

DESENVOLVIMENTO DE UM PROGRAMA DE AVALIAÇÃO DO SISTEMA ESTADUAL DE ENSINO: o exemplo de Minas Gerais

HERALDO MARELIM VIANNA*
ANA LÚCIA ANTUNES**
MARIA ALBA DE SOUZA**

1.0 - Avaliação da Rede Estadual - considerações gerais

A educação em Minas Gerais, a partir de 1991, começou a sofrer transformações que afetaram a sua linha tradicional e conservadora. A oportunidade do momento criou condições para estruturar um sistema descentralizado de ensino, com autonomia administrativa, financeira e pedagógica; em consequência, procedimentos como a eleição de Diretores, com mandato por tempo limitado, e a criação do Colegiado da Esco-

* Do Departamento de Pesquisas Educacionais da Fundação Carlos Chagas, São Paulo

** Da Secretaria Estadual de Educação de Minas Gerais, Superintendência de Desenvolvimento do Ensino, Diretoria de Avaliação do Ensino

la, com a participação efetiva de professores e pais, possibilitaram o surgimento de um novo clima na escola, oferecendo condições propícias à implementação de um projeto de avaliação do sistema estadual de ensino (Vianna, 1992).

1.1 - Experiências anteriores a partir dos anos 20

A área educacional em Minas Gerais possuía relativa experiência no campo da avaliação da aprendizagem, graças à influência de professores europeus na década de 20 e à formação de professores mineiros nos Estados Unidos, cuja ação educacional foi significativa nos anos 50 e 60. A análise do problema, segundo uma perspectiva histórica (Goulart, 1992), deixa bastante clara a ênfase atribuída à avaliação no contexto da sala de aula, para a identificação de questões ligadas à aprendizagem. Após vários decênios nessa mesma linha de atuação, somente no início da década de 90 começou a surgir a idéia de uma avaliação não apenas do aluno, mas do próprio sistema de ensino da escola pública; ou seja, uma avaliação da qualidade da educação. Assim ocorreu em Minas Gerais e, possivelmente, essa mesma linha terá ocorrido em outros Estados, com variações resultantes de condições específicas locais.

2.0 - Qualidade da educação: esboço de um paradigma

O Programa de Avaliação do Sistema Estadual de Ensino de Minas Gerais partiu do pressuposto de que a avaliação da qualidade da educação exige a consideração de algumas variáveis relacionadas ao *contexto* (atuação da escola), *a fatores não diretamente ligados à escola* (status sócioeconômico da família, nível de educação dos pais, recursos educacionais no lar, participação dos pais no processo educacional, atividades educacionais fora da escola, atividades de lazer e sociais, e atitudes dos estudantes), e a *informações ligadas à escola*, por intermédio da coleta de elementos de *entrada* (características do corpo docente, programas escolares e participação dos pais), do *processo* (currículos e práticas instrucionais), e, finalmente, de dados relativos ao *produto*, representado pelo desempenho escolar e a formação de atitudes.

Assim, o Programa de Avaliação da Escola Pública de Minas Gerais partiu do princípio de que não podia limitar-se à verificação do rendimento escolar, por ser este apenas um momento na caracterização da qualidade do ensino (Vianna, 1990), adotando essa orientação metodoló-

gica a partir de 1991, quando a Secretaria de Estado da Educação iniciou um conjunto de medidas com vistas à recuperação do ensino público.

2.1 - Decisão política de avaliar o sistema

A decisão política de avaliar o sistema resultou do conhecimento e da vivência dos seus principais problemas, assim como de uma análise crítica e isenta dos seus vários aspectos. Isso foi feito por intermédio de uma caracterização da realidade da educação e de um compromisso das autoridades governamentais em relação aos alunos, aos pais, aos professores e à escola e a uma definição de prioridades, que no caso de Minas Gerais (Mares Guia Neto, 1992) foram:

1. autonomia da escola;
2. fortalecimento da direção da escola;
3. programa de aperfeiçoamento e capacitação de professores, especialistas e funcionários;
4. avaliação do Sistema Estadual de Educação;
5. integração com os Municípios do Estado.

2.2 - Fundamentos legais: resoluções

A questão da qualidade e de sua avaliação ficou definida na Constituição do Estado de Minas (1989), que, no seu artigo 196 item X, estabeleceu como uma das condições para garantir a qualidade do ensino a avaliação periódica pelo próprio sistema, pelo corpo docente das escolas e pelos responsáveis pelos alunos (Antunes, 1992). Talvez seja este o único documento constitucional de Estado brasileiro que tenha destacado o significado da avaliação no aprimoramento do ensino.

A par dessa imposição legal, fez-se necessário o envolvimento de todo o sistema - diretores, supervisores, professores, alunos, funcionários e pais - no processo de avaliação; assim, antecedendo a cada avaliação, a Secretaria divulga, por intermédio do Diário Oficial do Estado, uma resolução especificando:

- as séries envolvidas na avaliação;
- os objetivos da avaliação em cada uma das áreas de conteúdo;
- a população alvo da avaliação;
- as datas de realização das avaliações;

- os turnos participantes do trabalho e sua definição;
- a competência das escolas na solução de problemas relativos às séries avaliadas;
- os conteúdos básicos sobre os quais versarão as avaliações;
- a documentação que servirá para a seleção dos conteúdos básicos;
- a existência de um questionário específico para os alunos, por série, e seu conteúdo: dados pessoais, familiares, escolares, sociais e opiniões sobre o curso;
- a existência de questionário de atitudes em relação a determinado aspecto, quando for o caso;
- a existência de questionário sobre a escola e seu contexto;
- a competência da Superintendência de Desenvolvimento do Ensino na realização dos estudos avaliativos;
- as competências da Diretoria de Avaliação do Ensino no que se refere à programação da avaliação; à produção dos instrumentos; à realização de treinamentos operacionais; à coordenação do processamento dos resultados e de sua divulgação; o acompanhamento das DREs e das Escolas em todas as fases da avaliação;
- a participação da Diretoria de Normas e Planejamento Curricular no processo de avaliação;
- a competência das Delegacias Regionais de Ensino na supervisão das atividades das escolas;
- a participação das Escolas no processo de avaliação;
- o envolvimento das demais Superintendências no âmbito de suas atribuições;
- a possibilidade de escolas municipais e particulares integrem o processo de avaliação.

2.3 - Objetivos da Avaliação - definições

O Programa de Avaliação procurou definir com precisão seus objetivos, que abrangeram quatro aspectos considerados da maior relevância para a caracterização da qualidade do ensino nas escolas estaduais:

I- conhecer o desempenho dos alunos em aspectos cognitivos dos conteúdos curriculares;

II- levantar dados e informações para servirem de subsídios na tomada de decisões sobre o desenvolvimento do processo de ensino;

III- identificar, nos conteúdos curriculares, pontos críticos que necessitam maior atenção e requerem intervenção imediata e prioritária para a melhoria da aprendizagem;

IV- fornecer, a partir dos dados levantados, subsídios para que os professores possam atuar com maior eficiência na condução do processo de ensino e a Secretaria da Educação possa colaborar para maior eficiência do sistema.

Os quatro objetivos têm servido de orientação para a estruturação do esquema de desenvolvimento de todo o programa de avaliação e permitido que a Secretaria de Educação possa:

1. compor um acervo de dados e informações consistentes das escolas públicas estaduais;
2. selecionar diversos temas que demandem novos estudos avaliativos mais aprofundados;
3. subsidiar grupos de desenvolvimento de recursos humanos e de currículo, indicando aspectos a merecerem atendimento prioritário;
4. oferecer subsídios a órgãos regionais, escolas e prefeituras municipais para a realização de novos estudos de natureza comparativa.

3.0 - Estrutura Funcional do Projeto: integração de níveis

A implementação do Programa de Avaliação do Sistema Estadual, por sua complexidade, envolve um número considerável de órgãos da Secretaria de Educação, além de organismos externos para a realização de tarefas específicas, como, por exemplo, o processamento de dados e tratamento estatístico, que no caso de Minas Gerais tem a colaboração da Companhia de Processamento de Dados de Minas Gerais (PRODÉM-GE), instituição responsável no Estado por atividades relacionadas com a informática.

As Secretarias de Educação nem sempre dispõem de órgãos dedicados exclusivamente à avaliação, sendo essa atividade integrada a outros setores, com funções múltiplas e diversificadas. A Secretaria da Educação de Minas Gerais, na estruturação de seu novo organograma (1992), considerou essa problemática e, assim, na **Subsecretaria de Desenvolvimento Educacional** criou a **Superintendência de Desenvolvimento do Ensino**, que tem como um de seus órgãos a **DIRETORIA DE AVALIAÇÃO DO ENSINO**, responsável pelo gerenciamento global de todas as atividades ligadas ao Programa de Avaliação.

A hierarquização administrativa é de suma importância no sucesso de um empreendimento avaliativo, conforme se constatou em Minas Gerais. Em nível regional, a Secretaria conta com **42 Delegacias de Ensino**, responsáveis, cada uma delas, por um número considerável de escolas dispersas em 756 municípios. Assim, tem havido necessidade de um perfeito entrosamento entre os vários escalões administrativos a fim de que as ESCOLAS, devidamente orientadas e treinadas, possam, por intermédio de um trabalho conjunto de Diretores, Supervisores, Professores, Alunos e do Colegiado da Escola, realizar o trabalho de aplicação dos instrumentos para a coleta dos dados e informações necessárias à avaliação.

4.0 - Planejamento da avaliação: processo decisório

Ao ser definido o Programa, várias decisões foram tomadas em função do contexto educacional brasileiro, sobretudo o de Minas Gerais. As decisões, longamente amadurecidas, podem, em alguns casos, ferir o purismo teórico dos mais apegados à ortodoxia; ou seja, em certos momentos, houve imperiosa necessidade de inovar, criando situações às vezes inexistentes no Brasil e em outros países com experiência em estudos de avaliação.

4.1 - Qual o sistema de avaliação a ser adotado?

A avaliação do rendimento escolar, no Programa de Avaliação do Ensino em Minas Gerais, partiu do princípio de que o **modelo de avaliação** deve ser integrado ao **modelo instrucional**; ou seja, entre ambos deve existir uma **unicidade**, sem a qual não se poderá afirmar que o ensino estaria concretizando seus objetivos, e nem que a avaliação estaria efetivamente constatando a experiência entre o ensino e o desempenho escolar. Considerou-se, também, que o modelo de avaliação, na medida do possível, deveria corresponder à realidade da escola mineira, fugindo, assim, à tentação de adotar um sistema possivelmente eficaz em outros contextos educacionais, com valores e objetivos diversos dos perseguidos pelos professores da escola fundamental em Minas Gerais. Assim, procurou-se um modelo de avaliação menos sofisticado na sua formulação, mas baseado na realidade sócio-educacional da região; contudo, antes de uma decisão final a respeito do instrumental a ser usado, verificou-se como o próprio sistema procedia, a fim de que a avaliação não fosse desvinculada do mundo da escola. A seguir, foram, também, considerados aspectos re-

lativos a modelos de avaliação existentes em outras áreas educacionais, especialmente a norte-americana (Kibler et alii, 1974), confrontando seus diferentes componentes com as características do sistema de avaliação nas escolas mineiras.

4.1.1 - Seria aconselhável usar prova/teste referenciado a critério?

As provas referenciadas a critério são "novas" (!) no domínio da avaliação e começaram a ser empregadas em alguns sistemas educacionais (Estados Unidos, sobretudo) a partir de meados da década de 60. Ainda que não sejam conhecidas pela quase totalidade dos professores brasileiros, examinou-se cada uma das suas características para verificar a possibilidade de aplicação imediata em Minas Gerais.

I- A principal função de uma avaliação por critério é verificar se o estudante domina um critério específico ou atingiu um determinado padrão de desempenho.

A escola brasileira não segue um ensino por objetivos comportamentais e com bastante frequência se observa grande dificuldade por parte de professores e administradores em estabelecer objetivos instrucionais precisos, prevalecendo objetivos gerais. O ensino não é individualizado e a avaliação não tem caráter formativo, salvo exceções, sendo, na verdade, somativa e episódica.

II- Objetivos instrucionais completos são fundamentais na construção de uma avaliação referenciada a critério.

Ainda que a teoria pedagógica tenha sofrido alguma "influência" behaviorista, a prática docente no país não se baseia em objetivos instrucionais explícitos e mensuráveis, como seria desejável. Toda a estrutura do ensino tende a concentrar-se na transmissão pura e simples de "conhecimentos". Ressalte-se, ainda, que muitos professores sentem dificuldades na definição de seus próprios objetivos instrucionais em termos operacionais.

III - O critério relativo a um domínio (*mastery*) deve ser estabelecido com antecedência para uso na avaliação por critério.

Essa não é a prática na realidade brasileira, que, inclusive, não considera, na avaliação, o conceito de "domain" (conjunto de conteúdos, comportamentos, habilidades, destrezas e atitudes em certa área curricular), como ocorre em contextos que adotam instrumentos referenciados a critério.

IV- Os itens, nos testes referenciados a critério, são construídos para mensurar um predeterminado nível de proficiência.

Seria impossível, num sistema que não prioriza o atendimento individual dos alunos, estabelecer um nível predeterminado de capacitação sob risco de invalidar o processo de avaliação, com a fixação de níveis inteiramente irrealistas.

V- A variabilidade é irrelevante e não é condição necessária para uma boa avaliação referenciada a critério

Todo o sistema brasileiro atual fundamenta-se na discriminação, elaborando itens que em média sejam de dificuldade média, a fim de aumentar a variância dos resultados obtidos e promover uma melhor discriminação dos vários níveis de desempenho. É impossível, a curto prazo, modificar uma idéia consolidada na maioria dos professores, para que surja uma nova mentalidade relativa ao processo de ensino, em que a questão da variabilidade não constitua elemento prioritário, conforme seria desejável.

VI- Os resultados, numa avaliação referenciada a critério, apresentam-se num sistema binário: satisfatório x insatisfatório; sucesso x fracasso, etc.

Esta forma de apresentação dos resultados de um teste não diz grande coisa aos professores, que se sentem mais gratificados com as informações expressas em uma escala decimal ou centesimal, ainda que, na realidade, isso pouco diga sobre o desempenho do aluno.

VII- A elaboração de provas referenciadas a critério exige um número elevado de itens para a seleção de questões após a sua pretestagem.

Assim, admitindo-se a seleção de 10 conteúdos básicos e que a cada conteúdo corresponda um total de 5 objetivos, ter-se-ia, ao final, um conjunto de 50 objetivos. O ideal, em termos de testes referenciados a critério, seria a construção de 10 itens por objetivo, o que forçaria a construção de 500 questões para uma prova. A prática revela, entretanto, que são elaborados entre 3 e 5 itens por objetivo. Se fossem elaboradas 4 questões, no exemplo, teríamos que construir 200 itens; entretanto, considerando, ainda, que há uma perda, às vezes elevada de questões, após a pretestagem, é sugerido por alguns que a esse total seja agregado um número de itens correspondente a 100% do definido, que no caso em espécie resultaria em 400 questões a serem pretestadas. O problema para a aplicação do total de itens após a pretestagem poderia obedecer às seguintes estratégias:

1º - aplicar parceladamente os itens, durante vários dias, aos mesmos alunos, perturbando a vida escolar e levando as crianças à estafa;

2º - criar conjuntos de itens, compondo blocos que seriam subtestes do conjunto das questões, e aplicá-los a uma amostra de alunos, segundo a técnica do **matrix sampling**, com todas as suas implicações estatísticas;

3º - fazer uma amostra aleatória estratificada por conteúdo e objetivo, reduzindo o total de questões a uma prova de 30 questões;

4º - proceder, como fazem muitos, a uma seleção dos objetivos mais importantes, 10 ou 12, e selecionar 3 questões por objetivo, tendo, ao final, uma prova de 30 ou 36 questões.

É necessário considerar a viabilidade desses procedimentos e, também, quais os seus custos-benefícios e se não se poderia obter as mesmas informações por intermédio de processos menos estafantes e complicados na sua execução, além de extremamente caros.

4.1.2 - Seria aconselhável usar prova/teste referenciado a normas?

A população escolar, possivelmente, nunca ouviu falar em instrumentos referenciados a normas. Todos os testes são **ad hoc**, elaborados para serem empregados em circunstâncias específicas. Não há, também,

tradição de instrumentos padronizados. A última, e quase exclusiva, experiência com teste padronizado na área educacional ocorreu em 1962 – há mais de trinta anos –, quando o Projeto Fundação Getúlio Vargas/Fundação Ford procurou desenvolver um instrumento semelhante ao **Iowa Basic Skills**; contudo, o empreendimento não teve maiores consequências, em virtude de alterações curriculares nas escolas do Rio de Janeiro.

I- A principal função de uma avaliação referenciada a norma é estabelecer a posição relativa do estudante em um grupo normativo.

Na avaliação por critério o referencial é o indivíduo, enquanto na avaliação por norma o parâmetro é o desempenho do grupo. A avaliação na escola brasileira valoriza o resultado bruto, apresentado em escala de 10 ou 100 pontos, permitindo inferências bem restritas sobre o desempenho grupal, face às limitações dos dados informativos.

II- Objetivos conceituais ou objetivos precisos precisam ser especificados para fins de construção de testes referenciados a normas.

Habitualmente, os testes elaborados em nossas escolas são um conjunto de questões com a finalidade de verificar o conhecimento substantivo, sem nenhuma preocupação com a identificação dos comportamentos, ficando a validade amostral do instrumento comprometida por verificar apenas a dimensão conteúdo.

III- O critério de domínio (*mastery*) geralmente não é especificado em testes referenciados a normas.

As avaliações realizadas nas escolas brasileiras não se preocupam em promover uma avaliação referenciada a critério ou a norma, mas simplesmente constatar o que o estudante "sabe" em uma determinada área. É uma avaliação eminentemente voltada para o conteúdo substantivo.

IV- Os itens nos testes referenciados a normas são construídos para discriminar os estudantes.

Essa é uma preocupação constante nas avaliações nacionais, a ponto de haver uma certa idolatria em relação à curva normal gaussiana, que supostamente traduziria o "sucesso" do processo instrucional.

V- A variabilidade dos escores é desejável como uma forma de possibilitar interpretações significativas.

Os escores brutos, ainda que às vezes variáveis, não permitem maiores interpretações no caso das avaliações em escolas brasileiras, mas a variabilidade é sempre desejável por ser considerada como demonstração de uma "boa" mensuração.

VI- Os resultados em avaliações referenciadas a normas são passíveis de uma transposição para um sistema tradicional de escores.

Outros contextos usam vários tipos de escores: percentis, escores Z e outras formas de escores padronizados, mas, no caso específico das avaliações nacionais, o resultado traduz apenas o número de questões consideradas corretas. As avaliações usam simplesmente escores brutos, sem tratamento estatístico.

4.1.3 - Como proceder? - uma tentativa de solução realista: testes referenciados a objetivos.

A realidade educacional brasileira desconhece os instrumentos referenciados a critério e a normas, salvo algumas tentativas individualizadas movidas por curiosidade intelectual e interesse educacional. Procurou-se, então, construir um sistema de avaliação que atendesse à nossa realidade e oferecesse resultados que refletissem os objetivos específicos do programa de avaliação do sistema de ensino em Minas Gerais.

O passo inicial foi obter uma amostra de provas aplicadas em todo o sistema, a fim de conhecer sua filosofia e sua estrutura. A seguir, foi levantado o conteúdo programático utilizado em todo o sistema, para determinar os elementos comuns em escolas de todas as Delegacias Regionais de Ensino. A partir desses elementos, definiu-se, então, o planejamento da prova, que teve as seguintes características:

1. avaliar o desempenho dos alunos em pontos relevantes do programa, com a finalidade de determinar deficiências na aprendizagem;
2. avaliar, juntamente com a mensuração do desempenho cognitivo, aqueles comportamentos – **conhecimento, compreensão e uso do conhecimento** – que teriam um desempenho comprometido;

3. definir os objetivos em termos de comportamento final observável a ser mensurado por intermédio de questões sobre conteúdos relevantes;

4. sem determinar um ponto fixo, que refletiria "sucesso" ou "fracasso", estabelecer o percentual de acertos por questão, com vistas a identificar possíveis dificuldades no programa escolar;

5. ainda que a variabilidade não seja intencionalmente procurada, construir instrumentos de forma a identificar diferentes níveis de desempenho e, desse modo, fazer comparações;

6. expressar os resultados em escores percentuais de acertos, e definir promédios, apresentando-os em termos de porcentagens para permitir comparações de cada escola com os desempenhos do Município, da Delegacia Regional e do próprio Estado;

7. os aspectos anteriormente apontados visaram a contornar dificuldades que seriam encontradas na aplicação pura e simples de "testes referenciados a critérios" ou de "testes referenciados a normas"; assim,

7.1 deu-se ênfase à seleção dos conteúdos, procurando abranger os aspectos mais representativos da parte substantiva do currículo;

7.2 procurou-se, em cada conteúdo, identificar os objetivos mais relevantes – **conhecimento, compreensão e uso** –, elaborando-se questão para cada um deles;

7.3 evitou-se elaborar um número grande de objetivos e um número igualmente grande de questões (3, 5 ou 10) por objetivo, por ser um processo demorado e cansativo para o aluno;

7.4 expressaram-se os resultados em termos percentuais de acertos por questão/objetivo, permitindo a identificação, portanto, da concretização ou não de um número grande de objetivos relacionados a conteúdos (30 na maioria das provas), como nos testes referenciados a critério;

7.5 visou-se, por outro lado, a ter estatísticas de normas, que dão significado às medidas de critério, para que as escolas, pela análise de seus professores, pudessem se situar no contexto dos vários grupos, a nível municipal, regional (DRE) e estadual.

O modelo, como todo e qualquer padrão referencial, tem, entretanto, suas limitações. Não é uma avaliação inferencial, que permita generalizar para todos os aspectos da totalidade dos vários domínios substantivos, como supostamente ocorreria na avaliação por critério, a qual não deve ser confundida com a avaliação referenciada a objetivo, planejada e

posteriormente executada no presente programa de avaliação do ensino em Minas Gerais.

O Programa preferiu, assim, adotar um sistema de avaliação simples, sem sofisticação tecnicista, mas ajustado à realidade de um contexto cultural em que os professores procuram vencer inúmeras dificuldades, com grande esforço pessoal, a partir de sua própria criatividade.

4.2 - Qual o número de itens em cada instrumento de avaliação?

A pergunta pode ser simples na sua formulação, mas demandou a consideração de vários elementos, tendo em vista a validade de conteúdo (dimensões comportamento/conteúdo) e, subsidiariamente, a fidedignidade dos resultados. A solução do problema levou em consideração o seguinte:

1. o estudante da escola fundamental não tem por hábito fazer provas longas;
2. as provas deveriam ser aplicadas em espaço de tempo limitado;
3. algumas provas teriam várias áreas de conteúdo, como as de Conhecimentos Gerais;
4. os alunos do Ensino Médio deveriam responder a um teste com um número maior de questões;
5. os alunos do Ensino Fundamental responderiam, junto com os itens da prova, questões de informações que variavam de 10 a 30;
6. os alunos do Ensino Médio responderiam, além das provas abrangendo várias áreas de conteúdo, questionários com número variável de questões (30 a 50);
7. uma questão de redação deveria ser respondida por todos que fizessem prova de português;
8. a aplicação dos instrumentos de avaliação não deveria exceder a um dia escolar para cada série.

A partir desses pressupostos, as várias equipes de avaliação, com a participação obrigatória de professores com experiência em sala de aula, identificaram 10 tópicos importantes em cada uma das áreas de conteúdo, elaborando, em princípio, três objetivos por conteúdo e a partir deles construíram uma questão por objetivo. Assim, as provas, ainda que variando conforme o número de conteúdos, seguiram a orientação dos testes referenciados a objetivos, segundo a conceituação de Hambleton et alii (1978) e, ao adotarem a relação ob-

jetivo/questão, procederam conforme o padrão norte-americano no final dos anos 60 e início da década de 70, de acordo com a observação de Milmann (1974). A avaliação planejada procurou, assim, verificar a congruência entre objetivos e desempenhos, de acordo com a colocação de Tyler (1942).

4.3 - Quais as séries a avaliar?

A análise de ocorrências ao longo da seriação escolar e de dados empíricos recomendou a inclusão no processo de avaliação da 3ª, 5ª e 8ª séries do Ensino Fundamental, da 2ª série do Ensino Médio, assim como da 3ª e/ou 4ª série da Habilitação Magistério das Quatro Primeiras Séries do Ensino Fundamental. A avaliação de alunos do Ensino Fundamental, no início da 3ª série, tem permitido verificar em que medida o Ciclo Básico de Alfabetização (CBA) estaria conseguindo atingir com êxito seus objetivos. A 5ª série representa uma mudança radical na vida do aluno, que deixa a primeira fase do Ensino Fundamental, ingressando em uma sala de aula onde o ensino é diversificado em seu conteúdo, nas metodologias de ensino e na pluralidade dos professores. A 8ª série por ser terminal de um nível de ensino que não tem continuidade para uma parcela expressiva do alunado que não prossegue seus estudos em nível de Ensino Médio. A 2ª série do Ensino Médio, que ainda não se acha inteiramente contaminada pela influência dos cursos preparatórios para o acesso ao ensino de 3º grau, foi selecionada para integrar o projeto, fornecendo subsídios sobre o sistema em algumas importantes áreas de conteúdo nesse nível: Português, Matemática, Física, Química e Biologia, além de informações sobre o funcionamento da escola. A inclusão da 3ª/4ª série da Habilitação Magistério resultou de razões óbvias: conhecer aspectos da formação dos futuros professores que irão atuar nas quatro primeiras séries do Ensino Fundamental.

4.4 - Quais as áreas de conteúdo a avaliar?

Houve consenso que para todas as séries pesquisadas haveria uma prova de PORTUGUÊS, com vistas a avaliar o entendimento e o uso da língua nacional, e uma prova de REDAÇÃO, que demonstrasse em que medida os alunos seriam capazes de expressar o pensamento com clareza e correção gramatical, além, naturalmente, de uma

prova de MATEMÁTICA. A Tabela 1 especifica os conteúdos e o número de questões das várias provas/têstes usados no Programa de Avaliação ora em discussão.

Tabela 1- Provas e demais instrumentos aplicados no Programa de Avaliação do Sistema Estadual de Ensino em Minas Gerais, 1991-93.

ÁREAS DE CONTEÚDO	NÚMERO DE QUESTÕES POR SÉRIE				
	3 ^a	5 ^a	8 ^{as} *	2 ^a /2 ^o grau*	3 ^a /4 ^a Hab. Mag.
1. Português + Redação	30	30	30	30	-
2. Matemática	30	30	30	-	-
3. Ciências	30	-	-	-	-
4. Conhecimentos Gerais	-	-	-	-	-
4.1 História	-	15	15	-	-
4.2 Geografia	-	15	15	-	-
4.3 Ciências	-	20	20	-	-
5. Estudos Sociais	-	-	-	-	-
5.1 História	-	-	-	20	-
5.2 Geografia	-	-	-	20	-
6. Matemática/Física	-	-	-	20+20	-
7. Química/Biologia	-	-	-	20+20	-
8. Conteúdos Básicos para Magistério	-	-	-	-	50
9. Questionário para aluno	10	30	50	-	50
10. Questionário de Atitude	-	-	-	34	-

* Provas equivalentes, mas não iguais, para os turnos diurno e noturno.

Todas as escolas responderam a um questionário de informações com 50 questões.

4.5 - Quais os alunos a avaliar: amostra ou população?

Ainda que a avaliação pudesse ser realizada a partir de uma amostra, obtendo resultados igualmente confiáveis, graças à precisão das técnicas estatísticas, e com menores custos, optou-se por uma avaliação censitária, envolvendo, conseqüentemente, toda a população de alunos das séries escolares a serem avaliadas, com a intenção deliberada de:

1º - criar uma cultura da avaliação em todo o sistema, considerando que a prática avaliativa é bastante episódica na escola. A exclusão de alunos e escolas, possivelmente, não geraria um maior comprometimento dos professores;

2º - influenciar, através dos instrumentos utilizados (provas), sobre a qualidade das práticas avaliativas que cada escola usa, bastante precárias em muitas escolas de diferentes regiões;

3º - possibilitar que todas as escolas, seus professores e seus alunos realizem uma auto-avaliação, comparando os resultados da sua escola com os de outros municípios, da região e do Estado;

4º - levar os professores a uma atividade cooperativa dentro da própria escola e com membros integrantes do Colegiado da Escola, que, segundo as normas estabelecidas, também participam da avaliação, atuando, todos, no controle do ensino de qualidade e elaborando relatórios da avaliação;

5º - identificar as escolas que, face às deficiências constatadas, necessitam de assistência que lhes garanta melhores condições materiais e de aprimoramento de seus recursos humanos;

6º - positivar, no conjunto do sistema, as escolas de melhor desempenho a fim de que seus professores e demais participantes do processo educativo possam ter o seu mérito reconhecido; e, finalmente,

7º - permitir que as opções avaliativas propostas no Programa de Avaliação contribuam, efetivamente, para que todas as escolas alcancem sua autonomia, sobretudo pedagógica.

A partir desses objetivos, o quadro numérico do programa de avaliação ficou definido da seguinte forma:

Quadro 1 - Número de DREs, Municípios, Escolas e Alunos participantes do Programa de Avaliação do Sistema Estadual de Ensino em Minas Gerais. 1991-93

	3ª série	5ª série*	8ª série	2º/2º grau*	3º/4º Hab. Mag.*
Delegacia Regional de Ensino	42	42	42	42	42
Municípios	723	720	721	520	361
Escolas	5.645	2.353	2.173	813	434
Alunos matriculados**	325.000	407.443	124.339	69.163	6.763
Alunos avaliados	312.811	-	99.810	-	-

*Avaliação a ser realizada em dezembro de 1993.

**Matrículas no início do ano letivo

Embora todos os alunos da série escolar selecionada sejam submetidos à avaliação, cada um deles responderá apenas uma das três (ou quatro) provas constantes da avaliação. Para determinar que prova fará cada aluno, procede-se a processo aleatório simples de enturmação, específico para a ocasião da aplicação das provas, e diferente do adotado pela escola durante o ano letivo.

Justificam a adoção dessa alternativa censitária parcial e a decisão por uma enturmação específica em que cada aluno tenha igual probabilidade de responder a qualquer das provas:

- a intenção de desvincular, nas primeiras avaliações, os resultados do rendimento dos alunos de uma escola do desempenho profissional de seus professores;

- a necessidade de se evitar que os resultados do desempenho dos alunos nas provas fossem influenciados positiva ou negativamente o que fatalmente ocorreria caso se adotasse a enturmação da escola geralmente processada pelo chamado "critério de homogeneidade", em que o agrupamento dos alunos em turmas é feito por níveis similares de aproveitamento escolar;

- o atendimento à deliberação de se ocupar somente um dia do calendário escolar em que seriam aplicadas simultaneamente as provas de todas as disciplinas incluídas na avaliação de cada série escolar.

4.6 - Quais as etapas a seguir no processo?

A fim de sistematizar as ações, definiram-se as seguintes etapas comuns a todas as experiências de avaliação:

- I- Elaboração de projeto de Avaliação;
- II- Aprovação técnica e administrativa do Projeto;
- III- Edição de ato legal definindo objeto das experiências e divulgando orientações pertinentes;
- IV- Divulgação da avaliação por intermédio do órgão oficial e da mídia em geral;
- V- Reunião com a comunidade educacional e mídia;
- VI- Construção dos instrumentos de medida;
- VI^a - elaboração de um fluxograma para os trabalhos de construção;
- VII- Preparo do apoio logístico;
- VII^a - elaboração de um fluxograma para a logística;
- VIII- Treinamento de pessoal dos órgãos envolvidos: DREs e Escolas;

- IX- Aplicação dos instrumentos nas escolas;
- X- Preparo, nas escolas, do material para remessa às DREs;
- XI- Remessa do material aplicado à SEE pelas DREs;
- XII- Controle na SEE do material das DREs;
- XIII- Processamento e computação dos resultados;
- XIV- Divulgação dos primeiros resultados, interna e externamente, em nível central e regional;
- XV- Treinamento de DREs e Escolas para elaboração de relatórios;
- XVI- Elaboração de relatórios por Escola;
- XVII- Resumo dos relatórios por Município e DRE;
- XVIII- Síntese global dos relatórios das DREs no Estado;
- XIX- Divulgação e disseminação dos Resultados Finais pela SEE/DRE e Escolas;
- XX- Utilização dos resultados pelas Escolas, DREs e Secretaria da Educação.

4.7- Quais os períodos de realização das avaliações?

A avaliação de um sistema não se realiza **ex-abrupto**, demanda tempo para planejar, preparar, aplicar, analisar e divulgar as informações. Procurou-se, assim, no presente programa, organizar as várias etapas de forma a que houvesse possibilidade de garantir a sua execução e continuidade, independentemente de possíveis alterações administrativas.

Assim, ficou decidido que:

1º - nos **anos pares**, seria feita a avaliação do Ciclo Básico de Alfabetização e da 5ª série; e,

2º - nos **anos ímpares**, far-se-ia a avaliação da 8ª série do Ensino fundamental, da 2ª série do Ensino Médio, e da Habilitação Magistério.

Nos anos pares, far-se-ão, também, os pré-testes das avaliações que serão realizadas nos anos ímpares e vice-versa.

O rendimento dos alunos das séries selecionadas será avaliado, portanto, a cada dois anos, permitindo que, no interregno entre uma e outra testagem, haja tempo suficiente para que sejam implementadas medidas de intervenção para sanar deficiências de ensino diagnosticadas e para que melhorias, porventura ocorridas, possam se manifestar nos resultados das experiências subseqüentes.

4.8 - Qual o envolvimento com órgãos externos?

A complexidade do processo de avaliação do sistema de ensino exige a realização de serviços técnicos e especializados que as Secretarias de Educação nem sempre possuem, ou, que, face a dispositivos legais, foram obrigadas a terceirizar. Assim, no caso específico de Minas Gerais, a impressão de centenas de milhares de provas, manuais de instrução e material de divulgação é feita, por exigência legal, por instituição do Estado – a **Imprensa Oficial** –; da mesma forma, atividades como cadastramento de escolas, divisão de alunado em turmas e por disciplina, identificação de folhas de respostas e outros materiais, assim como processamento de todos os dados coletados e emissão das estatísticas uni e bivariadas, e que geram uma massa substancial de informações, são realizadas pela PRODEMGE – **Companhia de Processamento de Dados do Estado de Minas Gerais** –, empresa do Estado incumbida de concretizar ações na área de informática.

4.9 - Qual a assessoria técnica a utilizar?

Ainda que as Secretarias de Educação disponham de recursos humanos qualificados, muitas vezes a sobrecarga de trabalho exige que pessoas estranhas aos serviços públicos sejam contratadas para atividades de assessoria. Um programa de avaliação como o descrito requer muitas vezes o concurso de pessoal técnico em diversas áreas (Vianna, 1992): Medidas Educacionais, Avaliação, Estatística, Pesquisa e Informática para as seguintes atividades:

I- ÁREA DE MEDIDAS

Atividades:

- 1) Desenvolver programas de medidas na área cognitiva e não-cognitiva;
- 2) Orientar estudos com base na literatura técnica sobre medidas;
- 3) Definir com precisão os objetivos e a natureza dos instrumentos de medida;
- 4) Analisar, criticamente, propostas de programas de medidas;
- 5) Desenvolver planejamentos adequados à realização de medidas necessárias aos objetivos da avaliação;

- 6) Estabelecer critérios para desenvolver instrumentos de medidas;
- 7) Estruturar plano de análise dos resultados apresentados pelos instrumentos.

II- ÁREA DE AVALIAÇÃO

Atividades:

- 1) Formular hipóteses, questões ou problemas: a serem verificados na avaliação;
- 2) Apresentar e definir os objetivos do programa de avaliação;
- 3) Especificar a natureza dos dados a serem levantados no programa de avaliação;
- 4) Desenvolver projetos de avaliação com vistas a responder às questões ou problemas formulados;
- 5) Orientar o planejamento da avaliação e controlar fatores que possam comprometer sua validade;
- 6) Transformar objetivos gerais em operacionais para fins de avaliação;
- 7) Analisar, criticamente, propostas de planos de avaliação educacional;
- 8) Formular plano de análise dos dados da avaliação para fins de elaboração de relatórios;
- 9) Orientar estudos com base na literatura técnica sobre avaliação educacional.

III - ÁREA DE ESTATÍSTICA

Atividades:

- 1) Definir programa de quantificação dos resultados da avaliação;
- 2) Selecionar, se for o caso, amostras da população para a qual os dados da avaliação serão generalizados;
- 3) Identificar a natureza das variáveis a serem quantificadas;
- 4) Estabelecer procedimentos para análise da validade e da fidedignidade dos instrumentos de medidas;
- 5) Definir a metodologia para o levantamento dos dados;
- 6) Selecionar e aplicar técnicas estatísticas adequadas à análise dos dados:

- 7) Orientar, juntamente com o grupo de processamento, o tratamento estatístico dos dados;
- 8) Orientar estudos com base na literatura técnica sobre estatística educacional.

IV - ÁREA DE PESQUISA

Atividades:

- 1) Desenvolver e fundamentar metodologias de pesquisa;
- 2) Analisar e discutir projetos de pesquisa elaborados por terceiros;
- 3) Apresentar modelos e abordagens adequados à pesquisa de avaliação;
- 4) Formular e discutir possíveis planos de pesquisa;
- 5) Administrar o desenvolvimento de programas de pesquisa, identificando e corrigindo possíveis desvios do planejamento;
- 6) Estabelecer procedimentos para a análise dos dados levantados;
- 7) Interpretar e estabelecer conclusões fundamentadas para a elaboração de relatórios;
- 8) Elaborar e discutir as implicações dos resultados das pesquisas;
- 9) Orientar estudos com base na literatura técnica sobre pesquisa educacional.

V - ÁREA DE INFORMÁTICA

Atividades:

- 1) Atuar de forma articulada com o pessoal de medidas, avaliação, pesquisa e estatística;
- 2) Elaborar programas de "software" que atendam às necessidades dos programas de avaliação;
- 3) Desenvolver pacotes estatísticos para tratamento dos dados;
- 4) Utilizar pacotes estatísticos, sobretudo os programas SAS, SPSS, STATGRAPH, entre outros;
- 5) Orientar, no Centro de Informática, as várias fases do processamento dos dados até a emissão do relatório final;
- 6) Analisar e discutir planos de processamento propostos pela área de medidas e avaliação;

- 7) Orientar a formação de competência técnica do pessoal da avaliação na operação de microcomputadores.

5.0 - Objetivos e especificações dos instrumentos

5.1. Provas: objetivas e não objetivas

As provas do Ciclo Básico de Alfabetização, no início da 3ª série, são de natureza semi-objetiva e exigem que as crianças demonstrem o seu grau de alfabetização. As provas das demais séries têm caráter objetivo e todas procuram verificar, basicamente, três capacidades: 1 - **conhecimento**, 2 - **compreensão** e 3 - **uso do conhecimento**. Ao lado das provas de conhecimento, uma outra de redação é exigida com vistas a estabelecer em que medida os alunos possuem capacidade de expressão escrita.

A elaboração de todas as provas segue o seguinte fluxo:

- 1 - Definição de objetivos instrucionais
- 2 - Identificação de áreas de conteúdo
- 3 - Fixação de padrões mínimos de desempenho
- 4 - Especificação do instrumento: teste/prova
- 5 - Elaboração dos itens
- 6 - Crítica interna dos itens
- 7 - Crítica independente (externa) dos itens
- 8 - Revisão dos itens/questões
- 9 - Montagem do teste/prova
- 10 - Revisão editorial do teste/prova
- 11 - Impressão do teste/prova
- 12 - Aplicação dos instrumentos
- 13 - Correção dos testes/provas
- 14 - Análise dos resultados: escores/notas
- 15 - Avaliação do rendimento: relatório

A partir de 1994, haverá pré-testagem dos instrumentos; ou seja, um número maior de itens será elaborado e aplicado a uma amostra da população de alunos, a fim de determinar as características psicométricas das provas, eliminar os itens possivelmente defeituosos e aprimorar a qualidade das questões, controlando, também, o seu grau de dificuldade e o seu poder discriminativo, para maior fidedignidade dos resultados na versão final do teste.

5.2. - Questionário do aluno

O questionário para as crianças da 3ª série incide sobre alguns poucos assuntos básicos, contendo apenas 10 perguntas de resposta imediata. Para os alunos das demais séries, entretanto, procura-se obter um maior número de informações, versando, basicamente, sobre:

- permanência na série (novato/reprovado);
- sexo e idade;
- tamanho da família (irmãos) e nível de instrução dos pais;
- exercício de atividade remunerada: início, horas de trabalho, renda mensal, e descrição ocupacional do pai;
- hábito de ver TV, ler, número de livros em casa;
- horas dedicadas às lições de casa;
- escola que frequenta (pública/privada) e posição em relação ao futuro, inclusive tipo de curso a realizar;
- importância das disciplinas que estuda, atividades ligadas ao seu ensino, dificuldades, leitura, atividades práticas, número de provas que realiza, atitude em relação à disciplina, dificuldades que encontra, importância da disciplina na vida profissional e na vida diária: – em relação a cada uma das áreas de conteúdo avaliadas.

As frequências marginais e o seu cruzamento com o desempenho escolar fornecem subsídio para a compreensão do desempenho dos alunos em cada uma das áreas cognitivas avaliadas.

5.3. - Questionário da escola

Ainda dentro da linha estabelecida inicialmente com relação à qualidade da educação, procura-se, por intermédio do questionário de escola, conhecer, entre outros aspectos:

- o número de alunos por turma e os critérios para a formação dessas turmas;
- a qualificação dos professores, os critérios de indicação para as turmas e sua distribuição por turma;
- o absentéismo dos professores e sua influência no processo ensino-aprendizagem;
- a questão do fracasso escolar, sua taxa, maior concentração por disciplina;
- os procedimentos de avaliação, objetivos e meios utilizados;
- recursos didáticos e pedagógicos utilizados em sala;

- participação dos pais e da comunidade na escola - sua integração;
- mudanças no corpo docente; uso da biblioteca;
- ação da Secretaria de Educação e das Delegacias Regionais de Ensino;
- programas de ensino e livro usado; prática da redação e maiores dificuldades encontradas pelos alunos;
- tipo de aula ministrado e a reação dos alunos.

Os itens incidem especialmente sobre as áreas de conteúdo objeto da avaliação, o que contribui sobretudo para que se compreenda a relação escola/desempenho dos estudantes nas várias avaliações.

5.4. - Construção de outros instrumentos

Atendendo às circunstâncias e tendo em vista a avaliação de futuros professores, assim como a importância da ciência na moderna sociedade tecnológica, foram incluídos no programa de avaliação: a) para os alunos da Habilitação Magistério, um questionário especial sobre sua formação e expectativas profissionais; b) para os estudantes da 2ª série do 2º grau um Questionário de Atitude em relação à Ciência.

5.4.1 - Questionário para Alunos da Habilitação Magistério

Além de características gerais: sexo, idade, estado civil, número de filhos, com quem mora, número de membros da família, exercício de atividade remunerada, renda familiar, ocupação paterna/materna e remuneração, procurou-se, também, incluir um levantamento de dados mais específicos sobre:

- conhecimento a respeito do curso de magistério;
- motivos que levaram à escolha desse curso;
- reação dos pais à escolha do curso de magistério;
- atividades após a conclusão do curso de magistério;
- curso de 3º grau que pretende realizar após a conclusão do magistério;
- grau de satisfação em relação ao magistério;
- opinião sobre o curso de magistério que realiza;
- proposta de modificações para o curso de magistério;

As questões foram fechadas, exigindo, assim, em muitos casos, a apresentação de um grande número de respostas possíveis, cuja frequên-

cia será posteriormente estabelecida pela determinação de diferentes vetores. Procurou-se, desse modo, evitar a multiplicidade de respostas abertas, difíceis de categorização e tabulação.

5.4.2. - Questionário de Atitudes em relação à Ciência

Foi desenvolvido um questionário de atitude a partir de um conjunto de 54 questões estímulo, que, depois de pretestadas em uma amostra de 500 alunos de Minas Gerais e analisadas as estatísticas da pretestagem, ficaram reduzidas a 34, distribuídas entre 5 possíveis relações com a Ciência:

- Ciência e Experiência,
- Ciência como Profissão,
- Ciência e Sociedade,
- Ciência como Processo,
- Ciência na Escola.

A escala de atitude permitirá calcular um escore, variando entre 34 a 170, sendo considerados como tendo uma atitude altamente favorável à Ciência os que se situarem a partir do escore 130, correspondente, em princípio, ao terceiro quartil na escala de distribuições dos escores. A escala permite uma correção objetiva, a partir de respostas em folha ótica.

5.4.3. Revisão dos questionários

A partir de 1994, após o término do 1º ciclo de avaliações (1992-93), e antes do início do 2º ciclo, os questionários serão revistos para modificação das perguntas que não funcionaram adequadamente e inclusão de outras que procuram conhecer em que medida houve, na escola, mudanças no processo de ensino-aprendizagem.

6.0 - Problemas de conteúdo e revisão

6.1 - Definição de áreas

A definição de áreas para um programa de avaliação do rendimento do ensino não constitui realmente um problema já que se achavam definidas por dispositivos legais, o mesmo não ocorrendo em relação aos conteúdos, que variavam grandemente, apesar da existência da chamada "proposta curricular". O problema está, na verdade, no conteúdo progra-

mático dos currículos, que variavam de escola para escola e até mesmo de professor para professor em uma mesma escola, gerando, assim, uma multiplicidade de situações, que, muitas vezes, são ampliadas pela introdução de mais um elemento: – o livro didático, conforme demonstrou a experiência.

Superada essa situação, configurou-se uma outra ligada ao conteúdo, sobre a qual nem sempre existiu consenso entre administradores, avaliadores e professores: – a natureza dos vários tópicos dos conteúdos e sua importância relativa. Tudo isso é de grande relevância para a determinação da validade de conteúdo dos instrumentos, e qualquer que seja a decisão, será objeto de críticas, pois todo processo é impregnado de subjetivismo, havendo variação de grupo para grupo e de professor para professor. É um assunto que exige reflexão e consideração sobre as possíveis influências na validade dos instrumentos. Minas Gerais optou, de um modo geral, por 10 sub-áreas de conteúdo, formando sub-testes. Em algumas situações, contudo, o número de conjuntos pode variar em função dos objetivos definidos.

6.2 - Critérios de êxito

O quadro apresenta uma maior complexidade quando se deve decidir sobre os critérios de êxito que indicariam o sucesso do aluno ou o seu fracasso na questão. Habitualmente, o professor não estabelece critérios de êxito ou fracasso, limitando-se a uma nota, que, ao final, se integrará a uma contabilidade duvidosa e da qual resultará uma aprovação ou reprovação. Ao ser solicitado a posicionar-se para estabelecer um critério relativo a um aluno que se situa na **borderline**, o professor tem dificuldade de quantificar essa situação, porque sempre se reporta aos extremos, exigindo um máximo de desempenho e não um marco que defina uma linha divisória entre o aceitável e o não aceitável. A tendência extrema reflete quase sempre uma avaliação subconscientemente punitiva, que parece ser a regra geral aceita pelo sistema. **Adequar o critério de êxito ao real desempenho do aluno vem se constituindo, assim, em uma questão extremamente crítica no processo de desenvolvimento do programa de avaliação que ora se desenvolve em Minas Gerais.**

6.3 - A questão da revisão

A revisão dos instrumentos foi feita em três níveis: técnico (psicométrico), substantivo (conteúdo) e editorial. O problema concentrou-se

na revisão substantiva, quando o conteúdo foi verificado na sua correção científica, na adequação ao nível de escolaridade dos avaliados e na identidade entre objetivos fixados e questões propostas.

Encontrar pessoas que possuam as qualificações e conhecimentos necessários e que se disponham a fazer o trabalho de revisão tem sido difícil; contudo, o trabalho de revisão tem sido feito por ser indispensável à garantia da qualidade do instrumental desenvolvido. Sempre que possível e oportuno tem sido utilizada a revisão em pequenos grupos de que participam especialistas com formação e experiências variadas, todas ligadas ao ensino, extraíndo-se deles, muitas vezes, decisões de consenso.

7.0 - Treinamento de pessoal

O programa de avaliação baseia-se na interação de todos os níveis administrativos do sistema, partindo da ação da Secretaria de Educação, passando pela atividade orientadora das Delegacias Regionais de Ensino para concretizar-se, ao nível da escola, nas ações dos professores, os quais trabalham as informações iniciais que constituem a base da documentação posteriormente consolidada por municípios e regiões, dando origem a 42 relatórios das DREs, os quais, mais tarde, são consolidados pela Diretoria de Avaliação do Ensino para disseminação no sistema, nas comunidades e no Estado.

A fim de que tudo isso ocorra são necessárias várias fases intermediárias de treinamento de recursos humanos nos níveis regional (DREs), municipal (inspetores e supervisores) e na escola (diretores, supervisores e professores), sendo esta última fase a que mais exige dos treinadores e treinandos, por condicionar todo o sucesso do projeto, uma vez que a essência da avaliação, a coleta de dados, se processa na escola quando da aplicação dos instrumentos (provas e questionários) e da apuração dos resultados.

O treinamento de pessoal constitui, pois, uma fase crítica, necessitando de planejamento detalhado e de uma execução bastante cuidadosa, sem esquecer os mínimos detalhes, a fim de evitar inovações passíveis de ocorrer em um processo que, no caso específico de Minas Gerais, chegou a empregar mais de 20.000 professores em uma só avaliação, o que poderia vir a comprometer o êxito do desenvolvimento do projeto.

8.0 - Elaboração de manuais

A complexidade dos processos de avaliação demandou a elaboração de um conjunto de manuais destinados à coordenação, supervisão municipal, aplicação das provas, apuração dos resultados, controle de qualidade e elaboração dos relatórios, que foram usados, em diferentes momentos, para treinamento dos vários responsáveis pela implementação do projeto de avaliação.

A elaboração de manuais para a coordenação geral e a supervisão, em nível municipal e regional, provê instruções detalhadas sobre as várias equipes e suas competências. Além disso, os manuais apresentam orientações sobre:

- seleção e preparo de recursos humanos,
- divulgação da avaliação,
- realização da avaliação,
- remessa do material,
- elaboração de relatórios,
- divulgação dos resultados,
- utilização dos resultados,
- gerenciamento de recursos financeiros,
- cronograma de atividades.

O manual do aplicador da prova mereceu um tratamento especial. Isto porque a aplicação dos instrumentos de forma adequada constitui a principal fase do fluxo da avaliação para garantia da validade e da fidedignidade dos resultados. Ainda que óbvia, quase tautológica a observação, muitas vezes cuida-se da elaboração do instrumental, com a preocupação de seguir toda a tecnologia recomendada pela docimologia, esquecendo-se, porém, de tomar cuidados com a aplicação e, assim, coletam-se dados de forma imprecisa, que, por sua vez, levam a conclusões não revestidas de credibilidade. O manual contém importante informações que devem ser dadas aos professores responsáveis pelo trabalho: além de explicar o objetivo da avaliação e a filosofia de cada uma das provas e sua composição, detalha as várias atribuições, nos diferentes momentos que antecedem à aplicação da prova e os comportamentos exigidos dos professores no ato de aplicação dos instrumentos e as várias atribuições a serem desenvolvidas após a prova. No caso da apuração manual, o que se verifica em relação às provas do Ciclo Básico de Alfabetização, instruções especiais são elaboradas sobre a organização de equipes, os critérios de ava-

lição e a transposição de resultados para folhas óticas a serem usadas na fase de processamento.

Importante no Programa de Avaliação foi a introdução do **Controlador da Qualidade** de todo o processo de aplicação, função esta exercida por pais ou responsáveis pelos alunos, membros integrantes do Colegiado da Escola e indicados pelo próprio Colegiado para o exercício dessa função. As atribuições dos Controladores da Qualidade da Avaliação, em diferentes momentos, inclusive na fase de apuração, permitiram estabelecer um estreito contacto entre a comunidade e a escola.

Um manual especial foi redigido para garantir a homogeneidade dos relatórios elaborados pelos professores, nas escolas. Sem procurar coibir as individualidades, o manual procurou chamar a atenção sobre:

- os dados do questionário,
- os dados do desempenho geral (média e variabilidade),
- os dados do desempenho geral (distribuição dos escores),
- os dados do desempenho em redação,
- a abordagem dos pontos críticos,
- as conclusões a partir dos dados coletados.

Os relatórios, por área de conteúdo, são redigidos com a participação do Diretor, Supervisor e/ou Coordenador, professores e representantes dos pais. A partir desses relatórios parciais, por escola, são gerados novos relatórios por municípios, DREs e, finalmente, há uma síntese em um relatório geral elaborado pela Diretoria de Avaliação do Ensino, da Secretaria de Estado da Educação.

9.0 - Estatísticas para análise

O Programa de Avaliação previu diferentes grupos de estatísticas, considerando seus usuários:

- 1 - Escolas e Delegacias Regionais de Ensino (DREs),
- 2 - Secretaria Estadual de Educação,

9.1 - Estatísticas para as escolas e DREs:

- porcentagem de acertos para cada uma das questões por prova, por escola, por município, por Delegacia e total do Estado;
- média e desvio padrão/coeficiente de variabilidade por prova, por escola, por município, por Delegacia e total do Estado;

- porcentagem de respostas às informações do **questionário dos alunos** por escola, por município, por Delegacia e total do Estado;
- porcentagem de respostas ao **questionário da escola** por município, por Delegacia e total no Estado.

9.2 - Estatísticas para a Secretaria Estadual de Educação:

- estatísticas gerais (média, mediana, moda, variância, desvio padrão, escores máximos e mínimos, amplitude, quartil inferior e quartil superior, amplitude interquartilica, assimetria, desvio padrão da assimetria, curtose e desvio padrão da curtose, r-bisserial por item e Alfa de Cronbach por prova);
- estatísticas bivariadas, apresentando relação entre variáveis do questionário do aluno e rendimento, e entre o questionário da escola e o desempenho escolar;
- estatísticas multivariadas poderão ser usadas numa fase posterior, se for estudada a influência de diferentes variáveis sobre o desempenho escolar;
- outras estatísticas foram usadas na pretestagem do Questionário de Atividades (2º série do Ensino Médio), envolvendo análise de variância e testes de significância, entre outras.

10.0 - Disseminação dos resultados

Foi definida uma política de disseminação dos resultados a fim de que todo o sistema pudesse entrar em contato com as informações levantadas e decidir com base em elementos comprobatórios. A divulgação dos resultados do programa de avaliação foi iniciada logo após a sua primeira fase, com os dados do Ciclo Básico de Alfabetização.

A disseminação dos resultados da avaliação tem sido feita mediante publicação de relatórios e outros modos de apresentação adequados aos diferentes tipos de usuários, que são definidos quando se responde à questão: A quem interessa os resultados da avaliação? Professores? Pais? Educadores? Políticos? Empresários? Tais categorias são reunidas, para efeito de divulgação, em usuários diretos e indiretos, internos e externos, gerais e específicos, uma vez que a cada grupo correspondem tipos e modos de apresentação diferenciados e adaptados a seu interesse e especificidade da utilização que pretende dar à informação produzida.

Ressaltam-se como usuários diretos e internos em termos institucionais, a Escola (Direção + Professores + Colegiado), a Delegacia Regional de Ensino e a Secretaria de Educação, considerados como grupos a que estão afetas, cada um em seu nível, a administração da educação e a prática da ação pedagógica propriamente dita. Para esses grupos, além do acesso a todas as estatísticas produzidas, são endereçados relatórios detalhados que analisam o que os resultados mostram, chegando, conforme o caso, a comentarem o que os resultados sugerem em termos de intervenção direcionada para melhoria dos serviços educacionais prestados.

Nas escolas, os resultados são objeto de estudo detalhado e cada uma produz seu relatório, cumprindo-lhe divulgá-lo na comunidade a que atende. Merece destaque a divulgação que se recomenda seja feita aos pais que, no nível da escola, constituem importante grupo a que interessam os resultados.

Estudos e apresentação de relatos também acontecem nas Delegacias Regionais de Ensino que se responsabilizam pela divulgação, interna e externa, dos resultados nos municípios e na região, atingindo aos usuários indiretos e externos, como autoridades municipais, representações comunitárias e associações diversas.

No órgão central, compete à Diretoria de Avaliação do Ensino a divulgação dos resultados para os diferentes grupos de usuários. Internamente, são encaminhados estatísticas produzidas e relatórios específicos, sobretudo a três áreas: Planejamento Curricular, Capacitação de Professores e Estudos e Pesquisas. A cada uma dessas unidades administrativas, por sua vez, cabe utilizar os resultados em sua área de atuação.

Entre os usuários externos incluem os legisladores (Assembléia e Câmaras), autoridades do mundo político, Sindicatos, Associações diversas, Empresários, Imprensa e público em geral. Para cada grupo são produzidos informes sobre resultados. Dependendo do grupo a que se destina, os resultados são traduzidos em modos de apresentação menos técnicos e mais atrativos sob o ponto de vista de apresentação que predisponham à leitura e à interpretação, como é o caso de gráficos, "folders", relatos sumários, reportagens jornalísticas (entrevistas, "press-releases" etc.).

11.0 - Considerações finais

O Programa de Avaliação do Sistema Estadual de Ensino em Minas Gerais foi estruturado de forma a garantir sua continuidade no tempo, independentemente da possível influência de fatores vários de natureza im-

previsível; desse modo, nessa linha de raciocínio, em 1993, já começaram a ser planejados os instrumentos para avaliação da 3ª, 5ª e 8ª séries, a serem pretestados e aplicados no primeiro semestre de 1994.

O surgimento de uma cultura da avaliação, ainda que incipiente, começou a ser observada, graças a manifestações de alunos, professores e pais, assim como nas iniciativas do próprio sistema em seus três níveis. Denota-se que os resultados da avaliação do sistema têm sido naturalmente incorporados ao conjunto dos indicadores que tem fundamentado planos e realizações da Secretaria de Educação. A definição de municípios de alto risco em questões educacionais, a proposta de conteúdos básicos de 1ª a 4ª séries do Ensino Fundamental, a programação da capacitação docente, para citar apenas algumas, constituem ações recentes da Secretaria de Educação de Minas Gerais em que os resultados das avaliações já realizadas foram significativamente considerados.

Um novo programa de formação de recursos humanos na área da avaliação passou a ser implementado, com a ida de um conjunto de técnicos da Secretaria de Educação, em julho de 1993, para cursos sobre Métodos de Pesquisa em Educação, Avaliação (*Evaluation e Assessment*) e Teoria do Currículo na School of Education da University of Bath, na Inglaterra. Outro grupo, em setembro e outubro desse mesmo ano, participou de intenso *workshop* sobre *Developing a National Assessment Program*, no Educational Testing Service (ETS), em Princeton, USA, juntamente com outros especialistas de várias nações, inclusive representantes dos chamados "tigres asiáticos" (Coréia, Japão, Malásia e Cingapura). A partir de 1994, haverá outros programas de treinamento a curto e médio prazo, no Brasil e no exterior, com a colaboração de assessores nacionais e internacionais, com o objetivo de formar recursos humanos qualificados ainda carentes numa área de extrema importância: – a da avaliação educacional.

A escola pública de Minas Gerais, por sua vez, a partir da análise dos seus resultados, na avaliação externa, começa a apresentar uma nova postura face às metodologias de ensino que utiliza, às estratégias educacionais que emprega e ao próprio processo de avaliação que adota, revelando, assim, um espírito crítico que, embora existente no passado, não era tão explícito como nos dias atuais.

12.0 - Referências bibliográficas

- ANTUNES, A.L. et alii (1992) - Avaliação do Rendimento de Alunos da Escola Estadual de Minas Gerais: o Ciclo Básico de Alfabetização. Estudos em Avaliação Educacional, jul-dez, nº 6. Fundação Carlos Chagas. São Paulo.**
- GOULART, I.B. (1992) - O Ciclo Básico de Alfabetização: Lições de uma Avaliação. Estudos em Avaliação Educacional, jul-dez, nº 6. Fundação Carlos Chagas. São Paulo.**
- HAMBLETON, R.K. et alii (1978) - Criterion-referenced testing and measurement: a review of technical issues and developments. Review of Educational Research, 48(1), pp. 1-48**
- KIBLER, R.J. et alii (1974) - Objectives for Instruction and Evaluation. Boston. Allyn and Bacon. Inc.**
- MARES GUIA NETO, W.S. (1992) - A realidade da educação em Minas Gerais. Estudos em Avaliação Educacional, jul-dez, nº 6. Fundação Carlos Chagas. São Paulo.**
- MILLMAN, J. (1974) - Criterion - referenced measurement. in W.J. Popham (Ed.) - Evaluation in Education: current applications. San Francisco, McCutchan.**
- TYLER, R. W. (1942) - General statement on evaluation. The Journal of Educational Research, XXXV, pp. 492-501**
- VIANNA, H.M. (1980) - A perspectiva das medidas referenciadas a critério. Educação e Seleção, nº 2. Fundação Carlos Chagas. São Paulo.**
- _____. (1982) - Avaliação educacional - problemas gerais e formação do avaliador. **Educação e Seleção, nº 5. Fundação Carlos Chagas. São Paulo.**
- _____. (1990) - Medida da Qualidade em Educação - apresentação de um modelo. **Estudos em Avaliação Educacional, jul-dez, nº 2, Fundação Carlos Chagas. São Paulo.**
- _____. (1992) - Avaliação do Ciclo Básico de Alfabetização em Minas Gerais. **Estudos em Avaliação Educacional, jan-jun, nº 5, Fundação Carlos Chagas. São Paulo.**

