

AVALIAÇÃO EDUCACIONAL NO BRASIL: experiências, problemas, recomendações

BERNARDETE A. GATTI¹

As reflexões apresentadas aqui baseiam-se na participação em diferentes projetos de avaliação nos quais equipes de pesquisadores da FCC têm se envolvido ao longo dos últimos anos. Ou seja, são considerações na perspectiva de quem não pertence ao sistema governamental mas que, de diferentes formas, tem atuado em parceria com este sistema.

No Brasil, temos vivido, no último decênio, algumas experiências de avaliação de desempenho do sistema educacional a nível do 1º grau realizadas de diferentes formas, na tentativa de associar rendimento escolar e alguns fatores, uns mais abrangentes, como por exemplo, regiões do país em que as escolas se situam, outros, incluindo, além disso, período de funcionamento da escola (diurno/noturno), ou condições das escolas, das famílias, dos professores, merenda escolar, custos etc... Várias disciplinas escolares foram incluídas nas avaliações realizadas, sendo constantes Língua Portuguesa e Matemática. Os modelos abrangeram parcelas dos sistemas (estaduais, municipais, particular), focalizando séries específicas (séries ímpares ou só 1ª ou 8ª série etc) e incluindo todos os alunos dessas séries ou apenas uma amostra deles. Esta avaliações se processaram por iniciativas governamentais com o propósito de se ter alguma informação sobre o nível de aprendizagem dos alunos e de seus condicionantes.

A primeira destas experiências – vamos falar apenas das mais abrangentes – na década de 80 foi a avaliação procedida no âmbito do **Projeto Edurural**.

1 Do Departamento de Pesquisas Educacionais da Fundação Carlos Chagas

Este projeto visava à melhoria das condições do ensino na zona rural dos Estados do Nordeste brasileiro. A pesquisa avaliativa do projeto (Gatti *et alii*, 1991) consistiu em um estudo, desenvolvido ao longo de cinco anos, com coleta de dados em uma amostra aleatória de 603 escolas em três Estados: Pernambuco, Piauí e Ceará. Constou de estudos sobre diversas variáveis (condições das escolas, merenda escolar, perfil de professores, impacto de treinamentos, condições das famílias dos alunos, etc.) e dentre elas o rendimento escolar das crianças. Foram testadas todas as crianças que frequentavam as 603 escolas nas 2^{as} e 4^{as} séries do primeiro grau em Língua Portuguesa e Matemática. Os testes foram constituídos com base em um levantamento das atividades curriculares efetivamente desenvolvidas nestas escolas, a partir do que se estabeleceu um conjunto de conhecimentos e habilidades que as crianças em cada um dos dois níveis deveriam ter minimamente dominado. Desta forma, os itens construídos para a prova visavam tão somente avaliar em que medida as crianças detinham esses conhecimentos, não apresentando esses itens dificuldades específicas para estabelecer discriminação entre elas. Nestas condições, esperava-se que as crianças obtivessem 100% de acerto ou próximo a isso.

Verificou-se, então, através das baixas taxas de acerto, que nas duas disciplinas, tanto para as 2^{as} como para as 4^{as} séries, a aprendizagem de conceitos básicos acha-se visivelmente prejudicada e segmentada. Não houve melhoria de desempenho, estatisticamente detectável, ao longo dos cinco anos; nem mesmo na última avaliação (1985) se notam ganhos nessa aprendizagem, apesar dos vultosos investimentos em construções, reformas, treinamentos, produção de material didático etc.. No modelo que associou todas as variáveis levantadas ao rendimento escolar (como variável dependente) não se conseguiu caracterizar com clareza fatores determinantes para este desempenho. No entanto, através dos dados de seis estudos etnográficos realizados em municípios destes três Estados, ao longo dos vários anos, foi possível trazer à tona algumas das razões subjacentes a este problema. Razões estas que permaneceram inalteradas durante todo o período da avaliação, e que perduram ainda, como se pode constatar por outros estudos realizados em anos subsequentes. Quais as razões encontradas para a permanência desse rendimento escolar tão baixo nas crianças desse setor social e desta região? A rotatividade das professoras nas escolas, cuja permanência apenas dependia da boa vontade e dos interesses pessoais e políticos dos prefeitos (treinada uma professora nada garantia que ela continuasse em sua escola; a designação da professora não se vinculava às suas condições de formação, mas a apadrinhamentos, etc.). As distribuições de livros, cadernos e merenda também dependiam da boa vontade e interesse dos prefeitos sendo inconstantes as condições desse atendimento. A complementação de salário prevista para as professoras nem sempre era providenciada a tempo pela prefeitura e mesmo o salário regular dessa professora (muito baixo, em geral 1/3 do salário mínimo) nem sempre era pago. As condições para o ensino multisseriado, característico destas escolas, per-

maneceram precárias: muitos alunos em classe, infra-estrutura deficiente, falta de condições do professor para o manejo de várias séries ao mesmo tempo. O tempo dedicado ao ensino propriamente dito era muito pequeno – duas a duas horas e meia por dia. Frequência irregular dos alunos em função de trabalho, doenças das crianças ou na sua família, condições difíceis para chegar à escola. Prevalência, na escola, da concepção de que aprender significava memorizar e, portanto, preponderava como "método de ensino" a repetição sem significado. Não incentivo à manifestação autônoma da criança; passividade induzida ao aluno. As professoras, presas às cartilhas ou livro didático (aquele disponível, em geral um só e não de sua escolha), não se utilizavam quer das experiências das crianças, quer dos espaços disponíveis no entorno da escola para desenvolver seu trabalho didático, não tendo sido nem incentivadas, nem orientadas para isto pelo sistema de supervisão pedagógica que se montou.

Tudo isto vem dar uma contundente validade à afirmação de que, sem esforço político consistente e adesão verdadeira de *lideranças locais* importantes, não há medida técnica e geral que resolva de fato, ou pelo menos em menor prazo de tempo, problemas educacionais das camadas sociais mais carentes.

Outra experiência foi a das avaliações de rendimento escolar segundo proposta do Ministério da Educação, e levadas a efeito entre 1987 e 1989.

Em 1987, o Ministério da Educação, através do Instituto Nacional de Estudos e Pesquisas Educacionais (INEP), pretendeu criar um programa externo de avaliação do 1º grau que subsidiasse as Secretarias de Estado da Educação com informações sobre problemas de aprendizagem escolar. Ao mesmo tempo, indiretamente, a proposta pretendia um maior envolvimento das Secretarias de Educação com as questões de avaliação em todo o sistema. Com o apoio dos Secretários de Educação foi celebrado entre elas e o Ministério da Educação um convênio para sua realização.

Ao definir a AVALIAÇÃO do RENDIMENTO em ESCOLAS DE 1º GRAU da REDE PÚBLICA em 1987, o MEC solicitou a colaboração da Fundação Carlos Chagas para a implementação do programa em todo o território nacional. A complexidade do projeto exigiu o seu desdobramento em fases distintas. Inicialmente, em caráter experimental, avaliaram-se 4.518 crianças de 10 cidades (10 capitais de Estado), matriculadas no 1º grau; a seguir, 8.069 alunos de 20 cidades diferentes em diferentes Estados foram avaliados no seu rendimento escolar, e, posteriormente, 14.868 crianças de 39 cidades, alunos de 157 escolas de 1º grau, foram pesquisadas em um terceira fase do projeto. No conjunto, foram avaliados 27.455 alunos de 238 escolas em 69 cidades localizadas em todos os Estados da Federação.

Os objetivos da pesquisa foram discutidos, inicialmente, com administradores e professores do 1º grau a fim de precisar a área da investigação e estabelecer os procedimentos para tratar e analisar os dados; ao mesmo tempo, foram levantadas informações relativas ao desenvolvimento dos programas de ensino e aos livros didáticos adotados.

A avaliação não procurou abranger a totalidade de alunos matriculados nem mesmo realizar uma amostra aleatória, o que, no presente caso, talvez inviabilizasse a realização do projeto, pela dificuldade de informações que embasassem uma boa amostra e a questão do acesso a muitas regiões. Optou-se por outro modelo. A investigação foi realizada em escolas escolhidas segundo alguns critérios previamente estabelecidos e os aplicadores foram contratados e treinados em São Paulo, na FCC, para garantir a confiabilidade dos resultados. Procurou-se abranger diferentes segmentos sócio-econômicos. O controle foi feito *a posteriori*, mostrando que a escola da rede pública, entretanto, era, à época, freqüentada na sua maioria por alunos oriundos das classes menos favorecidas da população, mesmo quando situadas em locais "nobres".

A amostra traduziu apenas um segmento do universo escolar em nível de 1º grau; desse modo, os resultados decorrentes da aplicação dos instrumentos de medida não visaram a estabelecer generalizações, mas a apresentar, nessa amostra, os pontos da aprendizagem em que os alunos não tiveram um desempenho satisfatório, para considerações do Ministério e Secretarias de Educação, dos técnicos e dos professores e a adoção de medidas que se fizessem necessárias.

A definição dos programas mínimos para balizar as provas partiu de um consenso entre professores com experiência no ensino de 1º grau. Não se partiu de um programa ideal, mas da programação vivenciada por professores e alunos. Ainda que existam programas estruturados por instituições oficiais, optou-se por uma abordagem diversa, ligada à realidade do dia-a-dia do magistério. Os programas, após a fase experimental, foram discutidos por elementos das Secretarias de Educação dos Estados participantes da pesquisa, reunidos em seminários durante dois dias em São Paulo, sendo incorporadas as sugestões apresentadas. A partir desses programas, foram elaboradas provas de PORTUGUÊS E MATEMÁTICA para as 1ªs e 3ªs séries, contendo cada uma 30 questões simples, que exigiam a apresentação de uma resposta. Provas de PORTUGUÊS, MATEMÁTICA E CIÊNCIAS, sob a forma de questões objetivas, também com 30 itens, foram construídas para os alunos das 5ªs e 7ªs séries. Para estas séries propôs-se, também, uma REDAÇÃO. Cada prova foi estudada através de coeficientes de fidedignidade e discriminação média, o que permitiu analisar o funcionamento desses instrumentos, que se mostraram adequados. A tendência geral das várias provas foi a de apresentar uma dificuldade mediana, fator que concorreu para uma boa consistência dos resultados e que permitiu definir diversos níveis de desempenho, facilitando a identificação de pontos críticos no quadro da aprendizagem dos alunos das diversas séries.

Ainda em 1987 realizou-se avaliação do rendimento de alunos de 2ªs e 4ªs séries de escolas oficiais de 1º grau do Estado do Paraná. Esta avaliação objetivou caracterizar o desempenho de alunos de 1º grau nas 2ªs e 4ªs séries nas escolas estaduais do Paraná nas áreas de Língua Portuguesa, Matemática,

Estudos Sociais e Ciências. A amostra abrangeu 41 escolas em sete cidades do Estado, tendo sido aplicadas 4.955 provas nas 2^{as} séries e 3.632 nas 4^{as} séries, num total de 8.587 provas. As provas foram construídas a partir de um programa mínimo definido para cada série e disciplina, a partir de programas levantados em escolas da cidade de Curitiba pelos técnicos da Secretaria. Como nos demais casos problemas no desempenho dos alunos foram encontrados. Na 4^{as} séries o rendimento escolar mostrou-se dramaticamente baixo nas provas de Português, Matemática e Estudos Sociais. Este fato foi amplamente discutido na Secretaria e nas Regionais de Ensino, com representantes de professores de todas as cidades envolvidas. Não se chegou a uma explicação satisfatória para a queda brutal de rendimento da 2^a para a 4^a série. Mas uma nova proposta curricular começou a ser discutida com os professores, estando hoje em vigor.

Um amplo estudo de rendimento escolar também foi realizado abrangendo outros segmentos da população, aqueles que têm condições de pagar pelo ensino de suas crianças e por isto optam por escolas da iniciativa privada. Vianna (1992) coordenou pesquisa em que um total de 5695 alunos do 1^o grau da rede privada de ensino, em 12 capitais de Estado, foram avaliados com os mesmos testes dos estudos na rede pública. A clientela das escolas estudadas foi classificada como de nível sócio-econômico médio e alto, composta por filhos de comerciantes, empresários, executivos e profissionais liberais. As escolas particulares, com poucas exceções, foram muito reticentes em permitir a aplicação das provas, alegando os mais diversos motivos e o estudo só foi levado a cabo por interferência do próprio Ministério da Educação e, assim mesmo, em um número muito menor de escolas do que o inicialmente pretendido.

De modo geral o desempenho destes alunos, em média, é superior ao desempenho dos alunos da escola pública. Isto indica tendencialmente uma relação entre condição social e rendimento escolar, aliás, já detectada pela heterogeneidade de desempenhos na escola pública nas diversas regiões do país. Observou-se que a diferença decresce à medida que os alunos progredem nas séries do 1^o grau, pois o rendimento dos alunos das escolas privadas decresce com o passar das séries. Os dados expostos por Vianna (1992) mostram que este desempenho, assim como os da escola pública, variam para melhor ou pior em algumas regiões. Os alunos do centro-oeste e sudeste apresentam melhores resultados. Há cidades em que os resultados dos alunos das escolas privadas, em algumas das provas, não se mostram significativamente diferentes dos da escola pública. Também nos resultados da prova de redação as diferenças entre as médias não foram significativas na 5^a série, notando-se que em algumas das cidades o desempenho da escola pública foi um pouco superior ao das crianças da rede privada. Mas cremos que há um dado interessante nestes estudos em que se verifica uma tendência geral de aumento da heterogeneidade de desempenho dos alunos das escolas privadas no avançar das séries, com uma maior inclinação para a homogeneidade das crianças da escola pública. Como sabemos, pelas estatísticas de reprovação e evasão, a seletividade é muito

intensa na escola pública, portanto ela atua mais sobre as camadas já desfavorecidas socialmente. Parece, então, que a homogeneidade crescente do rendimento dos alunos da escola pública é devido à exclusão impiedosa que dentro dela se opera, enquanto que na escola privada esta exclusão é bastante moderada, mantendo-se nesta escola alunos com desempenhos bastante discrepantes.

Através destes trabalhos parece de fato legítima a inferência feita de que o desempenho escolar associa-se a condições mais gerais das clientelas das escolas (públicas e privadas) e de que as diferenças regionais estão associadas às condições bem discrepantes de acesso a bens sociais nas diversas regiões do país, bem como em função do nível de desenvolvimento e industrialização que determinam não só melhores condições de renda, como também estruturas de relações de trabalho mais democráticas e vida política mais susceptível à opinião pública. Renda e escolaridade não existem num vazio, estão sim associadas a um contexto sócio-político que propicia determinadas condições de vida e de exigência de respostas das gerências públicas pelos usuários de seus serviços.

1. UMA EXPERIÊNCIA INTERNACIONAL

A Fundação Carlos Chagas participou ainda do Segundo Programa Internacional de Avaliação Educacional (IAEP)/1990. Esta participação se deu desde a fase de elaboração dos testes, de sua definição e consolidação, até sua aplicação e análise. No Brasil, avaliou-se crianças de 13 anos e os dados foram colhidos, por problemas de custos, somente em duas cidades brasileiras: São Paulo e Fortaleza. No caso, o processo de amostragem originalmente delineado teve de ser adaptado às condições do país e dos custos. É preciso lembrar que aqui uma grande quantidade de crianças aos 13 anos já está fora da escola (25%) e as que estão na escola distribuem-se por todas as séries (de 1ª a 8ª). Foram amostradas crianças de 13 anos, entre a 5ª e 8ª séries, nas cidades de São Paulo e Fortaleza, representando 3% da população-alvo. Os resultados mostraram, dentro dos critérios desta avaliação, um baixo desempenho das escolas brasileiras (públicas e particulares). Só para exemplificar, enquanto a Coreia obteve 73 pontos em Matemática e 78 em Ciências, em São Paulo se obteve 37 e 53 pontos, respectivamente.

2. AVALIAÇÃO DE RENDIMENTO DE ALUNOS PARA A SECRETARIA DE EDUCAÇÃO DO ESTADO DE SÃO PAULO

2.1 Avaliação das Escolas-Padrão

Em outubro de 1991, com a finalidade de recuperar o padrão de qualidade do ensino ministrado nas escolas públicas, o governo do Estado de São Paulo instituiu, na rede estadual de ensino, o Projeto Educacional das "Escolas-

Padrão", abrangendo o ensino fundamental e médio. Em relação a essa proposta, a Secretaria da Educação desencadeou uma série de ações de supervisão e acompanhamento, dando início a um programa de avaliação sistemática da qualidade de ensino oferecido pelas escolas-padrão. Esse programa abrange vários conjuntos de variáveis, entre elas a do rendimento escolar. Quanto a este aspecto, a avaliação, em sua primeira etapa, focalizou o rendimento dos alunos que estão freqüentando a última série do 1º grau (8ª). Para tanto, em junho de 1992, todos os alunos das 306 Escolas-Padrão, matriculados nessa série, submeteram-se à aplicação de uma prova em uma das seguintes áreas do conhecimento: Língua Portuguesa, Matemática, Ciências, História e Geografia (as três últimas agrupadas em uma prova de Conhecimentos Gerais). Foram testados alunos do período diurno e do noturno, com provas diferentes mas relativamente equivalentes. Além das provas, a avaliação incluiu dois questionários: um dirigido aos estudantes e outro que foi respondido pelos diretores das unidades escolares. As informações levadas por meio desses instrumentos deveriam permitir a identificação de algumas variáveis que poderiam estar influenciando o processo de ensino-aprendizagem. Complementando o levantamento de dados solicitou-se aos professores que tecessem comentários a respeito das provas aplicadas. Para esta avaliação solicitou-se, também, aos alunos, a produção de um texto narrativo-descrito adotando-se a forma "continue a história a partir do início proposto". Duas situações-estímulo foram criadas, uma para o grupo dos alunos do período diurno e outra para os do período noturno. Adotou-se para o julgamento dos textos um conjunto de critérios que foram expressos em notas numa escala com variação de 1 a 100 pontos. Realizaram as provas 25.198 alunos. As dificuldades encontradas confirmam o panorama de um problemático rendimento escolar e de dificuldades na fixação de aprendizagens básicas. O quadro é piorado no que se refere à área de Matemática. Quanto às redações, comparando-se os resultados verifica-se que o desempenho dos alunos do período noturno é muito heterogêneo e que eles obtêm resultados inferiores aos de seus pares das turmas do diurno, o que é compreensível até pelo fato de serem, em boa parte, alunos-trabalhadores e pelas condições em que se opera o ensino noturno.

Estes dados foram utilizados pelos professores para seu planejamento escolar específico, diferentemente do que ocorreu com os dados de avaliação do primeiro projeto do MEC. A utilização dos resultados pelos professores, através de uma análise orientada sobre o significado dos erros (construiu-se para isto um Manual) é um aspecto que tem causado impacto nessas escolas. Outro ponto importante foi a avaliação das provas pelos próprios docentes que, ao se manifestarem justificando sua análise, trouxeram contribuições importantes para a continuidade do processo de avaliação, além de terem refletido sobre a questão.

Obtidos os principais resultados relativos ao desempenho escolar e de posse de informações que caracterizavam as diferentes unidades escolares deu-se início a um tratamento estatístico (*análise de variância*) visando iden-

tificar que fatores poderiam estar influenciando nos resultados alcançados pelos alunos.

Os resultados obtidos indicam que para as turmas do período diurno melhores rendimentos são obtidos pelos estudantes das 8^{as} séries que freqüentam escolas que dispõem de um corpo docente mais estável e experiente. No noturno, o fator mais importante, que contribui para a existência de resultados mais favoráveis, é a presença de uma biblioteca contendo um acervo maior e mais diversificado.

Os demais fatores analisados não mostraram contribuição estatisticamente significativa na explicação dos resultados alcançados, embora se registre uma variabilidade que diferencia o desempenho dos alunos quando se focaliza o conjunto de escolas de determinadas Divisões Regionais de Ensino, ou seja, em algumas regiões há uma tendência à predominância de bom desempenho, ao passo que em outras ocorre o inverso.

2.2 Avaliação do Impacto da Jornada Única em São Paulo

Este projeto foi posto em ação no segundo semestre de 1992 e deverá ser desenvolvido até 1994. Vamos situar o contexto desta avaliação: em 1984 foi instituído em São Paulo o Ciclo Básico de Alfabetização, experiência que hoje vários Estados implantaram. No Ciclo Básico a criança faz as duas primeiras séries sem reprovação formal. Discutindo dados do Ciclo Básico, que foram sendo coletados durante toda a experiência, segundo documento da própria Secretaria de Educação, hipotetizou-se que, talvez, fosse mais interessante dar mais tempo de permanência diária na escola tanto ao professor como ao aluno. Então, em 1988, instituiu-se a Jornada Única para crianças e professores do Ciclo Básico. Ou seja, o professor das duas primeiras séries trabalharia por seis horas, regeria uma classe só e teria vantagens salariais para se dedicar a essa única classe.

Esperava-se que, para as crianças, vantagens pudessem existir na medida em que o professor teria tempo para corrigir trabalhos dos alunos, preparar aulas, participar de reuniões com os demais professores que lecionassem no Ciclo Básico, enfim, melhorar seu perfil profissional. As crianças não ficariam as seis horas com seu professor: haveria um período em que elas estariam fazendo a merenda, um período em que elas estariam desenvolvendo atividades de educação física e atividades de educação artística, o que daria tempo ao professor da classe para fazer outros tipos de trabalho. Então, a proposta da Jornada Única veio como uma proposta de um novo tipo de permanência na escola para o professor e para o aluno, visando propiciar ao professor, dentro da escola, um espaço para trabalho pedagógico, e, para o aluno, diversificar seu contato já com outros tipos de professores, participando de outros tipos de atividades que cooperam para o seu desenvolvimento: seriam as atividades ligadas à educação física e à educação artística.

A proposta envolveria ainda distribuição de material pedagógico e instruções para seu uso adequado, e cursos de reciclagem realizados quer no período de permanência do professor na escola, quer fora.

Hoje, praticamente noventa por cento das escolas de 1º grau em São Paulo já têm a Jornada Única implantada no Ciclo Básico. Então, a questão é: Ciclo Básico/Jornada Única causam algum impacto, algum diferencial no desempenho e no desenvolvimento intelectual das crianças? Há diferenciais no desenvolvimento das habilidades básicas que devem ser adquiridas? Altera-se a relação aprovação/reprovação/evasão?

Embora a Jornada Única esteja sendo desenvolvida em todo o Estado, o projeto de avaliação se restringe à Grande São Paulo. Todos os insumos que estão sendo propiciados para o Ciclo Básico e a Jornada Única melhoraram significativamente o nível de aproveitamento escolar dos alunos? Quais as evidências desta melhoria? Isto está sendo analisado em dois níveis: num nível macro, com as estatísticas de reprovação, evasão, fluxo de alunos etc., e em outro nível, através de uma avaliação de desempenho. Nesta avaliação de desempenho, prevê-se a aplicação de testes de rendimento escolar por três anos seguidos, comparando alunos do Ciclo Básico/Jornada Única com alunos que se matricularam em escolas que não têm Jornada Única. Os testes examinarão questões de Língua Portuguesa e aprendizado em Matemática. Este mesmo grupo de alunos será acompanhado no segundo e no terceiro ano, para que se possa verificar, neste contínuo, o que sucederá com eles em termos de aprendizagem de conteúdos escolares. O estudo está sendo feito por amostragem tendo envolvido de início 3.600 alunos distribuídos em 60 escolas.

Os focos de análise são: 1) os antecedentes sociais dos escolares, considerados como variável exógenas (renda familiar, trajetória de vida, renda per capita mensal, escolaridade dos pais, padrões sócio-culturais da família etc.); 2) a estrutura e dinâmica da experiência escolar, que são as variáveis endógenas (condições materiais da unidade escolar, perfil profissional e sócio-econômico do diretor, dos professores, dos coordenadores do Ciclo Básico, integração da equipe escolar, proposta pedagógica da escola para o Ciclo Básico, situação do fornecimento de merenda, relação escola-comunidade etc.).

Especial cuidado está sendo dedicado à coleta destes dados, inclusive no sentido de que eles sejam passíveis (posteriormente, já que vários são qualitativos na sua origem) de transformarem-se em indicadores que possam ser confrontados com o rendimento escolar.

Além deste estudo longitudinal, realiza-se um estudo tipo antropológico em seis escolas. Estas escolas foram selecionadas segundo alguns critérios específicos que definem heterogeneidade de condições. Pretende-se fazer uma análise em profundidade do processo ensino-aprendizagem (os métodos de ensino, o que o professor está fazendo dentro da sala de aula, qual o seu comportamento pedagógico, seu desempenho, como é que ele estimula as crianças etc.) através de um método observacional, em campo, com imersão

integral nestas escolas. Para uma sub-amostra de crianças nestas seis escolas avaliar-se-á o desenvolvimento cognitivo com um instrumental pesquisado e construído especificamente para este fim. Com este instrumento não se pretende avaliar aprendizagem de conteúdos escolares mas habilidades, funções gerais e específicas da cognição.

3. REFLEXÕES E RECOMENDAÇÕES SOBRE ESSES PROCESSOS DE AVALIAÇÃO

3.1 Considerações Específicas

Ao se trabalhar com amostras grandes – como era o caso da avaliação do Projeto Edurural onde se procurava garantir a representatividade dos resultados obtidos – nem sempre é possível assegurar condições ideais de trabalho. Notadamente nessa região – onde impera uma política clientelista – as dificuldades normalmente encontradas são exacerbadas. Assim, convém arrolar alguns dos principais impasses com os quais se deparou nessa avaliação, evidenciando que não se pode fugir ao fato de que elementos culturais devem ser considerados no desenvolvimento de modelos de avaliação: 1) Quanto ao delineamento da pesquisa: uma das exigências do Projeto foi a adoção de um modelo experimental, onde municípios recebendo insumos do Edurural fossem comparados a outros excluídos de tais benefícios. Entretanto, foi ficando claro, no decorrer da pesquisa, que os Estados enviavam todos os recursos recebidos – fossem eles ou não advindos do Edurural – para a Coordenadoria de Educação Rural, que os distribuía, segundo seus próprios critérios, para os municípios. Desta forma, ficou-se sem "grupo controle", fato que dificultou verificar o impacto comparativo do Edurural. 2) Seleção Aleatória das Escolas: foi frequente a constatação da presença de escolas "fantasmas", registradas só para fazer número. Desta forma, vários sorteios fizeram-se necessários para determinar quais seriam as escolas investigadas. Deve-se ter em mente que as escolas "fantasmas" ocasionaram, também, um problema para a amostra, no que diz respeito às suas bases estatísticas. 3) Seleção dos Aplicadores: constatou-se que os critérios estipulados nem sempre foram seguidos, dando-se preferência a pessoas que, de um ou outro modo, ligavam-se, por laços de compadrio, a pessoas influentes do município (prefeito, vereadores, grandes comerciantes ou proprietários de terras) 4) Instabilidade das escolas: verificou-se que, de uma para outra avaliação, as escolas antes visitadas deixavam de existir, sendo necessário selecionar – agora intencionalmente – outra que, nas cercanias, apresentasse perfil semelhante. 5) Aplicação das provas: durante a correção das provas foram observadas certas particularidades que evidenciavam a ocorrência de interferências externas no momento em que os alunos estavam respondendo aos testes, no sentido de ajudá-los, ou seja, havia indícios que punham em dúvida a validade da resposta do aluno. No caso, foi feita uma triagem que não nos garante totalmente que o problema tenha sido contornado. De qualquer

modo, tomadas as devidas precauções, as considerações tecidas sobre o desempenho dos alunos nas provas de Português e Matemática permitiram uma categorização inicial do processo de ensino e aprendizagem a que estavam submetidos os alunos da zona rural dos três Estados da amostra.

3.2 Considerações Gerais

A primeira questão a abordar é que a avaliação necessita considerar aspectos do meio educacional em que se desenvolverá, evitando a imposição de modelos nem sempre ajustáveis aos vários aspectos do sistema, ainda que bem estruturados do ponto de vista teórico. A realidade brasileira possui aspectos de grande singularidade, sendo preferível a adoção de algumas idéias gerais para orientação e fundamentação do trabalho de avaliação que permitam contornar áreas de atrito cultural e político, buscando preservar tanto a consistência dos dados (saber-se o que são de fato), como a própria continuidade do processo avaliativo. Assim, a pesquisa em avaliação educacional, inclusive quando restrita ao rendimento escolar, exige definições e procedimento *ad hoc*, para que possa ser realizada com pleno êxito. É evidente que, não havendo um consenso – difícil, às vezes, entre educadores –, ocorram, especialmente nessa área, incompreensões quase sempre centradas em aspectos não necessariamente relevantes para os propósitos da avaliação. Desta forma, é preciso conviver com as limitações. A amostragem na avaliação educacional, assim como em outros tipos de investigação na área de educação no Brasil, sofre restrições decorrentes dos dados estatísticos, além dos problemas de tempo, de pessoal e, especialmente, de custos. Há uma questão cultural também a considerar que é a da pouca credibilidade nos estudos amostrais entre administradores, técnicos, professores.

Importante, no caso específico de uma avaliação educacional, com vistas ao levantamento de habilidades escolares dos estudantes, é considerar a situação crítica das estatísticas sobre educação no Brasil. A estruturação de uma amostra para fins de avaliação educacional deveria contar com informações seguras da localização e número de escolas e dos índices de matrícula, repetência, promoção e evasão existentes; contudo, é preciso convir que as informações disponíveis são bastante falhas. A questão da amostragem em avaliação educacional é complexa e, às vezes, difícil de solucionar, porque, malgrado o rigor dos procedimentos lógico-estatísticos, sempre subsiste a pergunta: a amostra selecionada é representativa do universo investigado? qual o qualificativo real dessa representatividade?

O praticante da avaliação educacional, freqüentemente, necessita posicionar-se em relação a problemas aparentemente menores, mas que perturbam o seu trabalho, se não forem equacionados. As decisões sobre esses problemas são subjetivas na maioria das vezes. A vivência do avaliador e seu bom senso encontrarão uma solução para as situações que, imprevisivelmente, possam surgir. Assim, o avaliador, na construção dos instrumentos de

mensuração, vê-se às voltas com a definição de um conteúdo programático que corresponda à média do que ocorre nas escolas do sistema, no caso de uma avaliação do rendimento. Há várias programações, nenhuma obrigatória. Assim, o avaliador vê-se compelido muitas vezes a criar um programa mínimo, a partir dos aspectos comuns às várias programações. Uma solução alternativa e parcial para o problema estaria na análise dos livros didáticos utilizados, que se constituem em fonte subsidiária de informações sobre o que é desenvolvido em sala de aula pelos professores. O procedimento não foge à prática ortodoxa da avaliação, sendo usado, inclusive, em outros complexos educacionais, para a construção dos chamados testes padronizados. Ajusta-se também à nossa realidade, pois é sabido que, em muitas escolas, o livro didático é o verdadeiro "programa" curricular, como já foi comprovado em pesquisas nacionais.

A problemática de pessoal precisa ser considerada quando a natureza do trabalho exige material humano especializado, especialmente na fase de elaboração dos instrumentos. Pouquíssimos são os que, efetivamente, se dedicam à avaliação e têm uma formação adequada. Esta questão de pessoal habilitado é um dos aspectos mais sérios envolvendo a problemática do desenvolvimento de procedimentos avaliativos, na educação, entre nós. Quase não há cursos para prover esta formação e quando os há são insuficientes em sua programação.

A isso acrescentam-se os custos de um processo avaliativo a nível de sistemas, que o tornam às vezes quase inviável em termos orçamentários, apesar de perfeitamente desenhado segundo as técnicas de amostragem. Note-se que recursos como o correio e a aplicação, sem supervisão, pelo professor são impensáveis em nossa realidade pelas implicações na validade dos dados, entre outras, como já se pode verificar em vários estudos.

Embora se tenha em algumas regiões competências um pouco mais desenvolvidas não é o caso geral. Assim, além das condições materiais precárias com que se conta, soma-se a falta de qualificação dos recursos humanos disponíveis até mesmo para aplicar os instrumentos de avaliação, registrar seus resultados e, em especial, para coletar os dados com um mínimo de confiabilidade. Na verdade, avaliar adquire, no contexto brasileiro, um significado extremamente ameaçador, na medida em que, ao apontar aspectos frágeis da situação de ensino-aprendizagem, pode-se, em princípio, desencadear ações sancionadoras. Isto introduz problemas quanto à sua participação em processos avaliativos mais amplos.

Professores e técnicos de ensino, ao invés de perceberem na avaliação uma possibilidade de aprimorarem seu trabalho, construindo uma escola de qualidade, sentem-se, em geral, muito inseguros quanto ao seu futuro profissional: eventuais resultados negativos podem implicar um fracasso no exercício da profissão e apontar responsabilidade não cumpridas. Nesta medida, salvo raras exceções, a avaliação, no entender das equipes escolares brasileiras, não se refere a um processo interativo que deve informar tanto os alunos sobre suas

dificuldades como – e sobretudo – os professores, diretores e os órgãos centrais sobre a adequação ou não do trabalho que desenvolvem, tendo em vista reformulações com bases mais concretas. Há de certa maneira, no ambiente escolar, um clima que não favorece o desenvolvimento de processos avaliativos externos e de âmbito mais amplo e, neste sentido, há muito o que fazer para demonstrar o caráter não punitivo dos processos de avaliação e, sim, seu caráter pedagógico e auxiliar na superação de dificuldades.

A realidade educacional enfrentada pelos avaliadores é bem diversa do ideal, exigindo a adoção de práticas acauteladoras para o bom êxito do trabalho. A utilização de professores integrados no sistema educacional, para elaboração de itens, independentemente da região ou área geográfica em que exerçam suas atividades, pode contribuir para validar o processo de avaliação, desde que sua experiência resulte da prática objetiva do ensino, na realidade do dia-a-dia.

O uso de provas objetivas ou discursivas costuma ser apresentado como um problema para alguns avaliadores e para os professores em geral. É o espírito "anti-cruzinha". A questão, em si, é um falso dilema, considerando-se que existe literatura, com base empírica, demonstrando, de modo insofismável, a adequação de ambos os tipos de prova para medir praticamente os mesmos traços, sendo, pois, inteiramente ociosa qualquer discussão sobre o assunto. Ao avaliador o problema deve se apresentar de uma outra forma, considerando a natureza da investigação e a população avaliada. Dependendo do contexto, talvez sejam recomendados instrumentos semi-objetivos apenas, ou, então, inteiramente objetivos, em função do tempo disponível, de quem vai responder (idade etc.), da problemática relativa à sua aplicação e aos custos operacionais. A discussão sobre itens discursivos/objetivos não é bizantina, como poderia parecer à primeira vista, merece reflexão, mas não deve ser colocada em termos extremados, pois depende de um concurso de variáveis que interagem, cabendo assim ao avaliador tomar a decisão que melhor se ajuste ao quadro investigativo.

Lembremos a dificuldade de se testar crianças das escolas particulares que, via de regra, por várias desculpas, não permitem o acesso a elas. Isto cria dificuldades para se pensar numa avaliação geral de sistema.

3.3 Considerações Finais

Inegavelmente, nas pesquisas relatadas, observou-se a mobilização de pesquisadores, técnicos de ensino, professores e alunos, todos dedicando seu tempo e esforços para que se conseguisse aquilatar a condição de aprendizagem oferecida por segmentos do sistema educacional. Gastaram-se recursos na esperança de que os centros de decisão se sensibilizassem para promover uma educação de maior qualidade onde, pelo menos, fossem garantidos o domínio da leitura, da escrita, da contagem básica e dos conhecimentos essenciais sobre o meio físico, social e cultural em que vive. Aliás, a origem destes estudos foram esses centros de decisão. E, aqui, há outro problema a ser levantado. A sucessão de Ministros e Secretários de Estado da Educação, com ou sem mudança de

governo e de partido, tem gerado uma dramática descontinuidade de informação e de utilização de resultados de projetos desenvolvidos no seu âmbito administrativo. Os que chegam não querem saber o que foi feito ou está sendo feito a partir da gestão anterior e dados e análises úteis se perdem nas prateleiras ou depósitos dos órgãos públicos. Dados para cuja coleta o poder público pagou.

Os resultados alcançados estão aí, à espera de quem os utilize de forma adequada, a serviço da Educação. Mostram uma dura realidade: se a escola brasileira não vem cumprindo nem mesmo sua função mais singela, o que se dizer a respeito da formação do espírito crítico, da engenhosidade, da elaboração do novo, da crença na capacidade de cada um – e de todos conjuntamente – transformarem o sofrido cotidiano deste país através da plena inserção na cidadania? Esta escola não deve – e nem pode mais – continuar assim. É preciso ação – e ação comprometida com o futuro da nação – para que se ofereça, efetivamente, *educação básica* para a população brasileira.

Todas essas avaliações oferecem um terreno fértil para se repensar o sentido que se quer atribuir à escola fundamental: aonde nos propomos verdadeiramente chegar, quando proclamamos a obrigatoriedade e, supostamente, nos comprometemos com o oferecimento de 8 anos de escolaridade básica? Como pretendemos, efetivamente, viabilizar uma proposta realista de ensino fundamental? Quanto ao processo de avaliação, somente a sua prática continuada oferecerá não só dados mais válidos, como seu aperfeiçoamento através da discussão construtiva das diferentes possibilidades e caminhos trilhados para o levantamento dos dados. Ignorar ou maquiar as alterações e percalços impostos ao modelo inicial pelo contexto real em nada contribui para o desenvolvimento teórico-prático da área de avaliação educacional.

Acreditamos que uma avaliação de conhecimentos de alunos, sob que modelo for, e das escolas só pode estar vinculada a um processo contínuo, o qual de fato sirva aos professores e ao sistema como parâmetro de aperfeiçoamento da qualidade do ensino. Isso requer um plano de ação de médio prazo, cuja viabilização esteja garantida por uma equipe designada para responder por ele, com cronograma de metas e responsabilidades definidos. Um processo dessa natureza só se implementa com tempo suficiente de maturação, discussão e divulgação, a partir de premissas claras quanto às propostas pedagógicas e de respaldo administrativo ágil.