

AVALIAÇÃO EDUCACIONAL NO BRASIL: PONTUANDO UMA HISTÓRIA DE AÇÕES

*Bernardete A. Gatti**

PALAVRAS-CHAVE: avaliação educa-
cional; avaliação de rendimento
escolar; avaliação institucional;
modelos avaliativos; avaliação de
sistemas educacionais.

*Pesquisadora da Fundação
Carlos Chagas e Professora na
PUC-SP.

RESUMO: O objetivo deste artigo é registrar e discutir experiências com avaliação educacional de porte mais extenso, em sistemas educacionais ou parte destes, desde os anos mil novecentos e sessenta para cá. Abordam-se processos desenvolvidos na educação básica e no ensino superior. Discutem-se a carência de especialistas no campo e a falta de cursos, no país, para a formação mais adequada na área de avaliação educacional. Disto decorrem alguns problemas que nesse campo de conhecimento se enfrentam em nível tanto teórico quanto técnico e social.

Avaliação Educacional hoje não é apenas um campo com teorias, processos e métodos específicos, mas também um campo abrangente que comporta subáreas, com características diferentes: avaliação de sistemas educacionais, de desempenho escolar em sala de aula, de rendimento escolar com objetivo de macroanálises, de programas, avaliação institucional e auto-avaliação. Admite ainda diferentes enfoques teóricos como avaliação sistêmica, avaliação iluminativa ou compreensiva, avaliação participativa etc.

No entanto, geralmente quando se fala em Avaliação Educacional, o que vem à mente é a de rendimento escolar, ou de desempenho, confundida com a idéia de medida pontual. Não sem razão, visto que esta é a modalidade de avaliação mais presente no cotidiano das pessoas. Como nossas escolas emergiram sob a égide da preparação de elites, a avaliação seletiva no cotidiano escolar firmou-se, por centenas de anos, como cultura preponderante. Mesmo em momentos nos quais aparecem propósitos de democratização das escolas e os primeiros debates sobre a expansão do sistema de ensino, a cultura que dá forma aos processos avaliativos na escola é a que tem como referência critérios altamente seletivos, associados a padrões ditos rigorosos, mas sempre pouco claros.

Essa seletividade só se tornará objeto de discussão pública quando um

número maior de alunos, na década de 60, começa a procurar o ensino superior e encontra a grande barreira dos vestibulares, problema levantado pelos que conseguiram superar vários obstáculos, tais como: ter tido acesso à escola, conseguir permanecer nela e concluir o ensino médio, o que, em termos de população brasileira, eram proporcionalmente muito poucos. Note-se que, nos anos 60 e início dos 70, ainda não se discutem a reprovação escolar em massa que se processava no ensino fundamental e a evasão de alunos, que tinha foros dramáticos. Nas escolas, fazer alunos 'repetirem' o ano, por 'avaliações rigorosas', tornou-se 'natural'. O fato de se eliminarem alunos das escolas, especialmente os de baixa renda, pelo insucesso ininterrupto, não era questionado.

Foi essa a cultura que se desenvolveu em torno dos processos avaliativos no cotidiano escolar e que, apesar das amplas discussões sobre a questão, permanece até hoje, tendo deixado uma forte marca na vida das pessoas e na representação que fazem da avaliação. Então, o fato de os processos avaliativos estarem sempre presentes no ambiente escolar, em todos os níveis, faz com que as pessoas se reportem de imediato a esse tipo de avaliação quando se fala em avaliação educacional, restringindo seu significado.

Além disso, exames como os vestibulares, o provão e o Enem, com o destaque que a grande mídia lhes dá, reforçam outra associação: avaliar é medir conhecimento. Mas o campo da Avaliação Educacional é mais abrangente, como já dissemos, e seu papel deve ser mais construtivo, até porque avaliar é muito mais do que medir.

Deve-se notar também que, ao se analisarem esses processos avaliativos, nas escolas ou em exames externos, o conteúdo das observações e as perspectivas de análise de questões a eles ligados são, em geral, empobrecidos por um fator histórico específico de nosso país: durante anos e anos não formamos pessoas com base sólida em avaliação educacional; não se tem, portanto, domínio adequado do campo e de seus avanços. Poucos se dedicaram aos estudos sobre avaliação educacional, que praticamente desapareceu como área de estudo e formação em nossas universidades e cursos superiores, e a produção científica relevante sobre a questão foi pequena e esparsa por um longo período.

Mais recentemente, em razão das políticas educacionais instituídas nacionalmente e, em alguns estados, pela ampliação do debate sobre os problemas que envolvem a reprovação escolar das pessoas social e culturalmente desfavorecidas,

a área vem merecendo mais atenção, relativamente às suas diferentes abordagens, e alguns esforços de formação de especialistas começaram a se fazer presentes.

Assim, apesar de se tratar de um processo vivido no cotidiano escolar, marcando nossas vidas e criando socialmente a imagem da Avaliação Educacional, essa área só mereceu atenção e análise crítica mais fundamentada, no Brasil, há pouco tempo, se comparada à atenção a ela dada em outros países. Tardamente presente nas discussões do campo da educação, sua valorização e desenvolvimento como campo teórico também sofreu os efeitos dos ressentimentos, do desprezo e da crítica ideológica, o que levou à carência, hoje, de massa crítica intelectual especializada que possa contribuir para a formação de uma consciência avaliativa de professores e demais educadores.

Ocorre que o desenvolvimento desse campo em nosso país foi truncado, num processo historicamente compreensível. Incipientemente introduzido nas primeiras décadas do século XX, sob a égide da psicomетria, não adquiriu relevância nas produções acadêmicas e no ensino nas décadas subseqüentes. Nos anos 60, no âmbito da valorização de perspectivas mais tecnicistas em educação, passa a ser incorporada no contexto das políticas desenvolvimentistas do período. No entanto, no fim dos anos 70 e início dos 80, a avaliação educacional, que ainda não tomara fôlego entre nós, é posta 'sob suspeita' como área do conhecimento, inclusive pelas grandes discussões sobre as avaliações vinculadas aos vestibulares para ingresso no ensino superior e suas repercussões na mídia. Diversos motivos contribuem para esse processo de desabonamento social e acadêmico, a confrontar os desejos, de um lado, de alguns estudiosos que queriam contribuir, por meio de seus estudos críticos, para a melhoria desse processo; de outro, os desejos da grande massa de estudantes, impulsionada pelos pais, ansiosa por uma vaga no ensino superior. Essa realidade gera a desqualificação dos processos avaliativos, que se estende à própria avaliação educacional, descaracterizando o seu verdadeiro objetivo como área de estudos e pesquisa. Além disso, a banalização, à época, do uso de 'testes objetivos' nos livros didáticos (se é que poderiam ser assim considerados, tal a precariedade de sua formulação) clamou por uma crítica densa, pois apresentava-se este tipo de modalidade avaliativa como solução hegemônica, cujo papel nos processos educacionais mostrava-se bastante duvidoso.

A partir de algumas iniciativas tímidas no fim dos anos 80, inicia-se a década de 90 com a introdução de políticas educacionais que vêm acompanhadas de preocupações avaliativas (por exemplo, a Escola Padrão, no estado de São

Paulo). No país, começa-se a estruturar um sistema nacional de avaliação da educação básica, e essa iniciativa do Ministério da Educação logo é acompanhada de iniciativas semelhantes em alguns estados da federação, cada qual com seu modelo específico. Fala-se na criação de uma nova 'cultura de avaliação', mas a falta de pessoas especializadas na área continua sendo um problema, apesar dos esforços de formação de algumas equipes. Para minimizar o problema, formam-se grupos que, a partir da pequena massa crítica existente, desenvolvem, ao mesmo tempo, os processos avaliativos, no caso avaliação de sistemas. Nessa época, a carência de formação básica na área – os cursos superiores não ofereciam (e não oferecem) – tornou-se obstáculo à continuidade desses processos, dificultando o trabalho das equipes e a formação de outras.

Decorrem daí a precária formação na especialidade e o número relativamente pequeno de especialistas com formação mais sólida no campo – pode-se dizer que desenvolvemos bem a crítica, mas pouco fizemos pela formação de quadros. Além disso, outras subáreas desse campo também não se desenvolveram suficientemente entre nós, com exceção talvez da avaliação institucional, que caminhou por modelos sistêmicos ou modelos compreensivos e participativos, criando competências em várias instituições.

Traçado esse quadro geral, descreveremos a seguir alguns dos caminhos que processos avaliativos de mais larga escala tomaram em nosso país, num primeiro momento na educação básica e, depois, no ensino superior.

Avaliações em larga escala na Educação Básica

Nossas preocupações focarão o desenvolvimento de processos avaliativos escolares de maior porte, ou sistemas, observando a dinâmica de sua evolução. Tomaremos como ponto de partida os anos 60, porque foi nessa década que surgiu a preocupação específica com os processos avaliativos escolares baseados em critérios mais claramente enunciados e em instrumentos que poderiam garantir, até certo ponto, seu nível de realização pelo uso das técnicas chamadas de 'objetivas'.

Teorizações sobre este tipo de instrumento de avaliação de desempenho escolar e suas respectivas aplicações já se haviam desenvolvido em alguns países, especialmente na Inglaterra e Estados Unidos. Aqui, algumas instituições e escolas

já o utilizavam, mas, até meados da década de 60, seu uso entre nós foi bastante restrito. No entanto, a iniciativa de sua adoção nos exames vestibulares para as universidades deu-lhes maior visibilidade, e os testes objetivos passam a ser utilizados em larga escala pelos educadores. Essa modalidade de instrumento orientou os exames vestibulares unificados em 1969. Os testes objetivos de avaliação de domínio de conhecimentos passam a ser mais conhecidos nas escolas de educação básica, associados com políticas de educação que começam a privilegiar a avaliação educacional ligada à operacionalização de objetivos instrucionais e à operacionalização de competências e níveis de habilidades, na maioria dos casos com base na taxonomia de Benjamin Bloom (1972). Livros didáticos incorporaram esses testes, alguns de modo absolutamente aligeirado e até de forma incompetente, gerando fortes críticas a esse tipo de instrumento. Neste processo, fica patente a incorporação de instrumental de avaliação que se expande sem a formação de competências de base para sua criação e validação, a não ser em grupos muito restritos.

Anos 60-90

Nos anos 60 e 70, alguns 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 –, no qual foram estudados os diferentes testes educacionais e onde se começou a produzir material de avaliação, composto de provas objetivas, para as últimas séries do ensino médio, nas áreas de linguagem, matemática, ciências físicas e naturais e estudos sociais. Neste estudo, incluiu-se um questionário sobre características socioeconômicas dos alunos e suas aspirações. No CETPP, desenvolveram-se cursos sobre elaboração de provas objetivas com especialistas estrangeiros e fizeram-se várias publicações sobre temas ligados à avaliação educacional. Em meados dos anos 70, temos, pela iniciativa do ECIEL – Programa de Estudos Conjuntos de Integração Econômica Latino-americana –, um estudo avaliativo de porte sobre os determinantes dos níveis de escolaridade e do rendimento escolar obtidos por alunos com diferentes características pessoais e socioeconômicas. Aos alunos foi aplicado questionário para levantamento de dados sobre a situação socioeconômica, atitudes relacionadas com o processo escolar e aspirações, e um exame de compreensão de leitura e ciências. Foram coletados

dados de diretores, professores e escolas. Este estudo foi realizado simultaneamente em vários países da América Latina e seus resultados amplamente discutidos.

Não houve, nos anos subseqüentes, outras iniciativas visando estabelecer avaliações mais abrangentes, embora alguns ensaios localizados tenham sido realizados. Quanto à formação na área, nos cursos de Pedagogia desenvolvia-se uma disciplina sobre medidas educacionais que, ao longo do tempo, foi sendo suprimida. Algumas instituições ainda mantiveram, em sua grade curricular, uma disciplina mais geral sobre avaliação educacional.

No início da década de 70, além da preocupação com o nível técnico dos grandes vestibulares unificados, sobre os quais alguns estudos são realizados, procura-se preparar pessoas em serviço para lidar com essas modalidades avaliativas. Em contrapartida, as administrações públicas não se mostram preocupadas com uma avaliação sistemática do rendimento escolar dos alunos, pois o foco era o gerenciamento do cotidiano da educação.

Marques (1981) relata estudo realizado nesse período, com a finalidade de chegar a um instrumento de medida que permitisse verificar a situação do ensino de 1ª série. Grandes amostras foram utilizadas para testar o instrumento em suas duas formas, com crianças amostradas em todas as regiões geográficas do país, abrangendo leitura, escrita e matemática. Este projeto derivou da experiência com outros estudos realizados durante a década de 70, no então estado da Guanabara, e outros feitos em convênio com o CBPE/Inep, desde a década de 60.

Em 1980, foi realizado um estudo avaliativo pelo Departamento de Planejamento da Secretaria Municipal de Educação da cidade de São Paulo, para verificar o nível de escolaridade dos alunos da Rede Municipal de Ensino, em provas que abrangeram língua portuguesa, matemática e ciências. Avaliaram-se todas as crianças do terceiro estágio de educação infantil até a 7ª série do ensino fundamental, e o ensino médio foi contemplado com a avaliação dos alunos do primeiro ano. Foi construída uma bateria de testes específicos, baseada no currículo da Secretaria Municipal de Educação, à época. A idéia era utilizar esses dados para planejamento das escolas e de atividades extra-escolares. No entanto, com a mudança de administração, em 1982, a equipe municipal responsável dispersou-se, e os resultados não foram utilizados.

Em outra vertente, a da avaliação de políticas educacionais e de programas, encontramos, entre 1978 e 1982, alguns estudos que se propõem a avaliar a política

nacional de educação, com enfoque sobre a seletividade da educação, utilizando dados demográficos e até estudos de caso (FCC, 1981; BARRETO, 1981; ROSEMBERG, 1981; MELLO, 1981; SILVA, 1981). Esses estudos abriram um outro leque de possibilidades em avaliação, numa direção diferente da avaliação de desempenho.

Como se observa, algumas competências vinham se formando, lenta e esparsamente, na área de avaliação, sobretudo quanto à medida de rendimento escolar. Algumas administrações públicas da educação mostraram interesse nesses estudos, mas a eles não deram continuidade, seja em relação à formação de equipes, seja no que se refere aos estudos. A descontinuidade era a marca não só com referência aos estudos avaliativos, como também a todos os aspectos de iniciativas de administrações anteriores. Observe-se que, apesar dos estudos e esforços formativos empreendidos, o número de profissionais dedicados a essa área era muito pequeno, considerado o país e o tamanho de nosso sistema de ensino. Os modelos quantitativos necessários também dificultavam a formação de quadros em área na qual os profissionais tinham um perfil de formação mais humanista. Nesse período – fim dos anos 70 e início dos 80 – o pensamento educacional volta-se para a crítica dos empirismos, tecnicismos e operacionalizações burocráticas, e a área de avaliação, que mal começara a formar quadros, sofre uma grande retração nas universidades.

Durante a década de 80, uma experiência que pode ser vista como um marco na história de desenvolvimento de estudos sobre avaliação de programas, envolvendo não só o rendimento escolar, mas também outros fatores, foi a do Projeto Edurural que se desenvolvia em todos os estados do nordeste brasileiro. A avaliação acompanhou a implementação e o desenvolvimento do projeto, sob vários aspectos, de 1982 a 1986. Analisaram-se as formas de gerenciamento geral desse projeto – sistema de monitorias, professores, organizações municipais de ensino, alunos e famílias – e, amostralmente, o local. Desenvolveram-se testes para avaliar crianças de 2ª e 4ª séries, considerando sua pertinência: zona rural em classes, geralmente, multisseriadas. As provas foram elaboradas a partir de amostras de exercícios e trabalhos colhidos nas escolas dos três estados onde a avaliação se desenvolvia – Piauí, Ceará e Pernambuco. Para isso, coletou-se material de alunos, trabalhou-se com o pessoal da região e, depois, deu-se forma final às provas. O que se tentava era construir um conjunto de provas tanto quanto possível adequado àquela realidade, buscando conferir maior validade aos dados de rendimento

escolar. As crianças foram avaliadas em 82, 84 e 86. Nos mesmos anos, os demais dados sobre o programa – gerenciamento, escolas, pessoal docente, famílias, etc. – também foram coletados, e análises integradas e multivariadas construídas com o objetivo de encontrar traços característicos relevantes para verificar sua contribuição socioeducacional. Análises qualitativas dos dados levantados foram desenvolvidas, agregando-se estudos de caso ao modelo avaliativo para análise e compreensão, em maior profundidade, de situações mais específicas. (UFC, 1988, 7 v.).

Esse projeto avaliativo foi realmente amplo; utilizaram-se meios de análise quantitativa mais sofisticados, sem abandonar o cuidado com a qualidade daquilo que se estava observando. No decorrer do processo, realizaram-se seminários e analisaram-se seus resultados. Além disso, muitas outras análises derivaram desses dados: das políticas públicas, do que acontece com a escola que se situa na própria casa da professora, ou do problema das escolas que estavam em assentamentos dos sem-terra.

Com metodologia clara e bem definida durante a coleta e análise do material, mostrou-se o que se poderia fazer com estudos dessa natureza, na direção de propugnar 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 presentes os pressupostos sobre os quais se assentava o modelo avaliativo, ofereceu uma oportunidade para, a partir de uma ação direta, formar pessoas e pensar a área criticamente.

No fim dos anos 80, no Ministério da Educação, discutia-se sobre vários problemas que ocorriam nos sistemas educacionais apontados pelos pesquisadores da área de educação. No centro da discussão, estava o fracasso escolar brutal que se constatava pelas estatísticas e estudos no interior das escolas. Uma das questões levantadas é que não se possuíam dados sobre o rendimento escolar dos alunos quanto aos sistemas e fatores a ele associados. Nesse momento, promoveu-se, no MEC, uma oficina de trabalho com um grupo de educadores para debater a questão. Aventava-se a possibilidade de que, com a mudança da Constituição, pudesse haver uma nova estrutura política, o que sinalizava a necessidade de uma mudança na política oficial em relação aos ensinos fundamental e médio. Isso levou o Ministério da Educação a intervir menos diretamente nestes níveis e a propiciar orientações e subsídios técnicos, assumindo um papel avaliador como referência para políticas na área da educação básica.

No fim de 1987, foi proposta uma avaliação de rendimento escolar em dez capitais, para aquilatar se um processo de avaliação mais amplo 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, com provas de português (incluindo redação), matemática e ciências – um estudo piloto para verificar a viabilidade desse processo, a receptividade da escola quanto a esse tipo de avaliação, a adequação das provas, etc. Utilizou-se a teoria clássica em avaliação, e o grande desafio foi mesmo a construção de provas adequadas aos diferentes públicos dos estados envolvidos, o que foi possível por meio de parcerias locais.

Com o relativo sucesso dessa etapa, as discussões sobre os resultados das provas e dos fatores a eles associados e os debates realizados 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 & GATTI, 1988; VIANNA, 1988; VIANNA, 1989a; VIANNA, 1989b). Paralelamente, com apoio da Secretaria de Ensino de Segundo Grau do MEC, realizou-se uma avaliação de rendimento escolar de alunos da 3^a série do segundo grau (hoje ensino médio), abrangendo todas as modalidades existentes (geral, normal, técnico industrial, comercial etc.). Nesse estudo, além dos alunos das escolas públicas, foram avaliados os de escolas privadas, incorporadas, segundo relatos, com alguma dificuldade.

Também em 1988, como extensão à iniciativa do MEC, a Secretaria de Estado da Educação do Paraná realizou uma avaliação dos alunos de 2^a e 4^a séries em suas escolas. 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 proposto pelo MEC, realizou-se, em escolas privadas de 11 estados e no Distrito Federal, o mesmo tipo de avaliação feito nas escolas públicas. Verifica-se, então, que o desempenho dos alunos da escola privada era pouco superior aos da escola pública. (VIANNA & FRANCO, 1993)

Essas avaliações de rendimento escolar foram realizadas entre 1987 e 1991 e, além do impacto dos resultados nas Secretarias de Educação e na mídia, propiciaram um certo aprimoramento de pessoal, instrumentos e formas de aplicação. No entanto, esse trabalho era desenvolvido fora das universidades, por instituições especializadas e pelos órgãos públicos. Com os resultados obtidos e os dados das escolas, professores e alunos, dispunha-se de muitas informações sobre as quais se

poderia debruçar, refletir e tirar inferências tanto para políticas de sistemas quanto para questões da aprendizagem de alunos. Embora tenha havido um pequeno aproveitamento por algumas administrações – o que está demonstrado em seus documentos – não houve aproveitamento, aprofundamento ou estudos críticos dos dados por parte de outros estudiosos. Estes primeiros estudos mais abrangentes de alunos e escolas serviram de base para o desenvolvimento, nos anos 90, do Saeb – Sistema de Avaliação da Educação Brasileira.

Anos 90

Logo no início da década de 90, o Brasil participa do segundo Programa Internacional de Avaliação de Proficiência Educacional. Esse programa internacional, que à época envolvia 27 países, pretendia ser comparativo. Uma equipe internacional de especialistas – chineses, árabes, americanos, suecos, holandeses, entre outros – elaborou testes específicos e adotou uma metodologia de aplicação que permitia o processo de validação recíproca. No Brasil, o estudo ficou restrito a duas capitais (Fortaleza e São Paulo), comprometendo a comparabilidade que se pretendia. Foram avaliadas crianças de 13 anos de idade, não importando a série. 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 nem aquelas com excessiva defasagem idade-série, problema que não se apresentava para a maior parte dos países participantes.

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. Realizada amostralmente em todos os estados, no início com provas objetivas seguindo o modelo clássico, essa avaliação passou a ser realizada anualmente em todas as séries dos ensinos fundamental e médio, aperfeiçoando-se seus métodos de montagem e aplicação. Finalmente, em 1995, adotaram-se os procedimentos preconizados pela Teoria da Resposta ao Item, que, no entender dos especialistas desta avaliação do sistema, poderia oferecer informações mais amplas sobre o repertório escolar das crianças e jovens e 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 prover informações para tomadas de decisão quanto a diversos aspectos das políticas educacionais e também para pesquisas e discussões, a partir da geração e organização de dados sobre o desempenho acadêmico dos alunos no sistema e fatores a ele associados.

Hoje, dispõe-se de extensa base de dados sobre as sucessivas avaliações feitas no país e relatórios que disponibilizam as várias análises realizadas. O Saeb compõe-se de dois grandes eixos: o primeiro voltado para o acesso ao ensino básico, no qual se verifica o atendimento à demanda (taxas de acesso e de escolarização) e a eficiência (taxas de produtividade, de transição e de eficiência interna); o segundo, correspondente à qualidade, implica o estudo de quatro dimensões relativas: 1) ao produto – desempenho do aluno quanto à aprendizagem de conteúdos e desenvolvimento de habilidades e competências; 2) ao contexto – nível socioeconômico dos alunos, hábitos de estudo, perfil e condições de trabalho dos docentes e diretores, tipo de escola, grau de autonomia e matriz organizacional; 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, instalações, equipamentos, recursos e materiais didáticos. Os instrumentos de coleta são as provas dos alunos; os questionários aplicados com eles, com professores e diretores; questionários sobre as condições da escola. Nos dois últimos anos, agregaram-se, ao modelo geral de avaliação, estudos de caso de situações específicas consideradas relevantes para análise mais aprofundada de fatores contextuais intervenientes nos resultados. (INEP, 2002; BRASIL, 1995; PESTANA, 1999; FONTANIVE & KLEIN, 2000)

O Saeb vem sendo objeto de vários estudos e discussões, inclusive de comissões do Ministério, na busca de seu aperfeiçoamento. Problemas técnicos têm sido arrolados, por exemplo, quanto à modelagem das provas, o teor dos itens e sua validade; quanto ao processo de amostragem, vem sendo aperfeiçoado; com relação aos problemas de divulgação, disseminação e apropriação dos resultados nos diferentes níveis de gestão do sistema e os referentes aos professores. Franco (2001) nos oferece observações interessantes sobre potencialidades e problemas relativos a essa avaliação, pontuando sugestões para tornar o Saeb mais relevante para gestores e professores. Ilustra, com exemplos claros, as dificuldades dos alunos; propõe alternativas pedagógicas em linguagem adequada, para explicitar melhor os

fatores que contam em educação (efeitos escola e sala de aula); apresenta as escalas em linguagem e esquemas gráficos facilmente inteligíveis para não-especialistas; propugna um relacionamento mais efetivo com secretarias estaduais e municipais para apropriação dos resultados, além de outras sugestões.

Iniciativas regionais

Durante os anos 90, diversas avaliações de sistemas escolares tomam corpo também em várias regiões do país, além do âmbito nacional já exposto. Assim, em 1992 foi realizada, no estado de São Paulo, uma avaliação dos alunos das Escolas-Padrão, projeto de inovação curricular desenvolvido pela Secretaria Estadual de Educação, inicialmente com 306 escolas no estado, ampliando-se, depois, para mais de mil, com o objetivo de abranger, aos poucos, todas as escolas do sistema, o que não chegou a acontecer. A intenção era verificar se, com os insumos educacionais diferenciados que escolas e professores recebiam, melhoraria a aprendizagem escolar. Para isso, foi realizada uma avaliação com todos os alunos de 8ª série, envolvendo Língua Portuguesa, Redação, Matemática, Ciências, História e Geografia, e aplicado um questionário para sondagem de características dos estudantes.

Essa avaliação seria o marco zero, para se comparar o futuro desenvolvimento educacional das crianças. Esperava-se, no início dos trabalhos, realizar avaliações sucessivas, durante vários anos, com dados da progressão dessas crianças vinculadas ao Projeto da Escola-Padrão. Pretendia-se que, com a implementação permanente do projeto, elas pudessem apresentar um ganho significativo em termos de aprendizagem escolar. Os primeiros resultados foram interpretados psicopedagogicamente, e recomendações curriculares, feitas e consolidadas em documentos distribuídos às escolas.

Em 1993, dando continuidade a esse programa, a Secretaria de Educação propôs algumas análises mais aprofundadas sobre aqueles dados que poderiam subsidiar algum tipo de ação específica mais dirigida. O primeiro estudo realizado foi a identificação das escolas em que o desempenho escolar das crianças se mostrava superior ao conjunto mediano, em cada disciplina e no conjunto delas. Procurou-se localizar as 10, 12, 15 escolas onde esse rendimento sobressaía e aquelas onde ele se apresentava mais comprometido, não para classificá-las, mas para analisar as condições específicas e características de seu funcionamento que estariam determinando a melhor capacidade de atuação, ou, então, os problemas

que poderiam existir para as crianças apresentarem rendimento deficitário, apesar de estarem numa Escola-Padrão. Aproveitaram-se os dados desta avaliação para um estudo sobre alguns municípios específicos, considerando seus contextos – Santos, São José dos Campos, Campinas e Grande São Paulo. Na Grande São Paulo, o estudo dividiu a análise dos resultados entre a cidade de São Paulo propriamente dita e os demais municípios para comparar o que acontecia nas zonas mais centrais e nas mais periféricas.

Uma terceira análise aprofundou os resultados do componente curricular Língua Portuguesa, identificando-se nas provas algumas habilidades específicas, a saber: habilidade para responder a itens de informação que dependiam muito de memorização, ou àqueles que envolviam compreensão de conceitos ou relações, ou aos mais aplicados a situações que simulavam realidades. Nos relatórios constata-se que, durante a análise dos itens da prova de Língua Portuguesa com professores da rede e alguns especialistas, encontrou-se, em certos casos, grande dificuldade de dizer se um item poderia ser classificado como de compreensão ou aplicação, ou mais de compreensão, ou mais de aplicação. Isso resultou em estudos e discussões que permitiram aperfeiçoar a compreensão do significado desses ‘tipos de habilidade’. Nesse momento, trabalhou-se com vários conjuntos de professores que se transformaram em ‘juízes’, para classificar o item como de ‘informação’, ‘compreensão’, ou ‘aplicação’. Após várias discussões, definiram-se itens consensuais, o que possibilitou examinar o comportamento das crianças. Se fosse verdade, como comumente se diz, que nossas escolas só desenvolviam a memória das crianças, se decorar fosse realmente o objeto de trabalho dos professores, então elas deveriam apresentar bom desempenho nos itens de informação, que exigiam uma data, o nome de uma pessoa e dados factuais. No entanto, não foi isso o que se encontrou: as crianças apresentaram, pelo menos nesse estudo, melhor rendimento nos itens de aplicação, ou seja, nos que envolviam um problema prático e uma certa compreensão conceitual para resolvê-los, e não nos itens de informação. Ainda nessa avaliação, procurou-se verificar o tipo de padrão de avaliação que os itens da prova apresentavam, a partir de uma análise dos itens que foram construídos. Para isso, aplicou-se, pela primeira vez no Brasil, a metodologia da Teoria da Resposta ao Item, que permitiu a verificação do comportamento desses itens e da potencialidade dessa metodologia, em termos de estimativa de desempenho em diferentes níveis de habilidades. (SEESP, 1993 e 1994; GATTI, 1996)

A continuação desse modelo avaliativo, que pretendia ser longitudinal, foi interrompida por mudança na administração da Secretaria de Estado da Educação, mas permitiu um estudo avaliativo comparativo entre as Escolas- Padrão – cujo número agora já estava bem ampliado – e as Não-Padrão. A avaliação, conduzida em 1994, foi feita com uma amostra que abrangia crianças da 4ª à 8ª série – as de 4ª série foram avaliadas em Língua Portuguesa, Redação e Matemática; as de 8ª série, em Língua Portuguesa, Redação, Matemática, História e Geografia, o que permitiu, mediante algumas condições, uma comparação com as crianças avaliadas dois anos antes.

Nesse momento, fim de governo, foram feitas análises do desempenho pelas Delegacias Regionais existentes à época, por Delegacias de Ensino e por turno (diurno e noturno), comparando-se sempre as crianças de Escolas-Padrão e Escolas Não-Padrão. Nessa etapa, aplicou-se também um questionário aos alunos, que foi um aperfeiçoamento daquele que havia sido aplicado, levantando-se características dos estudantes como sexo, tipo de escola onde estudaram, instrução dos pais, trabalho, repetência de séries, preferência por leituras, etc. Nenhum cruzamento foi feito desses questionários com os dados de rendimento; eles apenas se destinaram a fornecer um perfil compreensivo das condições de vida, de trabalho, familiares e de alguns valores dos alunos. Todos os dados foram divulgados para a rede. Enfatizava-se que se deveriam enfrentar, com critério e determinação, os pontos críticos revelados no processo de ensino. Em ambas as etapas, as provas foram analisadas, e implicações pedagógicas foram apontadas com recomendações quanto aos pontos críticos de cada uma das disciplinas avaliadas. (SEESP, 1994; MONDEL & MALUF, 1994)

No mesmo período, entre 1992 e 1994, outro projeto de avaliação vinha sendo desenvolvido, amostralmente, nas escolas estaduais do estado de São Paulo, como 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 três anos, com o objetivo complexo de analisar mudanças de aprendizagem e de características cognitivas ao longo do tempo. Esse tipo de pesquisa em avaliação foi desenvolvido pela primeira vez no país e incluiu um estudo antropológico paralelo, permitindo uma compreensão profunda do que se passava no cotidiano das escolas e no desenvolvimento cognitivo das crianças (NEUBAUER, DAVIS & ESPÓSITO, 1996). A partir de 1995, instituiu-se no estado,

como parte de um programa para a educação pública, o Saesp – Sistema de Avaliação do Rendimento Escolar no Estado de São Paulo –, que se realiza até hoje e abrange séries sucessivas, com questionários informativos e o estabelecimento das equalizações necessárias, ano a ano, para obter-se comparabilidade. Neste modelo, todos os alunos matriculados nas séries envolvidas são avaliados. (SEESP, 1996; SEESP, 1998; ESPÓSITO, 2000)

Também o estado de Minas Gerais começa, em 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, que incorporavam outras informações por meio de questionários. Em 1998, com a implantação da progressão continuada nas escolas, passou-se a fazer a avaliação todos os anos. Com este programa desenvolveu-se uma proposta formativa de quadros na própria Secretaria, promovendo-se cursos específicos para os técnicos, aqui e no exterior. (GOULART, 1992; ANTUNES, XAVIER & FREITAS, 1992; SOUZA, 1999)

Iniciativas dessa natureza estenderam-se para outros estados da federação:

- 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. O processo teve seqüência nos anos subseqüentes e, em 2000, introduziu, na avaliação 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 4ª e 8ª séries. As matrizes de referência das provas trazem inovações em sua concepção – em Língua Portuguesa, por exemplo, contempla cinco áreas: procedimentos de leitura, recursos expressivos e efeitos de sentido, variação lingüística, coerência e coesão e conhecimento lingüístico operacional. Análises diversificadas encontram-se consolidadas em relatórios publicados e distribuídos à rede. A disseminação desses resultados fez-se 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 foram feitas, no Boletim da Escola, uma síntese específica da situação da escola e uma descrição dos níveis de desempenho alcançados. (SEEPR, 2001)

- O estado do Ceará desenvolveu um sistema avaliativo de suas escolas a partir de 1996, com avaliação, por amostras, de desempenho de alunos do ensino fundamental. Realizou também um processo de avaliação institucional de escolas, com envolvimento da comunidade, o qual constituiu um modelo diferenciado em relação ao que se fez em outros estados.
- Na Bahia iniciou-se, em 1999, um projeto de avaliação externa de seu sistema de ensino, com a assinatura de um convênio com a Universidade Federal da Bahia e o suporte de seu Centro de Estudos Interdisciplinares do Setor Público (ISP). Foram avaliados conhecimentos em português e matemática de alunos da rede pública urbana de 45 municípios; em 2000, esta avaliação abrangeu 126 municípios e, em 2001, 274. Esses municípios são parceiros no Programa Educar para Vencer. Constituiu-se uma Agência de Avaliação UFBA-ISP/Fapex, e o projeto desenvolve duas ações básicas: 1) a avaliação de desempenho, que é anual e informa cada escola sobre o desempenho de seus alunos (agregado por programa de ensino, disciplina e série) a partir de uma orientação curricular comum, com o objetivo de a escola discutir os dados e adaptar seu plano de desenvolvimento (PDE); 2) a avaliação da aprendizagem procura informar a escola e cada professor sobre o desempenho dos alunos ao longo do ano letivo, aplicando três testes por ano; a informação é apresentada por domínio de conteúdo para que o professor possa diagnosticar as deficiências de seus alunos e definir ações de recuperação durante o curso. (PROJETO DE AVALIAÇÃO, s.d.)
- O estado do Rio de Janeiro também realizou uma avaliação externa do Programa de Governo Nova Escola. Esta comportou uma avaliação de desempenho escolar em anos sucessivos, no fim dos anos 90, e uma avaliação da gestão escolar. A avaliação de desempenho foi feita em séries dos ensinos fundamental e médio, pautada numa escala de habilidades cognitivas e competências que comportava seis níveis. Para cada série construiu-se uma escala contínua a partir dos percentuais de acerto dos alunos nos itens das provas, o que permitiu uma

descrição das competências e habilidades que eles demonstraram possuir ao se situarem em um dos níveis definidos para a série a que pertenciam. Realizaram-se uma análise geral e uma por escola. O processo avaliativo da gestão escolar das escolas centrou-se em oito funções administrativas, identificadas como as de maior importância para a construção coletiva de uma nova escola. Com base nelas, foram elaborados os instrumentos que, uma vez aplicados, permitiram a produção de indicadores específicos (planejamento, gerenciamento de pessoas, etc.) e de um indicador geral da gestão de cada uma das escolas. (SEERJ & F. CESGRANRIO, 2002)

- Em Pernambuco, desde 1987, iniciativas foram tomadas para avaliações relativas às 1ª e 3ª séries, nas áreas de Língua Portuguesa (compreensão de texto) e iniciação matemática, especialmente em regiões onde se detectavam mais dificuldades escolares. Realizaram-se prospecções também em 1988 e 1989, nas quais o processo avaliativo foi acompanhado de ações diretas que visavam a superar as dificuldades encontradas nas escolas, redundando em significativa melhora nos resultados dos alunos (WEBER, 1991). Esse processo é interrompido e retomado a partir de 1995, com suas características um pouco modificadas. Atualmente, com a instituição de um sistema de avaliação, tem-se aperfeiçoado o processo.

Contribuições e desafios

Essas experiências, de mais amplo espectro, contribuíram para a formação e o aperfeiçoamento de um certo número, embora ainda limitado, de pesquisadores e técnicos em avaliação de desempenho escolar e institucional e em sistemas de ensino, tanto nas Secretarias quanto fora delas. Depreende-se, pelos documentos consultados, que se procurava desencadear um trabalho sociocomunicativo que conferisse a avaliações dessa natureza o sentido que deveriam ter: alimento para revisões críticas, porém construtivas, e base de apoio à progressão dos alunos. Pensava-se, portanto, em provocar mudanças na representação dos processos avaliativos que tinham conotação punitiva e depreciativa, dando-lhes novo significado. Todos sabemos das dificuldades de mudanças dessa natureza – mudanças socioculturais –, mas tentava-se estimular, no nível das gestões, o início desse processo.

Fatores internos e externos interpuseram-se, e as contradições implicadas nesses procedimentos tomaram corpo. De um lado, ênfase muito forte nos melhores e nos piores desempenhos, com a mídia valorizando apenas a média das notas obtidas e seu respectivo resultado, e não outros até mais importantes; de outro, a parca utilização dos dados pelas redes, seja por processos inadequados de visibilidade, acesso e disseminação dos dados, seja pelo hábito de, na área educacional, não se lidar pedagogicamente com esse tipo de dado. Estes problemas de utilização dos dados por diretores de escola, coordenadores pedagógicos e professores vêm ocorrendo tanto nas avaliações estaduais quanto nas nacionais. A disseminação adequada, em suportes diferentes para audiências diferentes, entre elas o público em geral e os pais, e o aperfeiçoamento da divulgação jornalística também são pontos nevrálgicos a serem resolvidos no desenvolvimento desses processos.

Avaliações institucionais e de desempenho no Ensino Superior

As questões de avaliação da qualidade dos cursos superiores oferecidos entrou em pauta quando, no fim dos anos 70, a Capes propôs um sistema de avaliação dos cursos de pós-graduação *stricto sensu* (mestrados e doutorados) para referenciar a sua qualidade e subsidiar seu credenciamento. Este sistema estruturou-se e aperfeiçoou-se no tempo, não sem problemas, polêmicas e discussões, e mantém seu papel até hoje, servindo como referencial de uma de experiência bem sucedida com relação a seus objetivos. No entanto, atualmente o debate sobre este ponto é acirrado, e o modelo vigente vem sendo posto em questão por vários estudiosos. (ANPED, 1999; CORREA, 1998; MASCARENHAS, 1998)

Antes dos anos 80, não há iniciativas assinaláveis de avaliação institucional ou de sistema para os cursos superiores de graduação, mesmo porque esse sistema era bem pequeno e altamente seletivo. Em 1983, foi lançado, pelo Ministério da Educação, o Paru – Programa de Avaliação da Reforma Universitária –, que pretendia desenvolver análises em dois aspectos principais: um, ligado à gestão das universidades, e outro, à produção e disseminação do conhecimento. Entretanto, a iniciativa durou apenas um ano, tendo sido suspensa por questões de desconti-

nuidade na administração pública que geraram disputas internas no Ministério da Educação.

Nessa década, no ensino superior, toma curso a idéia de avaliação institucional, associada à perspectiva de auto-avaliação das instituições. Na avaliação institucional, o objeto é a instituição, com sua estrutura, dinâmica, elementos relacionais e bases para o desempenho de suas funções. Além da compreensão, busca-se uma visão crítica sobre a instituição na comunidade em que se insere, segundo os fins a que se propõe. Grande impulso foi dado a essa perspectiva pela experiência pioneira da Universidade de Brasília, iniciada na década de 80. Indicadores foram desenvolvidos com a participação voluntária dos departamentos, num processo interativo com a equipe coordenadora da proposta. Outras instituições caminharam nessa direção, utilizando modelos próprios, como a Universidade de São Paulo e a Unicamp.

Foi, no entanto, com o Paiub – Programa Nacional de Avaliação Institucional das Universidades Brasileiras – que essa modalidade avaliativa se desenvolveu no ensino superior nos anos 90. Em 1993, o Ministério da Educação criou a Comissão Nacional de Avaliação das Universidades Brasileiras, composta de representantes das principais entidades associativas do ensino superior, para coordenar esse processo avaliativo e definir suas diretrizes. Depois de realizados seminários, discutidos os temas e feitas várias reuniões, a Andifes – Associação Nacional dos Dirigentes das Instituições Federais de Ensino Superior – elaborou um documento em que apresentava uma proposta, que, debatida, suplementada e enriquecida, gerou o Documento Básico orientador do Paiub. Foi então posto a público o primeiro edital convidando as universidades interessadas a apresentarem projetos avaliativos para o período 1994-1995, e outro em 1996. Os projetos que, submetidos à análise de um comitê assessor, foram aprovados, receberam apoio financeiro do MEC. Esse programa foi implantado em parceria com as instituições e recebeu o aval de seus representantes. A proposta era que, aos poucos, todas as universidades desenvolvessem projetos próprios, de acordo com suas características, embora houvesse indicadores básicos que deveriam ser elaborados; o momento adequado para iniciar a avaliação, seus procedimentos e/ou etapas ficavam a cargo da própria instituição. Até meados de 1996, cerca de 94 universidades estavam envolvidas com uma avaliação institucional, segundo as orientações do Paiub. Não obstante, o Ministério encerrou o projeto logo depois.

De todo modo, o programa possibilitou a formação de equipes nas universidades que procuraram maior especialização em teoria e métodos ligados à avaliação institucional, fazendo com que as preocupações com a qualidade institucional se expressassem dentro e fora das instituições, propiciando à avaliação instituir-se como processo público, com transparência, no interior do movimento maior do conjunto das universidades envolvidas ou não (MEC/SESU, 1994; TRINDADE, 1996; LUZ, 1996; RISTOFF, 1996; SANT'ANNA & VERAS, 1997). O projeto Paiub, entretanto, deixava de fora as instituições de ensino superior não universitárias que, à época, formavam a maioria dos graduados nas diversas áreas do país. Sobre estas, nenhuma avaliação havia incidido, e análises indicavam que as condições de muitos cursos nelas oferecidos eram precárias, sem corpo docente e infra-estrutura adequados, além de currículo empobrecido. A política aventada pelo Ministério da Educação a partir de 1995, e implantada em 1996, foi a de instituir um sistema de avaliação de todas as IES, com três grandes indicadores básicos: um indicador sobre a adequação de sua infra-estrutura, um sobre a qualificação do corpo docente e outro dos conhecimentos básicos de curso que os alunos deveriam ter aprendido na graduação. Implementado o modelo, os dois primeiros indicadores passaram a ser obtidos mediante questionários apropriados, e o último, por uma prova obrigatória, no fim do curso, sobre conhecimentos fundamentais relativos ao currículo do curso avaliado; não seria, pois, uma avaliação de conhecimentos profissionais, como o Exame de Ordem da Ordem dos Advogados do Brasil. Este último indicador – avaliação de conhecimentos –, apelidado de provão, seria o único a merecer atenção pública. Esse modelo de avaliação de cursos superiores implica a participação, no processo, de uma comissão de avaliação por curso. Essas Comissões têm representantes do Conselho de Reitores das Universidades Brasileiras, Associações Nacionais representativas de áreas de ensino e dos Conselhos Profissionais, compondo-se com professores das diferentes universidades. Essas comissões fazem as indicações para a elaboração das provas. Agregou-se ainda a esse trabalho uma análise dos cursos de graduação conduzida por comissões de especialistas.

Esse modelo, especialmente o provão, não foi implementado sem polêmicas e fortes reações; discutiram-se desde questões de princípios até problemas mais técnicos. Como todo processo regulador, ele apresenta algumas facetas que o justificam e outras que o condenam. Se, de um lado, tornou visível a fragilidade de condições de ensino de muitas instituições de ensino superior, levando mesmo

a um movimento de reformulação de currículos, de atividades docentes e infraestrutura em várias delas, de outro, esse processo mostrou-se um controlador do conhecimento, até com potencial reducionista para as aprendizagens, na medida em que propõe ajustes arbitrários em processos de formação de profissionais na graduação. Por isso, esse modelo demanda mais estudos sobre seus aspectos técnicos e suas implicações político-sociais na conjuntura em que se desenvolve, pois, conforme bem assinala Leite (1998: 60), fazer uma reflexão “sobre a educação superior, sobre a formação profissional que ela produz, implica também levar em conta as relações estruturais que configuram o ensino e a avaliação, a forma como a atividade educativa é determinada historicamente”.

Considerações finais

Tratando-se de um tema de tal amplitude, pelo fato de representar um campo que ainda carece de especialistas e pelo próprio processo de exacerbação das mudanças sociais das duas últimas décadas, não se pode dar finalidade conclusiva a este tópico, mas indicativa para a temática deste periódico – Avaliação Educacional e Mudanças. Podemos dizer que, para além das questões de ordem teórico-metodológicas e das estratégias do processo avaliativo conduzidas pelas políticas de educação no País ao longo das décadas aqui examinadas, convém demarcar nossas idéias dentro de um quadro mais referencial e inter-relacional com os contextos de mudanças, por vezes ruptural, pelas quais vem passando a sociedade brasileira.

Se, por um lado, nosso foco recaiu mais sobre a questão da avaliação no processo de ensino e aprendizagem, por outro, não podemos deixar de anotar que o terreno da avaliação não pode ser visto sem o conjunto das mudanças sociais, econômicas, políticas, culturais e epistemológicas – as chamadas mudanças de paradigma que alteram as formas de aprender, viver e produzir socialmente. Sabemos que continuamos ainda com os pés fincados nas linhas teóricas tradicionais, de forte carga psicologizante¹, quando pensamos em avaliação de ensino e aprendizagem e mesmo no aspecto mais geral.

As profundas mudanças que têm impactado as sociedades, induzidas principalmente pelo desenvolvimento dos meios de acesso e transmissão de mensagens, e pelas profundas alterações nos modos de produzir bens (inclusive os culturais),

¹Não se pode fazer a respeito da avaliação uma caracterização simplista que equalize as idéias de *centrada no sujeito* com abordagem rogeriana, de *centrada na situação* com abordagem ambientalista, de *tradicional* com abordagem behaviorista, de *progressista* com a abordagem construtivista, ou de *autoritária, anárquica, democrática, personalista*, com traços demarcados em uma específica abordagem teórica. Tudo isso indica o rompimento das demarcações isoladas das ciências e portanto do conhecimento. Cf. CARRARA KESTER. Avaliando a Avaliação: óticas teóricas e processo de construção da cidadania. In: ____ (org.). *Avaliação sob exame*. São Paulo: EAA/Fapesp, 2002.

²Fazemos uso aqui – não sem alguma resistência – da terminologia que se tornou comum nos documentos da Organização das Nações Unidas, Agências de Desenvolvimento e/ou Fomento, Organizações Não-Governamentais e Organizações Sociais Civis de Interesse Público – pessoa humana.

têm estimulado, como não poderia deixar de ser, reflexões acerca dos papéis – conseqüentemente, das exigências e expectativas – das agências educacionais e, extensivamente, da posição da pessoa humana² em sua relação com o planeta. Têm sido postos em cheque tanto as atribuições quanto as práticas pedagógicas e conteúdos que deveriam ser discutidos nestes espaços culturais, e a própria condição de instituição ‘civilizadora’ que definia a escola parece ter perdido força. Trata-se da velha questão: educar para quê?

Sob que parâmetros devemos, então, referenciar a pedagogia escolar – os do mundo do trabalho, os da família e comunidade, os dos homens e mulheres de sucesso, os da inteligência universitária, os da política...? Quem nos civilizará? Parece que somos obrigados a constatar que as sociedades estão prestes a imergir num processo penoso, longo e incerto de avaliação de seus sujeitos (o que implicará auto-avaliação, reflexão sobre a identidade e a ética do sujeito) e de suas organizações (que significa, entre outros procedimentos reflexivos, a discussão do papel social das empresas, da ética nas organizações, da submissão do trabalho).

Trata-se de um processo que se dá sem parâmetros definidos, sem valores e conceitos irrevogáveis, sem certezas e pontos definidos de chegada... pura aposta de que será necessário que nos submetamos à avaliação permanente, seja nas organizações coletivas, seja na consciência individual.

KEY WORDS: educational evaluation; evaluation of scholar productivity; institutional evaluation; evaluation standards; evaluation of educational systems.

ABSTRACT: The aim of this article is to write and discuss some experiences gathered from educational evaluation situations, more specifically from the largest ones, which have been applied to either educational systems or part of them since the 1960s. Besides, it not only focus on the processes that have been developed in elementary education and university teaching but also discusses the lack of both educational specialists and courses to teach people adequately for the field of educational evaluation, which causes some intellectual problems that have been in both theoretical and technical as well as social levels of this area.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANPED. *A avaliação da pós-graduação em debate*. São Paulo, 1999.
- ANTUNES, Ana Lúcia; XAVIER, Edir P. C.; FREITAS, Maria A. de. Avaliação do rendimento do aluno da escola estadual de Minas Gerais: o CBA. *Estudos em Avaliação Educacional*, São Paulo, FCC, n. 6, p. 29-42, jul./dez.1992.
- BRASIL. Ministério da Educação e do Desporto. *Resultados do Saeb 1995: a escola que os alunos freqüentam*. Brasília, 1995.
- BLOOM, Benjamin S et al. *Taxionomia de objetivos educacionais*. Porto Alegre: Globo, 1972. 2 v.
- CASTRO, Cláudio de M. *A escola que os brasileiros freqüentaram em 1985*. Ipea, 1986.
- CASTRO, Cláudio de M. & SANGUINETTY, Jorge A. *Custos e determinantes da educação na América Latina: resultados preliminares*. Rio de Janeiro, mar.1977.
- CASTRO, Cláudio M.; FRIGOTTO, Gaudêncio; MARTINS, Ricardo R.; CORDOVA, Rogério. *A educação na América Latina: estudo comparativo de custos e eficiência*. Rio de Janeiro: FGV, 1980.
- CORREA, Franklin R. Capes, Pronex, CNPq: por que usam critérios diferentes de avaliação? *Jornal da Ciência*, SBPC, 6 nov. 1998, p. 10.
- ESPOSITO, Yara L. (coord.). *Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo – Saresp 98: conhecendo os resultados da avaliação*. São Paulo: SEE/FDE, 2000.
- FONTANIVE, Nilma S. & KLEIN, Ruben. Uma visão sobre o Sistema de Avaliação da Educação Básica do Brasil – Saeb. *Ensaio*, Fundação Cesgranrio, v. 8, n. 29, p. 409-439, 2000.
- FRANCO, Creso. O Saeb – Sistema de Avaliação da Educação Básica: potencialidades, problemas e desafios. *Revista Brasileira de Educação*, n. 17, p.127-133, maio/jun./ago. 2001.
- FUNDAÇÃO CARLOS CHAGAS. *Educação e desenvolvimento social*. São Paulo, 1981.

8 v.

- FUNDAÇÃO GETÚLIO VARGAS. *Testes e medidas na educação*. Rio de Janeiro, 1970.
- GATTI, Bernardete A. Avaliação da jornada única em São Paulo. *Estudos em Avaliação Educacional*, São Paulo, FCC, n. 5, p. 85-90, jan./jun.1992.
- _____. Desenvolvimento de projetos de avaliação do sistema educacional no estado de São Paulo. *Estudos em Avaliação Educacional*, São Paulo, FCC, n. 13, p. 19-26, jan./jun.1996.
- _____. (coord). *Avaliação da dimensão qualidade no contexto do projeto Edurural: estudo de caso: a escola rural num município do interior nordestino*. São Paulo: FCC/DPE, 1981.
- GOULART, Íris B. Apresentação – O ciclo básico de alfabetização – Lições de uma avaliação. *Estudos em Avaliação Educacional*. São Paulo. FCC, nº 6, p. 9-28, jul./dez.1992.
- INEP. Sistema Nacional de Avaliação da Educação Básica: Saeb. *Textos divulgados para a imprensa: coletânea 1966-2001*, Brasília, Inep/MEC, 2002. p. 133-155.
- LEITE, Maria Cecília L. Avaliação e relações de poder: Paiub e Exame Nacional de Cursos. *Avaliação*, Campinas,. Unicamp, 1 (6): 59-68, mar.1998.
- MASCARENHAS, Sérgio. Avaliação pelo critério da árvore. *Jornal da Ciência*, SBPC, 28 ago 1998, p. 8.
- MONDEL, Lúcia M. & MALUF, Mônica M. B. Avaliação do rendimento de alunos das Escolas-Padrão. *Estudos em Avaliação Educacional*, São Paulo, FCC, n. 10, p. 103-122, jul./dez.1994.
- PROJETO DE AVALIAÇÃO. *Educar para vencer, Seeba*. Mimeo.
- PESTANA, Maria Inês G. S. O sistema de avaliação brasileiro. *Revista Brasileira de Estudos Pedagógicos*, Brasília, Inep/MEC, v. 79, n. 191, p. 65-73, 1999.
- PINHEIRO, Lúcia M. Medida de conteúdo de uma prova de Comunicação e Expressão: análise universitário. *Educação e Seleção*, São Paulo, FCC, n. 4, p. 87-108, jul./dez.1981.
- SANT'ANNA, Heloisa H. N. & VERAS, Neide F.M. O Programa de Avaliação Institucional das Universidades Brasileiras (Paiub). *Avaliação*, Campinas, Unicamp, 1 (3): 53-60, mar. 1997.
- SÃO PAULO. Secretaria da Educação. *Sistema de avaliação de rendimento escolar do*

- Estado de São Paulo – Saresp: relatório final dos resultados da 1ª aplicação.* São Paulo: SEE, 1996. 3 v.
- _____. *Sistema de avaliação de rendimento escolar do estado de São Paulo – Saresp.* São Paulo: SEE/FDE, 1998. 5 v.
- SEEPR, DG. COORDENAÇÃO DE INFORMAÇÕES EDUCACIONAIS. *Cadernos Ava,* 2000.
- SEERJ/FUNDAÇÃO CESGRANRIO. *Avaliação Externa do Programa Nova Escola, Relatórios.*
- SOUZA, Maria Alba de. A Experiência de Avaliação Educacional em Minas Gerais – 1992-1998. *Estudos em Avaliação Educacional,* São Paulo, FCC, n. 19, p. 57-76, jan./jun.1999.
- UNIVERSIDADE FEDERAL DO CEARA. *Edurural 1981-1985: avaliação da educação básica no nordeste brasileiro.* Fortaleza: UFC/FCC/DPE, 1988. Anexos. 7 v.
- VIANNA, Heraldo M. Avaliação do Rendimento de Alunos de Escolas de 1º grau da Rede Pública: um estudo em 39 cidades. *Educação e Seleção,* São Paulo FCC, n. 20, p. 5-56, jul./dez.1989.
- _____. Avaliação do Rendimento de Alunos de Escolas do 1º Grau da Rede Pública: um estudo em 20 cidades. *Educação e Seleção,* São Paulo, FCC, n. 19, p. 33-98, jan./jun.1989.
- VIANNA, Heraldo M. & GATTI, Bernardete A. Avaliação do Rendimento de alunos das 2ªs e 4ªs séries de Escolas Oficiais do Estado do Paraná. *Educação e Seleção,* São Paulo, FCC, n. 18, p. 5-62, jul./dez.1988.
- ____ & _____. Avaliação do Rendimento de Alunos de Escolas de 1º grau da Rede Pública: uma aplicação experimental em 10 cidades. *Educação e Seleção,* São Paulo. FCC, n. 17, p. 5-52, jan./jun. 1988.
- WEBER, Silke (org.). *Democratização, educação e cidadania: caminho do governo*

E
C
C
O
S

R
E
V.

C
I
E
N
T.

n. 1
v. 4

jun.
2002