

AVALIANDO A AVALIAÇÃO: da prática à pesquisa*

HERALDO MARELIM VIANNA**

A análise das atuais práticas de avaliação nos diferentes níveis das escolas brasileiras mostra a necessidade da adoção de novas políticas e novos procedimentos para que seja possível o aprimoramento do processo de avaliação dos estudantes. É imperativo que se estabeleçam meios para verificar se diferentes tipos de aprendizagem estão sendo promovidos, sobretudo daqueles que exigem do aluno capacidades mais complexas. O problema não é de fácil solução para a nossa escola, especialmente em relação à prática avaliativa, por não existir concordância sobre questões bem imediatas, como os tipos de instrumento a construir e como chegar aos objetivos a que se propõe o sistema educacional.

Quando se considera, em termos apenas teóricos, as diversas formas de utilização dos instrumentos usados na avaliação educacional, constata-se que poderiam servir a diferentes propósitos:

1. avaliar a eficiência de professores, currículos, sistemas e programas educacionais;

* Trabalho apresentado ao XXIV Encontro Nacional da Fundação AMAB para Educação e Cultura, no Instituto Granbery da Igreja Metodista, em Juiz de Fora, Minas Gerais (14-17.07.92).

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

2. identificar diferentes tendências quanto ao desempenho educacional;
3. determinar o progresso educacional do ponto de vista regional, nacional e até mesmo entre nações;
4. possibilitar a definição e o planejamento de currículos, assim como a definição de novas políticas educacionais (Nickerson, 1989).

A avaliação da eficiência de professores esbarra em interesses corporativistas e não faz parte da nossa tradição, que se limita a avaliar o aluno, assim mesmo de forma bastante precária. A avaliação de currículos, sistemas e programas começa a ser realizada, mas de forma restrita e limitada a algumas poucas experiências, como são o caso, no momento presente, da **Avaliação da Jornada Única e das Escolas-padrão**, em São Paulo, e da **Avaliação do Ciclo Básico de Alfabetização (CBA)**, em Minas Gerais. A experiência mais vigorosa e com maior amplitude é a de Minas Gerais, que até o final de 1994 pretende avaliar todo o sistema educacional da rede oficial, gerando, assim, competências sobre esse tipo de avaliação, que ainda não possuímos, e concorrendo para a formação de uma cultura da avaliação, que ainda não temos.

A nossa carência atual em termos de avaliação decorre em grande parte da falta de continuidade nos trabalhos realizados, por intermédio de um processo de disseminação, que poderia servir de estímulo a outros empreendimentos na área de avaliação. A **FUNBEC - Fundação Brasileira para o Ensino de Ciências** -, nos anos 60 e 70, realizou avaliação de currículos para a introdução de novas metodologias de ensino em Matemática, Física, Química, Biologia e Geociências, mas essa importante prática não teve continuidade em outras instituições, perdendo-se assim, parte do **know-how** adquirido com a colaboração de avaliadores de prestígio, como Hulda Grobman.

A experiência da Fundação Getúlio Vargas, no princípio dos anos 60, merece ser lembrada, tendo sido a primeira vez que se procurou construir um teste padronizado no campo da educação para avaliação de desempenho terminal. O projeto, dirigido por Ruth Schaeffer e Nícia Maria Bessa, no Rio de Janeiro, criou um instrumento que seguia as linhas gerais do **Iowa Basic Skills**, e se destinava à avaliação de capacitações ao término do atual 1º ciclo. O projeto contou com a colaboração de figuras expressivas como Anne Anastasi, Frederick Davis e Robert L. Ebel, que participaram de treinamento para a formação de **expertise** entre educadores brasileiros. Houve um esforço no sentido de criar condições para o desenvolvimento de qualificações, mas, por diferentes razões, o projeto não teve continuidade, restando dessa experiência apenas um manual de interpretação, elaborado com extremo cuidado metodológico, e uma pesquisa sócio-econômica, que são modelos de trabalho científico da melhor qualidade em avaliação educacional.

A década de 70 apresentou grande interesse por avaliação de currículo, assistindo-se, no plano teórico, à divulgação do modelo sugerido por Stufflebeam - contexto, **input**, processo e produto (CIPP) -; no entanto,

como é comum em nosso meio educacional, foi um momento transitório. Alguns trabalhos importantes, como os de Maria Amélia Azevedo e Clarilza Prado de Sousa, entre outros, foram realizados, mas também não tiveram continuidade.

Levantamento recente, a partir da revista **Cadernos de Pesquisa** (Vianna, 1992), editada pela Fundação Carlos Chagas, mostrou que um amplo **espectro** de assuntos na área da avaliação foi analisado, discutido e pesquisado, restando, entretanto, uma indagação: - esses estudos chegaram ao conhecimento do professor e influenciaram na sua prática docente? A avaliação, lamentavelmente, não faz parte da formação dos docentes, quando muito é um tópico isolado, uma aula ou talvez uma unidade, mas não uma área de concentração.

Tentativas de avaliação do sistema educacional foram promovidas pelo **Instituto Nacional de Estudos e Pesquisas Educacionais (INEP/MEC)**, no final da década de 80, em pesquisas realizadas em escolas da rede oficial em 69 cidades do País, com o objetivo de identificar pontos críticos na aprendizagem de crianças do 1º grau (Vianna e Gatti - 1988; Vianna - 1989; Vianna 1989b), e, posteriormente, idêntica pesquisa do rendimento de alunos da rede privada foi concretizada. Houve um levantamento de importantes dados sobre o desempenho em Português, Matemática e Ciências de alunos da escola de 1º grau, mas não se possui um sistema adequado de disseminação de informações, que custam a chegar ao professor, quando realmente chegam, e nem sempre têm ressonância na escola. Ainda no início dos anos 90, o MEC/INEP procurou implantar um **Sistema Nacional de Avaliação do Ensino Público de 1º grau**, envolvendo custo-aluno, rendimento e gestão escolar.

Os anos 70 assistiram ao uso indiscriminado dos testes objetivos, por influência do processo de seleção para o ensino superior. Elaborados sob o signo da improvisação, sem conhecimento de sua complexa tecnologia e dos seus fundamentos estatísticos, acabaram no descrédito. Ao longo desses anos foram ignorados os avanços da psicometria, área da estatística aplicada aos instrumentos de medidas educacionais e psicológicas que apenas tangencia os currículos para a formação de docentes. Seria inteiramente procedente, nesse momento, que se indagasse: para que servem os testes e/ou provas?

Os testes/provas (Nickerson, 1989), quando considerados em relação aos alunos, podem servir para diferentes fins:

1. possibilitar o acesso aos vários níveis de escolaridade em diferentes escolas ou programas de ensino;
2. permitir a promoção em uma seqüência educacional, nas suas várias fases, até a conclusão;
3. positivar deficiências a superar através de uma ação terapêutica;

4. identificar a possibilidade de acelerar (ou retardar) determinado programa;
5. orientar o processo instrucional por intermédio de uma avaliação contínua do desempenho, com a promoção de ajustamentos necessários à concretização da aprendizagem;
6. qualificar professores para o exercício de suas atividades docentes.

A análise dos instrumentos usados em nosso contexto educacional, muitas vezes construídos por instituições com excelente reputação, e, por isso mesmo, supostamente adequados às necessidades da avaliação, mostra elementos críticos, sobretudo no destaque de aspectos ligados à memorização e pouca ênfase no indicativo da capacidade de compreender e saber usar conhecimentos importantes em situações nóveis, revelando o real domínio do conhecimento graças à capacidade de aplicação.

A ausência de instrumentos capazes de medir compreensão e aplicação, objetivos maiores do planejamento educacional, mostra que os projetos de avaliação nem sempre têm condição de positivar se a escola - o sistema educacional - está realizando com sucesso, por intermédio do processo instrucional, um dos aspectos relevantes da tarefa a que se propõe. Isso gera de certa forma uma cadeia de responsabilidades - os instrumentos não são capazes de determinar com êxito capacidades significativas, que a própria escola não as desenvolve conforme seria desejável. Assim, por via de consequência, a avaliação acaba prejudicada.

Existe toda uma cadeia de elementos interrelacionados - os instrumentos de medida refletem a filosofia instrucional seguida em classe pelos professores e esses mesmos instrumentos (deficientes) determinam a forma pela qual os estudantes adquirem supostas capacidades. É necessário ressaltar que as avaliações são prejudicadas na medida em que os professores ensinam para o teste/prova, considerando que os seus resultados muitas vezes são usados para determinar a eficiência do professor.

Ainda que compreensível do ponto de vista do professor, o **ensinar para a prova** acaba por comprometer o processo de avaliação no que tange à sua validade preditiva. É preciso levar em conta que não faz sentido **preparar para o exame**, tendo em vista os objetivos do processo instrucional. A idéia de ensinar para o teste, apesar de partir do pressuposto de que as provas determinariam o que os professores ensinam e os alunos estudam, não é defensável, se for considerado que os instrumentos de avaliação nem sempre avaliam o relevante e desejável; desse modo, estaria sendo dada ênfase a atributos menores, em detrimento de capacidades mais importantes que, porém, não foram desenvolvidas face à relação ensino-teste-avaliação.

A questão de avaliar aquilo que é importante não é de solução fácil e imediata, tendo em vista certas constatações: os instrumentos para a medida de pro-

cessos complexos são difíceis de construir, exigem pesquisas e investimentos, o que até agora não tem sido feito. A medida desses processos exigiria a avaliação de conceitos, princípios e relações, aspectos da aprendizagem nem sempre desenvolvidos em nossos currículos, que se preocupam com o domínio dos conhecimentos mais simples, quase sempre factuais e imediatos.

É preciso pensar e repensar o problema da avaliação no contexto brasileiro: - avaliar é difícil, avaliar no campo educacional é extremamente difícil, e muitos não se dão conta das suas implicações pedagógicas e das amplas ressonâncias, inclusive no social e no econômico. A repetência, por exemplo, afeta a auto-estima, e nem sempre é uma garantia de futuro êxito; ao contrário, as possibilidades de sucesso diminuem consideravelmente. Essa é uma preocupação do atual Governo de Minas Gerais (1992), que em documento oficial declara **haver todo um esforço no sentido de assegurar que o número de concluintes do ensino fundamental, com no máximo duas repetências, se eleve dos 15-18% atuais para 40% em cinco anos e para 60% em dez anos, quando, então Minas Gerais se terá equiparado ao México de hoje.** A um custo médio de 220 dólares por aluno-ano, a repetência em Minas Gerais custa 110 milhões de dólares.

A psicologia educacional, nos dias fluentes apresenta grandes e novas concepções sobre aptidão, aprendizagem, desenvolvimento e rendimento escolar (Snow, 1989; Snow e Lohman, 1989.) Tudo isso está tendo amplas repercussões na área da avaliação educacional, especificamente nas pesquisas sobre validade de construto. Essa nova situação leva a crer que diferentes níveis e diversos modelos de avaliação devem ser adotados, sem abandono dos atuais, mas com um significado inteiramente diferente. O próprio professor precisa ser igualmente pesquisado, porque nada de importante se conseguirá sem a devida compreensão do seu papel e sem a sua irrestrita adesão aos trabalhos de avaliação.

A avaliação cognitiva é complexa e a sua complexidade vem aumentando na medida em que a psicologia cognitiva apresenta novos desafios, desvendando a multiplicidade de estratégias que levam a criança a diferentes tipos de aprendizagem. As crianças sabe-se hoje em dia (Siegler, 1989), usam diferentes estratégias relativamente à soma, à subtração, à multiplicação, à soletração, à identificação de palavras, entre outros aspectos. A consequência disso é a valorização de um ensino voltado para as diferenças individuais e para uma avaliação individualizada, do tipo formativo e baseada em critério e não em normas.

A avaliação em nosso contexto educacional acha-se na fase da pré-história, preocupada com problemas menores e sem significado efetivo para o sucesso de uma aprendizagem realmente conseqüente para o ser humano diante dos desafios que a sociedade constantemente apresenta. Assim, como avaliar as diferentes estratégias usadas na área cognitiva? A indagação está apresentada,

mas dificilmente ter-se-á uma resposta satisfatória e imediata. É um **challenge** que devemos aceitar, esperando que algum dia se possa chegar a uma solução definitiva para o problema, que possivelmente será obtida por meios não convencionais, em relação às atuais práticas de avaliação.

A avaliação procura entre seus objetivos verificar a capacidade de pensar criticamente, sem se dar conta de que somente se mede e avalia o que se pode definir operacionalmente, o que torna difícil a mensuração desse construto, face ao caráter extremamente vago do conceito e à impossibilidade de fixar padrões de julgamento (Norris, 1989). O assunto, por sua importância, precisaria ser pesquisado a fim de aprofundar o problema e identificar formas de avaliar o pensamento crítico no contexto dos problemas da realidade, o que significa dizer que há necessidade da determinação de um número bastante considerável de variáveis intervenientes nesse complexo problema.

É praticamente impossível discutir avaliação sem tratar, necessariamente, do problema da validade (Norris, 1989; Fredericksen e Collins, 1989). Até que ponto a preocupação com os múltiplos problemas da validade chegou até nós é um questionamento que deveria ser objeto da reflexão de todos os avaliadores, inclusive de professores. Os testes aplicados em sala de aula são realmente válidos? Os instrumentos usados em pesquisas educacionais apresentam comprovada validade? Quais as evidências empíricas de que os resultados obtidos são realmente válidos? O que validar: - o instrumento, os resultados ou a idéia que constitui a infra-estrutura de todo o processo de avaliação? Quase sempre avaliamos construtos: rendimento escolar, capacidade de expressão escrita etc., mas qual a prova que se possui da validade de construto? É preciso atentar para as conseqüências de uma avaliação realizada sem base em uma fundamentação teórica (**rationale**) e sem evidências empíricas que sustentem possíveis inferências e decisões com seus múltiplos resultados práticos, às vezes irreversíveis.

Os diversos problemas da avaliação precisam ser pesquisados para que a avaliação realizada na escola venha a ter um papel importante no processo ensino-aprendizagem, no qual está integrada. Somente a pesquisa sobre avaliação e a prática constante da avaliação serão capazes de criar uma cultura da avaliação e dar credibilidade à avaliação no contexto das atividades educacionais, sujeitas a constantes desafios. A informática, por exemplo, começa a ser introduzida em nossas escolas de maneira ainda incipiente, mas o uso de computadores na área educacional significará um novo repto à avaliação, que se verá diante da possibilidade de poder avaliar como as pessoas compreendem e pensam (Frederick e Collins, 1989); contudo, ainda temos um longo caminho a percorrer, necessitamos formar novas competências, a fim de enfrentar os problemas em benefício de um ensino eficiente em nossas escolas, sobretudo as responsáveis pela educação fundamental.

BIBLIOGRAFIA

- FREDERICKSEN, J. R. e COLLINS, A. (1989). A Systems Approach to Educational Testing. *Educational Researcher*, 18(9) pp. 27-32.
- NICKERSON, R. S. (1989). New Directions in Educational Assessment. . *Educational Researcher*, 18(9) pp. 3-7.
- NORRIS, S.P. (1989). Can we test validity for critical thinking? *Educational Researcher*, 18(9) pp.21-26.
- SIEGLER, R.S. (1989). Strategy Diversity and Cognitive Assessment. *Educational Researcher*, 18(9) pp. 15-20.
- SNOW, R.E. (1989). Toward Assessment of Cognitive and Conative Structures in learning. *Educational Researcher*, 18(9) pp. 8-14.
- SNOW, R. E. e LOHMAN, D.F. (1989). Implications of cognitive psychology for educational measurement, in R.L. LINN (ed.). *Educational Measurement*. 3rd. ed. pp. 263-331. New York, Macmillan.
- VIANNA, H.M. e GATTI, B.A. (1988). Avaliação do Rendimento de Alunos de Escolas do 1º Grau da Rede Pública: uma aplicação experimental em 10 cidades. *Educação e Seleção*, jan./jun., nº 17, Fundação Carlos Chagas, São Paulo.
- VIANNA, H.M. (1989a). Avaliação do Rendimento de Alunos de Escolas do 1º Grau da Rede Pública: um estudo em 20 cidades. *Educação e Seleção*, jan. /jun., nº 19. Fundação Carlos Chagas. São Paulo.
- _____ (1989b). Avaliação do Rendimento de Alunos de escolas do 1º Grau da Rede Pública: um estudo em 39 cidades. *Educação e Seleção*. jul./dez., nº 20, Fundação Carlos Chagas, São Paulo.
- _____ (1992). Avaliação Educacional nos Cadernos de Pesquisa. *Cadernos de Pesquisa*, 80 pp. 100-105. Fundação Carlos Chagas. São Paulo.

