuldades nas redes de lidar, analisar, interpretar esses dados pedagogicamente. Os problemas de utilização dos dados por diretores de escola, coordenadores pedagógicos e professores vem ocorrendo tanto na utilização das avaliações estaduais como das avaliações nacionais. A disseminação dos dados em formas mais adequadas, diferenciadas conforme a audiência — público em geral, pais, professores, gestores — é ponto nevrálgico a ser resolvido no desenvolvimento desses processos. O IDEB e o IDESP têm sido melhor

compreendidos, porém, é necessário que aspectos pedagógicos ligados aos desempenhos sejam mais amplamente expostos e discutidos. O grande desafio é, ainda, a apropriação por parte das escolas dos resultados obtidos por seus alunos e a utilização desses resultados para orientar as atividades de ensino.

Aspectos mais técnicos, ligados à construção de provas de avaliação de desempenho, tanto relativos à fidedignidade como à validade, também foram objeto de discussão. As matrizes de referência têm sido sistematicamente criticadas, em que pese a busca de seu aperfeiçoamento contínuo. Outro aspecto polêmico é a utilização da Teoria da Resposta ao Item nos principais processos avaliativos nas grandes redes de ensino, onde se questiona a pertinência do modelo e a adequação das análises e sua consistência.

FINALIZANDO

Pode-se afirmar que vem mudando a representação dos processos avaliativos de sistemas educacionais no Brasil, em relação aos quais houve inicialmente uma reação contrária muito forte, e que uma cultura de avaliação educacional está se consolidando, na idéia de *accountability*. Salutares são os debates e as contraposições aos diversos processos avaliativos implementados em vários dos níveis educacionais, e nos vários sistemas de administração educacional do país (União, Estados e Municípios). A preocupação com os resultados dos processos de ensino está presente atualmente nas administrações públicas da educação e nas escolas, dando margem a iniciativas como aperfeiçoamento dos currículos escolares, formação continuada de professores, revisão da formação básica de docentes, produção de materiais didáticos novos em vários tipos de suporte (impressos, virtuais, DVDs, etc.). O impacto dessas avaliações começa a ser sentido na educação básica esperando-se que as avaliações sejam vistas como estímulos à mudança em processos educacionais, e, não como punição.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANDRADE, Dalton F. & VALLE, Raquel C. (1998). Introdução à Teoria da Resposta ao Item. São Paulo, *Estudos em Avaliação Educacional*, 18, pp. 13-32.
- ANTUNES, Ana Lúcia; XAVIER, Edir P. C. & FREITAS, Maria A. de. (1992). Avaliação do rendimento do aluno da escola estadual de Minas Gerais: o CBA. São Paulo, *Estudos em Avaliação Educacional*, 6, pp. 29-42.
- ANPED — ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO E PESQUISA EM EDUCAÇÃO (1999). *A Avaliação da pós-graduação em debate*. São Paulo: ANPED.
- ANPED — ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO E PESQUISA EM EDUCAÇÃO (2005). *40 Anos da Pós-Graduação em Educação*. *Revista Brasileira de Educação*. São Paulo: ANPED/Autores Associados.
- BRASIL. MINISTÉRIO DA EDUCAÇÃO E DO DESPORTO (1995). *Resultados do SAEB 1995: a escola que os alunos freqüentam*. Brasília: MEC.
- BRASIL. MINISTÉRIO DA EDUCAÇÃO (1996). Programa de Avaliação Institucional das Universidades Brasileiras — PAIUB. *Avaliação: Revista da Rede de Avaliação Institucional da Educação Superior*, 1, 1, pp. 43-68.
- BRASIL. MINISTÉRIO DA EDUCAÇÃO (2002). Sistema Nacional de Avaliação da Educação Básica: SAEB. In BRASIL. Ministério da Educação, *Textos divulgados para a imprensa: coletânea 1966 — 2001*. Brasília: MEC/INEP, pp. 133-155.
- CASTRO, Cláudio de M. & SANGUINETTY, Jorge A. (1977). *Custos e determinantes da educação na América Latina: resultados preliminares*. Rio de Janeiro: INTED.
- CEARÁ (2009). *Sistema Permanente de Avaliação da Educação Básica — SPAECE*. Fortaleza: Secretaria de Estado da Educação.
- ESPOSITO, Yara L. (coord.); SÃO PAULO (Estado) & SECRETARIA DA EDUCAÇÃO (2000). *Sistema de avaliação de rendimento escolar do Estado de São Paulo — SARESP 98: conhecendo os resultados da avaliação*. São Paulo: SEE/FDE.
- FLETCHER, Philip R. (1991). Avaliação do perfil cognitivo da população brasileira. São Paulo, *Estudos em Avaliação Educacional*, 4, pp. 27-64.
- FLETCHER, Philip R. (1995). Propósitos da avaliação educacional: uma análise de alternativas. S. Paulo, *Estudos em Avaliação Educacional*, 11, pp. 93-112.
- FONTANIVE, Nilma S. & Klein, Ruben (2000). Uma visão sobre o sistema de avaliação da educação básica do Brasil — SAEB. Rio de Janeiro, *Ensaio: Avaliação e Políticas Públicas em Educação*, 29, 8, pp. 409-439.
- FRANCO, Creso (2001). O SAEB — sistema de avaliação da educação básica: potencialidades, problemas e desafios. *Revista Brasileira de Educação*, 17, pp. 127-133.
- FUNDAÇÃO CARLOS CHAGAS (1981). *Educação e desenvolvimento social*. 4v. São Paulo: FCC/DPE.
- FUNDAÇÃO GETÚLIO VARGAS (1970). *Testes e medidas na educação*. Rio de Janeiro: FGV.
- GATTI, Bernardete A. (1992). Avaliação da jornada única em São Paulo. *Estudos em Avaliação Educacional*, 5, pp. 85-90.
- GATTI, Bernardete A. (1996). Desenvolvimento de projetos de avaliação do sistema educacional no Estado de São Paulo. *Estudos em Avaliação Educacional*, 13, pp. 19-26.
- MONDEL, Lúcia M. & MALUF, Mônica M. B. (1994). Avaliação do Rendimentos de Alunos das Escolas Padrão. *Estudos em Avaliação Educacional*, 10, pp. 103-122.
- NEUBAUER, Rose; DAVIS, Cláudia & ESPÓSITO, Yara L. (1996). Avaliação do processo de inovações no ciclo básico e seu impacto sobre a situação de ensino-aprendizagem na Região Metropolitana de São Paulo. *Estudos em Avaliação Educacional*, 13, pp. 35-64.
- PESTANA, Maria Inês G. S. (1999). O Sistema de avaliação brasileiro. Brasília, *Revista Brasileira de Estudos Pedagógicos*, 191, 79, pp. 65-73.
- SEEPR — PARANÁ. SECRETARIA DA EDUCAÇÃO. COORDENAÇÃO DE INFORMAÇÕES EDUCACIONAIS (2001). *Cadernos Ava*. Curitiba: SEEPR.
- SEESP — SÃO PAULO (Estado) SECRETARIA DA EDUCAÇÃO (1996). *Sistema de avaliação de rendimento escolar do Estado de São Paulo — SARESP: relatório final dos resultados da 1ª aplicação*. São Paulo: SEE.
- SEESP — SÃO PAULO (Estado) SECRETARIA DA EDUCAÇÃO (1998). *Sistema de avaliação de rendimento escolar do estado de São Paulo — SARESP*. São Paulo: SEE/FDE.

- SOUZA, Maria Alba de (1999). A Experiência de Avaliação Educacional em Minas Gerais: 1992-1998. *Estudos em Avaliação Educacional*, 19, pp. 57-76.
- UNIVERSIDADE FEDERAL DO CEARÁ (1988). *Edurnal 1981-1985: avaliação da educação básica no nordeste brasileiro*. (7 v.). Fortaleza: UFC/FCC/DPE.
- VIANNA, Heraldo M. (1988). Avaliação do Rendimento de Alunos de Escolas de 1º grau da Rede Pública: uma aplicação experimental em 10 cidades. *Educação e Seleção*, 17, pp. 5-52.
- VIANNA, Heraldo M. (1989a). Avaliação do Rendimento de Alunos de Escolas de 1º grau da Rede Pública: um estudo em 39 cidades. *Educação e Seleção*, 20, pp. 5-56.
- VIANNA, Heraldo M. (1989b). Avaliação do Rendimento de Alunos de Escolas do 1º Grau da Rede Pública: um estudo em 20 cidades. *Educação e Seleção*, 19, pp. 33-98.
- VIANNA, Heraldo M. & GATTI, Bernardete A. (1988). Avaliação do Rendimento de alunos das 2ªs e 4ªs séries de Escolas Oficiais do Estado do Paraná. *Educação e Seleção*, 18, pp. 5-62.

