

Problemas e Impasses da Avaliação de Projetos e Sistemas Educacionais: Dois Casos Brasileiros

Bernardete A. Gatti¹
Heraldo Marelim Vianna¹
Cláudia Davis¹

A década de 80, no Brasil, foi especialmente rica no relativo à elaboração e implantação de programas educacionais, notadamente nos Estados mais desenvolvidos, cuja pressão por mais e melhores escolas se fazia mais presente. Surgiram, desta forma, inovações nos sistemas públicos de ensino dos Estados, que se voltaram, basicamente, para as primeiras séries da escola básica. O objetivo principal de tais inovações era superar a maciça repetência aí observada, propiciando às escolas condições de abrigar as crianças por mais tempo, oferecendo-lhe, simultaneamente, melhores condições de aprendizagem. Destacam-se, por sua abrangência e impacto, nas redes públicas de seus respectivos Estados, as medidas envolvidas no Cíelo Básico (SP), na Jornada Única (SP), nos Centros Integrados de Educação Pública (CIEPs) (RJ) e as reformas nos currículos e nos sistemas de gerenciamento decorrentes de amplo debate com o conjunto dos educadores, em Minas Gerais e em Pernambuco.

Esta efervescência no cenário educacional, que se caracterizou por um afã em contornar problemas crônicos da educação brasileira, não se fez acompanhar por medidas que permitissem avaliar se as ações adotadas estavam atingindo os objetivos propostos, permitindo eventuais alterações em pontos de estrangulamento. Desta forma, conta-se, na maior parte dos casos, tão somente com medidas grosseiras, a saber as estatísticas oficiais que se referem às taxas de evasão, repetência e aprovação. É possível, desta forma, considerar que falta ao Brasil uma cultura avaliativa que permita conhecer a real situação educacional do país e que informe as políticas públicas na área.

Evidentemente, em especial nas regiões mais desenvolvidas do país, conta-se com um corpo técnico atualizado e competente para elaborar sofisticadas pesquisas avaliativas, cujos resultados possam subsidiar reformas curriculares, treinamento de professores, alterações nos métodos de ensino etc. Por outro lado, as pre-

¹ Pesquisadores da Fundação Carlos Chagas.

cárias condições materiais com que se conta - dificuldades de acesso, hospedagem, abastecimento de combustível e alimentação, entre outras - somam-se à falta de qualificação dos recursos humanos disponíveis 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 político brasileiro, um significado extremamente ameaçador, na medida em que, ao apontar aspectos frágeis da situação de ensino-aprendizagem, pode, em princípio, desencadear ações sansionadoras.

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 responsabilidades 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.

Todos esses fatores, em conjunto e em estreita articulação, fazem com que se procure simplificar ao máximo o delineamento da pesquisa avaliativa. É inconcebível, por exemplo, pensar-se em avaliar criatividade ou o desenvolvimento cognitivo ou, ainda, o processo de construção da língua escrita em estudos de cunho nacional e/ou regional. Tais medidas requerem a elaboração de instrumentos avaliativos de extrema complexidade (notadamente num país onde as diversidades regionais são tão acentuadas) que, para produzirem resultados fidedignos, exigem profissionais com formação adequada, raramente disponíveis nos diferentes Estados e/ou municípios. Desta maneira, por frustrante que seja para o pesquisador, há que se ater ao mais simples e, mesmo assim, ao mais importante: detectar qual é a real situação educacional com que se conta para, inclusive, se estimar a magnitude do esforço a ser empreendido para transformá-la no sentido almejado.

É evidente que a passagem pela escola só tem sentido quando se supõe que se sairá dela diferentemente da forma como nela se entrou. O aproveitamento escolar, desta maneira, nada mais é do que a apropriação que se faz daquilo que a escola pretende ensinar. Dentre os muitos objetivos da escolarização e, portanto, do aproveitamento escolar, três parecem essenciais:

1) objetivos que se referem à obtenção de informações sobre as conquistas das gerações precedentes, de forma a se apropriar de um conjunto de dados de natureza física, biológica e social sobre a realidade que se vive e se enfrenta;

2) objetivos que se referem à construção de funções cognitivas que permitam pensar e atuar sobre o mundo físico e social de maneira independente, crítica e criativa, estabelecendo relações cada vez mais complexas entre as informações disponíveis;

3) objetivos que se referem à elaboração de atitudes e valores, de modo a se contar com cidadãos que, conhecendo as variedades da condição humana, escolham de maneira lúcida, consciente e responsável sua conduta pessoal e social.

Naturalmente, estes objetivos encontram-se todos mesclados, sendo difícil estabelecer limites precisos entre cada um deles. A construção da escrita, por exemplo, contribui não só para a elaboração do pensamento simbólico, o estabelecimento de generalizações, o uso da memória e do raciocínio, o emprego de análises e sínteses como,

também, para uma forma de conhecer o mundo, de se posicionar frente ao mesmo, de formar opiniões sobre o ambiente circundante. Informações, desenvolvimento cognitivo e aquisição de valores se dão a um só tempo, ainda que, para efeitos didáticos, um objetivo possa preponderar sobre os demais.

Isto posto, gostaríamos de relatar a experiência de dois projetos de avaliação de grande porte – ambos conduzidos por pesquisadores da Fundação Carlos Chagas (SP) – que visavam a apreender tão somente os resultados da aprendizagem escolar. O primeiro deles refere-se à avaliação do Programa de Expansão e Melhoria do Ensino no Meio Rural do Nordeste Brasileiro – Edurural – e, o segundo, à avaliação do Rendimento de Alunos de Escolas de 1^a grau da Rede Pública dos vários Estados da Federação. Pretendemos através destas duas experiências apontar aspectos relativos ao cotidiano das avaliações que os modelos teóricos, em geral, não levam em consideração.

I. O EDURURAL: VISÃO GERAL

O Edurural foi um projeto planejado no ano de 1977, a partir de propostas decorrentes de entendimentos e compromissos assumidos por técnicos das diferentes Secretarias de Educação com o Ministério de Educação e Cultura. Tais propostas refletiam, razoavelmente, a forma através da qual as equipes locais equacionavam a problemática do ensino rural e acabaram por se consubstanciar num programa de intervenção na área a ser custeado pelo governo central. Este programa privilegiava, acentuadamente, a melhoria do ensino rural por intermédio de uma série de ações que deveriam levar à capacitação pessoal, à assistência técnica e à produção de material didático, uma vez que os técnicos locais estavam convencidos de que as condições físicas da escola rural – apesar de sua extrema penúria – não influíam, substancialmente, sobre o desempenho da rede escolar. Nos últimos anos da década de 70 o BIRD assumiu, em negociações com o governo central, 1/3 dos custos do programa, fato que redundou em reformulações e exigências quanto à sua execução. Dentre estas estava a de que houvesse uma avaliação externa, capaz de fornecer subsídios para a definição dos programas de educação rural. Os estudos e projetos de avaliação ficaram a cargo da Fundação Cearense de Pesquisa – encarregada dos aspectos institucionais ligados ao Edurural – e da Fundação Carlos Chagas – que deveria se centrar naqueles concernentes ao rendimento escolar.

1. O estudo do Rendimento Escolar na Zona Rural do Nordeste

A pesquisa consistiu num estudo transversal, com coleta de dados realizada em 1981, 1983 e 1985, nos Estados do Piauí e Pernambuco (tidos como representativos das tendências da região) e no Ceará (por ser este o Estado-sede da equipe central de pesquisa). Os resultados alcançados em 1981 foram considerados como sendo a linha-de-base que permitiria verificar se teriam havido avanços na aprendizagem escolar em função das ações desencadeadas a partir da injeção de insumos na região nos anos subsequentes. Tomou-se, pois, como uma das dimensões da qualidade do ensino o de-

sempenho dos alunos, aferindo, através de provas de rendimento académico, o nível de domínio das habilidades básicas nas áreas de Português e Matemática em alunos de 2^a e 4^a séries do ensino fundamental.

Estas séries foram escolhidas levando-se em conta a necessidade de conter a amostra dentro de proporções mínimas que garantissem a viabilidade do trabalho de campo e que mantivessem, ao mesmo tempo, as quantidades suficientes para o tratamento de dados desejado. Desta forma, decidiu-se pela 2^a série porque esta permitiria: 1) uma avaliação de modo mais consistente do que na 1^a série dos primeiros efeitos da escolarização no que diz respeito às disciplinas selecionadas e 2) uma verificação dos efeitos do programa sobre o problema da retenção na 1^a série que, se diminuísse, implicaria o aumento do contingente de alunos que permaneciam na escola e o melhor fluxo do sistema de ensino. Quanto à 4^a série, sua inclusão justificase pela importância de um estudo de eficiência e da qualidade do ensino rural, na medida em que reside aí sua terminalidade.

Ao lado dessa sistemática de avaliação da qualidade do ensino oferecido pela escola rural, procurou-se apreender, também, a forma de atuação desta escola em seu contexto concreto, dentro dos limites das relações sócio-econômico-culturais locais. Optou-se, assim, por realizar seis estudos de caso que, ao identificar as variáveis conjunturais que atuam no processo de ensino-aprendizagem, iluminariam a compreensão do que se passa nas escolas rurais ao nível do desempenho dos alunos. Embora não abordados diretamente neste trabalho, os dados dos estudos de caso foram muito importantes para a análise dos resultados aqui apresentados.

1.1 Delineamento da Amostra

A determinação do tamanho da amostra processou-se em duas etapas. Na primeira, através de uma análise das características específicas da população, determinou-se o tamanho desejável da amostra no total, por Estado e por modalidade de programas: os beneficiados pelos recursos do Edurural e aqueles que se utilizavam de insumos advindos de outros programas. Obteve-se com isto um quadro de proporções que permitiu determinar o número de escolas por Estado e tipo de programa. Desta forma, estabeleceu-se que se iria trabalhar com 6% do universo total de escolas ($NA = 9918$)², ou seja 603 unidades escolares.

Na segunda etapa, procurou-se determinar as escolas, dentro de cada Estado, que seriam pesquisadas. Estabeleceu-se que, em cada um deles, a amostra de escolas iria se concentrar em 10 municípios atendidos pelo Edurural e em 10 atingidos por outros programas. A determinação dos municípios deveria se dar através de um estudo com os técnicos das respectivas Secretarias de Educação, com o objetivo de detectar as condições peculiares sócio-estruturais específicas das microrregiões de cada Estado, recomendando-se que se incluíssem no programa Edurural os municípios mais carentes. A seleção das escolas seria aleatória.

2 Total obtido mediante listagens fornecidas pelas Secretarias de Educação, ano base 1980.

1.2. Elaboração dos Instrumentos

A elaboração dos instrumentos foi feita levando-se em conta que o processo de avaliação deveria oferecer, prioritariamente, subsídios às Secretarias de Estado para planejamento de atividades diretamente relacionadas com o processo ensino-aprendizagem. Com este pressuposto, foram consultadas, para fins de seleção de conteúdos, as seguintes fontes:

1. os planos construídos pelas Secretarias e o currículo orientador dos programas desenvolvidos nas escolas das zonas rurais dos Estados da amostra, quando estes existiam;
2. técnicos e pessoal envolvido com os órgãos municipais de ensino;
3. professoras trabalhando na zona rural; e
4. diários de classe, livros adotados nas áreas de Português e Matemática, planos de aula e cadernos dos alunos.

Os dados assim obtidos propiciaram informações sobre: 1) as condições de funcionamento das escolas rurais; 2) os conteúdos de Português e Matemática desenvolvidos nas 2ª e 4ª séries e tipos de exercícios empregados na sistematização destes conhecimentos; 3) os critérios de avaliação utilizados e 4) as percepções dos professores sobre as escolas nas quais trabalhavam. Com base nestas informações, procurou-se identificar os conteúdos de Português e Matemática que eram comuns aos três Estados e elaborou-se uma primeira versão das provas a serem utilizadas na pesquisa. Esta primeira versão foi pré-testada de forma a verificar falhas em sua organização e estruturação, bem como a influência, maior ou menor, destas falhas, nas dificuldades apresentadas pelos alunos. Face aos resultados obtidos no pré-teste, os eventuais problemas foram corrigidos e a versão final da prova elaborada.

1.3. Organização das Provas

A provas de Português e de Matemática pretendiam aferir o conhecimento a respeito de determinadas habilidades que já deveriam fazer parte do repertório de alunos de 2ª e 4ª séries. Os testes foram, então, organizados levando-se em conta que cada questão deveria se enquadrar dentro de um objetivo maior, através do qual determinados conteúdos seriam avaliados. Uma vez que estes conteúdos não visariam a selecionar mas, sim, a avaliar conhecimentos e habilidades fundamentais de alunos de 2ª e 4ª séries, a prova não era referenciada no grupo, mas no objetivo avaliado. Desta forma, os itens incluídos nos objetivos não apresentavam dificuldades que permitissem uma discriminação específica dos alunos, referindo-se a conhecimentos básicos que deveriam ser totalmente dominados ao final da 2ª e da 4ª série. Estabeleceu-se, portanto, o critério de 100% de acerto em cada objetivo da prova, ou seja, esperava-se que os alunos respondessem corretamente a todos os seus itens.

Nos quadros a seguir estão indicados os objetivos da prova de Português e Matemática para a 2ª e 4ª séries e o peso que receberam em função de sua importância no estágio de escolarização em que o aluno se encontrava. Como pode ser visto, procurou-se valorizar nas provas de Português as questões relativas às habilidades de Leitura e Interpretação e Escrita. Na 2ª e 4ª séries, estes objetivos correspondem, respectivamente

te, a 70% e 66% do valor total da prova. Em Matemática, na 2ª série, as questões que objetivavam medir conhecimentos relativos à Numeração e Cálculos representaram cerca de 80% do total de pontos. Na 4ª série, as questões que avaliavam conhecimentos de numeração, cálculos e sua aplicação consistiram em 84% do valor total (26%, 33% e 25%, respectivamente). Estes objetivos de Português e Matemática, a que se atribuíram maiores pesos, referem-se a conhecimentos e habilidades essenciais, nessa fase da aprendizagem escolar.

Quadro 1

Objetivos da Prova de Português de 2ª e 4ª série e o peso que receberam

2ª SÉRIE		4ª SÉRIE	
Objetivo	Peso do Objetivo	Objetivo	Peso do Objetivo
Leitura e Interpretação	36 pontos	Leitura e Interpretação	30 pontos
Escrita	34 pontos	Gramática	34 pontos
		Escrita	20 pontos
Gramática	30 pontos	Redação	16 pontos

Quadro 2

Objetivos da Prova de Matemática de 2ª e 4ª série e o peso que receberam

2ª SÉRIE		4ª SÉRIE	
Objetivo	Peso do Objetivo	Objetivo	Peso do Objetivo
Numeração	28 pontos	Numeração	26 pontos
Conceito de Dezena e Centena	8 pontos	Medidas de Capacidade e Comprimento	4 pontos
Conceito de Dúzia	6 pontos	Operações de Multiplicação e Divisão	11 pontos
Relações Numéricas: Dobro, Metade, Par e Ímpar	6 pontos	Números Racionais	2 pontos
Operações Elementares	Adição e Subtração	Unidades de Medidas e Uso de Notas e Moedas	10 pontos
	Multiplicação e Divisão		
Quatro operações	24 pontos	Operações Matemáticas	22 pontos
Situações Problemáticas	15 pontos	Situações Problemáticas	25 pontos

1.4. Critérios de Correção

As provas foram organizadas de tal forma que os alunos acertavam ou erravam a questão, tendo sido estabelecidos critérios para avaliar o desempenho nos diferentes objetivos.

1.5. Procedimentos

A aplicação dos testes de rendimento deveria ter sido realizada por pessoal recrutado na sede de cada município investigado a partir dos seguintes critérios: 1) estar cursando ou já ter completado o 2º grau e 2) já ter participado de pesquisas do censo. Para familiarizá-los com os instrumentos e para treiná-los para a aplicação dos mesmos foi elaborado um minucioso manual, onde se procurava elucidar cada passo a ser seguido e esclarecer qual a ação a ser tomada em situações de dúvida. O treinamento dos aplicadores deu-se em cada um dos Estados da amostra, teve sempre a duração de uma semana e foi conduzido, sistematicamente, por pesquisadores dos dois Centros de Pesquisa envolvidos, em colaboração com os técnicos das Secretarias de Educação locais. Durante toda a coleta de dados, os profissionais acima mencionados supervisionaram o trabalho em andamento.

1.6. Tratamento dos Dados

Uma vez corrigidas as provas, procurou-se comparar os resultados obtidos pelos alunos nos diferentes objetivos que compunham as provas de Português e Matemática. Para tanto, o total de pontos em cada um deles sofreu uma transformação linear que redundou em uma escala de 100 pontos, dividida em quatro faixas de avaliação. Na primeira delas estavam distribuídos os alunos que obtiveram de 0 a 25 pontos; na segunda, aqueles que conseguiram de 26 a 50; na terceira situavam-se os alunos que alcançaram de 51 a 75 pontos e, na quarta faixa, localizavam-se aqueles que receberam de 76 a 100 pontos.

1.7. Resultados Alcançados

Ao se analisar o desempenho geral dos alunos em função dos objetivos das provas, uma primeira constatação a ser feita é a de que os resultados obtidos estavam muito aquém do mínimo esperado tanto para a 2ª como para a 4ª série nas duas disciplinas avaliadas, sendo que em Matemática a situação foi mais crítica nas duas séries. E que este desempenho não apresentou diferenças significativas de uma etapa de avaliação para outra, não evidenciando impacto particular de qualquer programa nessa instância³. Lembrando-se que se trabalhou com referência a critério, es-

3 Cf. Relatórios FCP/FCC, 1982, 1985, 1988.

perando-se, pois, uma aproximação dos 100 pontos em cada prova, as Tabelas I e II nos dão uma idéia geral dos resultados obtidos.

TABELA I

DESEMPENHO DOS ALUNOS DE 2ª SÉRIE DOS TRÊS ESTADOS DA AMOSTRA, NAS PROVAS DE PORTUGUÊS E MATEMÁTICA EM 1981, 1983 E 1985, SEGUNDO A SUA PARTICIPAÇÃO NO PROGRAMA EDURURAL

ESTATÍSTICAS	PROVA DE PORTUGUÊS					
	EDURURAL			OUTROS PROGRAMAS		
	1981	1983	1985	1981	1983	1985
\bar{X}	48,3	60,2	60,2	44,2	55,7	58,1
sd	27,6	23,3	25,3	28,2	23,8	25,0
md	53,0	66,0	67,6	46,0	62,0	63,7

ESTATÍSTICAS	PROVA DE MATEMÁTICA					
	EDURURAL			OUTROS PROGRAMAS		
	1981	1983	1985	1981	1983	1985
\bar{X}	45,4	52,8	50,5	41,3	48,0	46,3
sd	26,7	25,6	25,3	25,0	23,0	24,2
md	46,0	54,0	51,3	40,0	47,0	46,1

TABELA II

DESEMPENHO DOS ALUNOS DA 4ª SÉRIE DOS TRÊS ESTADOS DA AMOSTRA, NAS PROVAS DE PORTUGUÊS E MATEMÁTICA EM 1981, 1983 E 1985, SEGUNDO PARTICIPAÇÃO NO PROGRAMA EDURURAL

ESTATÍSTICAS	PROVA DE PORTUGUÊS					
	EDURURAL			OUTROS PROGRAMAS		
	1981	1983	1985	1981	1983	1985
\bar{X}	52,1	52,2	48,7	49,4	52,7	47,7
sd	20,0	17,8	18,2	21,0	18,7	18,5
md	53,0	53,0	49,4	52,0	55,0	47,4

ESTATÍSTICAS	PROVA DE MATEMÁTICA					
	EDURURAL			OUTROS PROGRAMAS		
	1981	1983	1985	1981	1983	1985
\bar{X}	50,0	49,4	51,1	44,0	46,0	48,0
sd	25,0	23,7	23,9	23,4	22,3	22,6
md	51,0	50,0	52,7	42,0	45,0	47,6

Face aos dados e às análises detalhadas empreendidas, é possível concluir que, nas duas disciplinas investigadas, tanto para a 2ª como para a 4ª série, a aprendizagem de conceitos básicos se mostra visivelmente prejudicada e segmentada. Os estudos de caso, por sua vez, iluminaram as razões subjacentes ao fracasso generalizado que se encontrou, razões estas que permaneceram como características inalteradas ao longo de toda avaliação (5 anos):

- as condições para o ensino multisseriado eram precárias: havia muitos alunos em sala de aula, a infra-estrutura era deficiente e a própria formação da professora não permitia o manejo adequado a este tipo de trabalho. Desta forma, a programação das atividades pedagógicas era tal que, quando o docente dedicava sua atenção a um aluno, os demais permaneciam desocupados. Como consequência, a modalidade de atuação selecionada reduzia significativamente a efetividade do ensino;

- o tempo dedicado a atividade "estrito senso" pedagógicas era extremamente reduzido quando se considera que, além dos problemas acima mencionados, ainda se gastavam momentos preciosos em atividades rotineiras e burocráticas, como o eram a "chamada", a organização para a escrita, a formação de filas e mesmo a distribuição da merenda. A situação se torna mais grave face ao fato de que a frequência, por parte dos alunos, não era regular;

- prevalecia, na escola, a concepção de que aprender significava memorizar e, portanto, preponderava, como "método de ensino", a repetição sem significação;

- nas atividades "estrito senso" pedagógicas não se observou praticamente nenhuma participação ativa dos alunos. A passividade era tal que inexistiam perguntas para esclarecimento de dúvidas, propostas de novas atividades, formas alternativas de se solucionar uma dada questão. Diálogo efetivo em torno dos conteúdos escolares não se davam entre professora e alunos;

- as situações de ensino-aprendizagem permitiram considerar que não se levava em conta a experiência dos alunos na rotina escolar. De igual modo, não se incentivava a manifestação autônoma e original do pensamento uma vez que esta podia, em última instância, colocar em cheque o conhecimento docente. Obediência era valorizada e atribuía-se o insucesso escolar aos alunos que "não estudavam" o suficiente;

- a atividade escolar dos alunos raramente era controlada pela professora. Como consequência, não se tinha idéia do andamento das tarefas e dificilmente se atuava sobre dificuldades identificadas. Quando correções e chamadas de atenção ocorriam, elas incidiam sistematicamente sobre a forma e não sobre o conteúdo;

- o ensino se realizava exclusivamente em sala de aula, não se utilizando, em nenhum momento, dos espaços sociais disponíveis. Presa aos conteúdos veiculados nos livros didáticos, a professora não fazia uso de situações cotidianas (trabalho, lazer, etc.) nem de materiais acessíveis no meio ecológico e social para incentivar a criatividade, expandir a compreensão ou mesmo para promover a observação e experimentação, elementos indispensáveis ao raciocínio e à crítica das situações vividas. E nem foi incentivada ou orientada para isto.

2. Algumas considerações sobre os dados e os problemas enfrentados

Ao se trabalhar com amostras grandes – como o era a do 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, evidenciando que não se pode fugir ao fato de que elementos culturais devem ser considerados no desenvolvimento de modelos de avaliação:

– *O delineamento da pesquisa*: uma das exigências do BIRD 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 – fosse eles ou não advindos do Edurural – para a Coordenadoria de Educação Rural, que os distribuía, segundo seus próprios critérios, para os municípios. Assim, na verdade, não haviam diferenças substanciais entre as ações desenvolvidas nos 20 municípios investigados. Desta forma, ficou-se sem “grupo controle”, fato que impossibilitou verificar se os insumos específicos do Edurural – caso tivessem sido utilizados conforme o acordado – teriam provocado as melhorias almejadas.

– *Seleção dos Municípios que integrariam o Programa Edurural*: no Estado do Piauí, os 35 municípios escolhidos pela Secretaria de Educação para integrarem o programa em questão eram os mais desenvolvidos do Estado, a despeito da recomendação clara de que se privilegiassem os mais carentes. Foi necessária a intervenção do BIRD para que nova composição fosse armada.

– *Seleção Aleatória das Escolas*: foi freqüente a constatação da presença de escolas “fantasmas”, cujo objetivo principal era fornecer salário para supostas “diretoras, professoras e merendeiras”, em geral pessoas ligadas ao prefeito e que, na verdade, não trabalhavam. 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.

– *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). Este problema não pode ser superado, uma vez que a equipe central deparou-se com um fato consumado: ao chegar ao centro de treinamento, os aplicadores oriundos de cada município já lá se encontravam, não havendo nem tempo hábil, nem recursos, para se proceder a uma nova seleção.

– *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.

– *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. Procedeu-se, então, a uma análise minuciosa de cada prova a fim de se eliminar do tratamento de dados aquelas que indiscutivelmente apresentavam indícios que punham em dúvida a validade da resposta do aluno. A despeito deste cuidado, nada garante, que, nas provas onde a incidência de fraude não ficou suficientemente clara para permitir sua eliminação da análise, o mesmo problema não tinha ocorrido. Adicionalmente, ficou patente a convicção do aplicador local com a tentativa de mascarar os reais resultados da aprendizagem escolar.

- *Análise dos Resultados*: face ao acima relatado, os dados alcançados devem ser vistos com certas reservas, na medida em que podem não retratar o conhecimento real dos alunos nos conteúdos avaliados.

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, o qual se mostrou, no geral, relativamente impermeável à ações tais como foram desenvolvidas pelos projetos educacionais em curso. Problemas reais foram apontados que talvez pudessem redundar em medidas concretas, o que parece não ter ocorrido. Cabe perguntar também sobre o descompromisso das diferentes instâncias públicas face a informações disponíveis que poderiam alimentar políticas de trabalho educacional. Assim, nesse momento, parece importante, raciocinando com Demo⁴, apontar que "aos pobres cabe uma educação pobre, dada na medida das sobras da parte privilegiada do sistema", ressaltando-se que ninguém, nas regiões de estudo, faz parte dessa camada privilegiada.

Finalmente, convém aqui colocar que a educação dada ao pobre talvez seja mais pobre ainda do que se imagina ou do que pode ser demonstrado neste trabalho. Convém lembrar a constatação de que houve tentativas de mascarar o conhecimento real dos alunos, procurando auxiliá-los a responder adequadamente às questões da prova. Se há margem de dúvida quanto à total confiabilidade dos resultados apresentados e, se estes ficaram tão aquém de qualquer expectativa sobre os efeitos da escolaridade oferecida mesmo pela escola pobre, a precariedade da qualidade de ensino na zona rural fica ainda mais patente. Face à manutenção dessa precariedade, patente fica também que os objetivos da avaliação de servir de realimentação e redirecionamento de programas educacionais foram totalmente frustrados.

II. AVALIAÇÃO DO RENDIMENTO DE ALUNOS DE ESCOLAS DE 1º GRAU DA REDE PÚBLICA: UM ESTUDO EM 69 CIDADES

Em 1987, o Ministério da Educação, através do *Instituto Nacional de Estudos e Pesquisas Educacionais* (INEP), pretendeu criar um programa extetno de avaliação do 1º grau que subsidiasse as Secretarias de Estado da Educação com in-

4 DEMO, P. *A Pobre Educação Pobre. Alguns Problemas da Falta de Recursos*. Texto Mimeografado. Brasília, D.F., 1980.

formações 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 eles e o Ministério da Educação um convênio para a realização da pesquisa.

Ao definir a AVALIAÇÃO do RENDIMENTO em ESCOLAS DE 1º GRAU da REDE PÚBLICA, o MEC solicitou a colaboração da FUNDAÇÃO CARLOS CHAGAS para 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, matriculadas no 1º grau, em 19 escolas⁵; a seguir, 8.069 alunos de 62 escolas diluídas em 20 cidades foram avaliados no seu rendimento escolar⁶, posteriormente, 14.868 crianças de 39 cidades, alunos de 157 escolas de 1º grau, foram pesquisadas em uma terceira fase do projeto⁷. Assim, no conjunto, foram avaliados 27.455 alunos de 238 escolas em 69 cidades localizadas nos vários Estados da Federação existentes à época da pesquisa.

1. Objetivos e Critérios de Trabalho

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. Procurou-se uma definição de objetivos que atendessem à pretensão maior do Ministério da Educação: subsidiar as Secretarias de Estado da Educação com informações relativas ao desenvolvimento dos programas de ensino, fornecendo, ao mesmo tempo, aos seus especialistas em currículo elementos para reflexão e, aos professores, dados que complementassem aqueles de que por ventura dispusessem, obtidos em suas próprias avaliações.

A pesquisa não procurou abranger a totalidade de alunos matriculados nem mesmo realizar uma amostra aleatória, aspecto às vezes desejável, mas que, no presente caso, talvez inviabilizasse a realização do projeto, pela dificuldade de acesso a muitas regiões e pela impossibilidade de um número representativo da totalidade do sistema educacional, nas várias cidades.

A investigação foi realizada em escolas indicadas pelas Secretarias de Educação, segundo alguns critérios estabelecidos, e os aplicadores foram contratados em São Paulo, para garantir a confiabilidade dos resultados. Procurou-se abranger diferentes segmentos sócio-econômicos. A escola da rede pública, entretanto, demonstrou ser freqüentada na sua maioria por alunos oriundos das classes menos favorecidas da população.

5 VIANNA, H.M. e GATTI, B.A. (1988) - Avaliação do Rendimento de Alunos de Escolas de 1º grau da Rede Pública: uma aplicação experimental em 10 cidades. *Educação e Seleção*, jan/jun - nº 19. Fundação Carlos Chagas. São Paulo (resumo da pesquisa).

6 VIANNA, H.M. (1989) - Avaliação do Rendimento de Alunos de Escolas de 1º grau da Rede Pública: um estudo em 20 cidades. *Educação e Seleção*, jan/jun nº 19. Fundação Carlos Chagas. São Paulo (resumo da pesquisa).

7 VIANNA, H.M. (1989) - Avaliação do Rendimento de Alunos de Escolas de 1º grau da Rede Pública: um estudo em 30 cidades. *Educação e Seleção*, jul/dez. nº 20. Fundação Carlos Chagas. São Paulo (resumo da pesquisa).

Os alunos que fizeram Matemática não realizaram a prova de Português ou a de Ciências. Ainda que em alguns delineamentos sejam desejáveis estudos de correlação, havendo necessidade, por conseguinte, de que os alunos de uma mesma série realizem duas ou mais provas, não se situou a análise de correlação como um dos objetivos da investigação. Esta enfrentaria sérios problemas se assim o fizesse, em virtude do absentefismo, especialmente em algumas áreas economicamente menos favorecidas, como na verdade ocorreu.

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 consideração do Ministério e Secretarias de Educação, dos técnicos e dos professores e a adoção de medidas que se fizessem necessárias.

Ao longo da pesquisa, a amostra sofreu alterações, especialmente na 1ª série, por influência de fatores externos e que impossibilitaram a sua reconstituição. Greves do professorado, prolongadas e sucessivas, às vezes com a duração de 160 dias, provocaram um esvaziamento em várias turmas. O mesmo fator exigiu que algumas cidades fossem substituídas por outras não afetadas pelo movimento reivindicatório; assim, alterações percentuais de alunos ocorreram especialmente na 1ª e 7ª séries. A pesquisa procurou trabalhar, ao final, com uma amostra não-probabilística por quotas, tendo em mira a representatividade em cada uma das séries.

2. Instrumentos de Medida e os Programas Curriculares

A definição dos programas 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 de prestígio internacional, como é o caso do CECISP, no Brasil, 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 (r_{xx}) e discriminação média (D), 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. As Tabelas 1 e 2 mostram, a título de ilustração, os dados correspondentes à última etapa do estudo realizada em 39 cidades de 15 Estados do

Brasil. Os dados desta etapa repetem o que ocorreu nas demais. A média para acerto de itens nas provas de Português para as séries estudadas fica em torno de 15-16 pontos (em 30 possíveis) e a porcentagem de alunos que não acertam sequer um terço das questões é considerável, sobretudo na primeira série. Em Matemática, os resultados apresentados são muito preocupantes: as médias são baixíssimas e a proporção dos que acertam um terço ou menos de questões é elevada, chegando a 70% nas 5^{as} e 7^{as} séries. Em Ciências, os dados não se mostram melhores, a média de acertos é de 13 pontos, em 30, tanto para a 5^a como para a 7^a série, e quase 40% dos alunos acertam um terço ou menos das questões das provas. Quanto à Redação, nas 5^{as} séries, os resultados são dramáticos: 60% dos alunos tiraram nota inferior a 5,0, numa escala de 1 a 10 pontos. Na sétima, esta proporção melhora, é de 23%, mas a média obtida pelo grupo estudado é de 5,5. Lembremos que a reprovação no sistema educacional até a quinta série é altíssima, sendo que os alunos que estão na sétima série já foram selecionados pelo sistema, onde as reprovações sucessivas imperam.

TABELA 1

RESULTADOS OBTIDOS EM PROVAS DE PORTUGUÊS,
MATEMÁTICA E CIÊNCIAS POR ALUNOS DE
1^a, 3^a, 5^a E 7^a SÉRIES EM 39 CIDADES DE 15 ESTADOS DO BRASIL

DISCIPLINAS	Nº MÁXIMO DE PONTOS	MÉDIAS ARITM.OBTIDAS				PORCENTAGEM DE ACERTOS PARA 1/3 E MENOS DE QUESTÕES			
		1 ^a S	3 ^a S	5 ^a S	7 ^a S	1 ^a S	3 ^a S	5 ^a S	7 ^a S
Português	30	16	17	16	15	30%	17%	14%	17%
Matemática	30	11	12	9	9	57%	55%	79%	79%
Ciências	30	-	-	13	13	-	-	38%	37%

TABELA 2

RESULTADOS OBTIDOS EM UMA REDAÇÃO POR ALUNOS
DE 5^a E 7^a SÉRIES EM 39 CIDADES DE
15 ESTADOS DO BRASIL

Séries	Média Aritmética (notas 0 a 10)	Porcentagem Dos Que Tiraram Notas Menores Do Que 5,0	Porcentagem Dos Que Tiraram Notas Menores Do Que 3,0
5 ^a SÉRIE	4,9	60%	32%
7 ^a SÉRIE	5,5	23%	6%

3. Reflexões sobre esse processo de avaliação

A avaliação necessita considerar aspectos do meio educacional em que se desenvolverá, evitando, assim, 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, como as que foram desenvolvidas por Cronbach⁸. Essas idéias são perfeitamente válidas em função de nosso contexto, mas não constituem, exatamente, um modelo. A adoção de um modelo transplantado de outras culturas pode criar uma situação artificial, sem desdobramentos práticos.

A avaliação tem por objetivo gerar conhecimentos que levem a decisões que tenham conseqüências imediatas na prática educacional. Desse modo, no caso particular do rendimento escolar, por exemplo, a avaliação procura obter dados que possibilitem a tomada de decisões relativas ao desempenho individual e/ou coletivo em face de um determinado programa curricular. A partir de elementos empiricamente coletados, a avaliação procura descrever, da melhor forma possível, o fenômeno considerado, para possibilitar a fundamentação do processo decisório com base em dados da realidade. O trabalho de avaliação exige, conseqüentemente, uma certa criatividade e independe da adoção de modelos desenvolvidos em outros contextos. Assim, a pesquisa em avaliação educacional, inclusive quando restrita ao rendimento escolar, exige definições e procedimentos *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.

A avaliação educacional, segundo a conceituação apresentada, que a situou como um conjunto de operações visando a levantar informações que possibilitem uma tomada de decisões, como usar determinado livro, modificar um currículo, identificar deficiências no processo de aprendizagem, positivar falta de sensibilidade à instrução, pode adotar como amostra diferentes tipos de conjuntos: um pequeno grupo de alunos, uma série completa, alunos de um ou vários tipos de colégios, estudantes de diferentes áreas geográficas etc. A amostragem na avaliação educacional, assim como em outros tipos de investigação na área da educação, sofre limitações decorrentes do tempo, de pessoal e, especialmente, de orçamento. Todos esses fatores devem ser considerados. A questão tempo é, muitas vezes, a mais crucial, pois a tomada de decisões exige uma coleta rápida para ação imediata, a fim de cortar a propagação de maiores danos em decorrência da demora da intervenção no processo educativo. A problemática de pessoal precisa ser considerada quando a natureza do trabalho exige material humano especializado, especialmente na fase de aplicação dos instrumentos. Poucos são os que, efetivamente, se dedicam à avalia-

8 Cronbach, L.J. Course Improvement Through Evaluation. *Teachers College Records*. Columbia University (64), 1963. New York.

ção. A situação atinge contornos dramáticos quando os trabalhos de avaliação envolvem regiões geográficas, às vezes de difícil acesso. A tudo isso acrescentam-se os custos da avaliação, que a tornam inviável em termos orçamentários, apesar de perfeitamente desenhada segundo as técnicas de amostragem. Note-se que recursos como o correio e aplicação pelo professor são impensáveis em nossa realidade pelas implicações na validade de dados, entre outras. O mesmo problema se apresenta no uso de aplicadores locais, como se constatou na avaliação do Edurural aqui relatado, e em outros levantamentos realizados por diferentes instituições e pesquisadores. Isto fez com que, em nosso caso, fossem deslocados aplicadores treinados externos às regiões e Secretarias. Essa situação aumenta os custos e traz problemas diretos para o tamanho da amostra e localização das escolas. O que privilegiar: a validade dos dados ou a representatividade estatística? Seria desejável que todas as áreas geográficas estivessem adequadamente representadas com seus diferentes tipos de escolas, formando estratos nos quais, posteriormente, de forma aleatória, seriam selecionadas as instituições e indivíduos que, finalmente, integrariam a amostra. Nem sempre esse procedimento - que, supostamente, garantiria a representatividade da amostra - possibilita a factibilidade da avaliação em termos operacionais. Ao avaliador cabe decidir entre o ideal, fundamentado na matemática da amostragem, mas irrealizável em termos concretos, e o possível em termos de realidade, que pode levar a inferências, a ilações, a implicações e, a partir dos dados levantados, chegar a *conclusões que oportunizam decisões e modificações da realidade*, objetivo final do processo de avaliação, com razoável confiabilidade.

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 discrepâncias entre as estatísticas educacionais coletadas por diferentes órgãos governamentais, em decorrência da diversidade das metodologias empregadas, são consideráveis, chegando muitas vezes a 40%, como é o caso das estatísticas sobre população em idade escolar fora do sistema educacional. 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 importante, em relação à atual estrutura de ensino em seus vários graus, é criar um sistema de avaliação externa que, de modo sistemático, informe aos responsáveis pelo ensino/educação os problemas da realidade pedagógica que ocorrem na escola e que se refletem nos diferentes níveis de capacitação cognitiva. A identificação de pontos críticos no desempenho escolar dos estudantes deve ser um dos objetivos do diagnóstico de conhecimentos, para que possam ser corrigidos os desvirtuamentos do processo ensino/aprendizagem.

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. As Secretarias de Estado da Educação poderiam colaborar na solução do problema, mas, quase sempre, acham-se impossibilitadas, por não possuírem cópias dos programas desenvolvidos nas escolas. A legislação permite que cada escola estabeleça seus próprios conteúdos, havendo, portanto, grande diversidade de seqüências curriculares em escolas de uma mesma cidade situadas próximas uma das outras. Algumas Secretarias possuem guias curriculares que oferecem uma certa ordenação ao conteúdo em suas linhas gerais, mas essa situação não é a regra comum. O avaliador vê-se compelido a criar um programa mínimo, não muito coerente às vezes, 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 contextos educacionais, para a construção dos chamados testes padronizados. Ajusta-se perfeitamente 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. O livro didático, ainda que elemento colateral de informação, é um instrumento importante para que o avaliador possa caracterizar o que, efetivamente, ocorre no mundo da escola e, especialmente, em sala de aula.

A participação de professores na avaliação educacional, por intermédio da elaboração de questões e de objetivos instrucionais para uso futuro em investigações, constitui motivo de preocupação para alguns especialistas. O problema precisa ser dimensionado levando em consideração o tipo de avaliação e seus objetivos. Se a avaliação é interna, com finalidade formativa ou somativa, ninguém melhor do que o próprio professor para avaliar seus alunos, tendo em vista que foi ele quem estabeleceu objetivos, usou estratégias para conseguir o máximo de aprendizagem e orientou o aluno ao longo do processo instrucional. Se a avaliação é externa, usar professores da própria escola envolvida na investigação seria criar um novo problema, que desvirtuaria os resultados, levantando elementos que comprometeriam a tomada de decisão. As medidas realizadas seriam, sem sombra de dúvida, fidedignas, grau talvez bastante elevado, mas a validade estaria irremediavelmente prejudicada, a não ser que se pudesse organizar um *pool* de professores igualmente capacitados, que gerariam um *pool* de questões de boa qualidade. Posteriormente, esses itens seriam selecionados de forma a constituírem em uma amostra representativa embasada na relação professores/alunos/escolas/currículos/objetivos/metodologias que integram a avaliação externa. A realidade educacional enfrentada pelos avaliadores é bem diversa do ideal, exigindo, assim, a adoção de práticas acauteladoras para o bom êxito do trabalho. A utilização de professores integrados no sistema educacional, independentemente da região ou área geográfica em que exerçam suas atividades, não contribui para desfigurar o processo de avaliação, desde que sua experiência resulte da prática objetiva do ensino, na realidade do dia-a-dia.

O problema do acesso às provas que serão aplicadas nas avaliações externas é conexo ao da elaboração de questões/objetivos pelos professores das instituições

avaliadas. Algumas experiências em avaliação educacional perderam o seu significado, justamente em face desse acesso. Ao tomar conhecimento das áreas de conteúdo objeto da avaliação, o professor, inconscientemente, tende a acentuar esses aspectos, provocando, portanto, um direcionamento favorável a seus alunos, em oposição aos demais sujeitos submetidos à avaliação. A avaliação externa, se não for bem compreendida pelos professores, gera uma certa ansiedade entre eles, em decorrência, possivelmente, do atual caráter punitivo de que, muitas vezes, se reveste a avaliação em nosso contexto. Não temos uma cultura de avaliação, que precisaria ser criada e desenvolvida em seu sentido mais progressista. A avaliação deve ser um instrumento de trabalho para o professor. A avaliação procura obter dados para orientar o professor em seu trabalho, indicando as distorções no processo de aprendizagem. Busca, também, outros elementos que facilitem o processo decisório, não devendo criar uma situação de constrangimento para o professor, a escola e o próprio sistema educacional.

A divulgação das provas usadas em pesquisa ligadas diretamente à avaliação costuma ser advogada por alguns elementos da comunidade educacional, criando uma situação de constrangimento para o avaliador que, a duras penas, construiu seu instrumental. Há uma tradição em nosso contexto social favorável à divulgação de provas e testes, inclusive por intermédio da imprensa, o que precisaria ser melhor discutido. Ao avaliador cabe decidir o momento apropriado para divulgação do instrumental, *no todo ou em parte*. Alguns relatórios reproduzem fragmentos com tipos específicos de questões, outros se limitam à descrição geral do instrumental e, ainda, um terceiro grupo o divulga na sua inteireza, impossibilitando, desse modo, a reaplicação dos mesmos instrumentos em sucessivas avaliações. Ao longo do tempo, com a ampla divulgação dessas provas e a grande rigidez curricular, torna-se difícil a construção de novos instrumentos de medida sem o comprometimento da sua validade. Por outro lado, a disseminação de provas favorece o treinamento para a avaliação, como já vem ocorrendo em outros contextos que, tradicionalmente, usam testes padronizados. Ao avaliador cabe julgar, portanto, após considerar diferentes variáveis, o acesso às provas e a divulgação das mesmas. A prudência recomenda, entretanto, que o sigilo seja mantido, se o objetivo for criar uma tradição relativa à avaliação externa, tornando-a parte integrante da vida escolar.

O uso de provas objetivas ou discursivas costuma ser apresentado como um problema para alguns avaliadores. A questão, em si, é um falso dilema, considerando-se que existe extensa literatura com base empírica, demonstrando, de modo infofismável, 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, 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.

A avaliação, segundo a perspectiva ora desenvolvida, é um processo sistemático de levantamento de dados relativos a um determinado fenômeno, a fim de possibilitar a tomada de decisões com base em julgamentos de valor. A avaliação educacional aplicada à situação específica do rendimento escolar procura, desse modo, identificar as necessidades instrucionais dos alunos, com base em pontos críticos do seu desempenho em provas de escolaridade. Busca coletar elementos para fundamentar a análise da eficiência do sistema de ensino. A avaliação educacional está, basicamente, comprometida com a melhoria da qualidade do ensino e com a compreensão da influência dos professores no desenvolvimento de programas educacionais.

III. CONSIDERAÇÕES FINAIS

Inegavelmente, nas duas 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 real 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, 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 eles mesmo pagaram.

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 ardidado 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.

Ambas as pesquisas 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? Isto significa pensar em uma política consistente, coerente e de longo prazo - que se sustente a despeito da instabilidade e das mudanças governamentais - sobre a formação dos professores, sobre o currículo ou o conhecimento desejável em cada área do

conhecimento, em métodos de ensino que, sabendo contemplar a diversidade, promovam a apropriação do conhecimento acumulado, o desenvolvimento cognitivo e o compromisso com o social.

Quanto ao processo de avaliação externa, 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 modificações impostas ao modelo inicial pelo contexto real em nada contribui para o desenvolvimento teórico-prático da área de avaliação educacional.