

O QUE VALE EM AVALIAÇÃO

Hermengarda Lüdke (1)

Em junho de 1981*, depois de um longo período de pós-doutoramento nos Estados Unidos, achei conveniente fazer uma revisão do que havia observado em termos de avaliação educacional, antes de voltar para o Brasil. Seria um pouco como se preparasse a minha bagagem de volta, selecionando os itens que deveria levar para o trabalho com avaliação. Claro que essa seleção refletiria muito mais as necessidades que sinto no Brasil, em matéria de avaliação, do que a vasta gama de recursos disponíveis dentro dessa disciplina nos Estados Unidos. Em outras palavras, eu não estaria tentando fazer um balanço geral da avaliação educacional americana, mas antes, dentro de uma visão bem pessoal, escolhendo suas contribuições mais significativas atualmente, para responder aos nossos problemas de avaliação.

Tendo deixado claro o caráter seletivo de minha listagem, organizarei minha discussão ao redor dos seguintes itens: 1 – “accountability”**, 2 – mensuração, 3 – testes, 4 – avaliação como profissão, 5 – novas abordagens em avaliação, 6 – o componente político da avaliação e 7 – o caráter julgador da avaliação.

1. “ACCOUNTABILITY”

Segundo a definição de Michael Scriven, “accountability” é “responsabilidade pela justificação das despesas ou do esforço de alguém. Portanto, os diretores de programas e professores deveriam ser, diz-se frequentemente, responsáveis (“accountable”) pelos seus custos, salários e tempo” (Scriven, 1980).

O conceito de “accountability” teve uma enorme influência no desenvolvimento da avaliação educacional nos Estados Unidos, nas duas últimas décadas. Ele ajudou a encontrar con-

(1) Professora da Pontifícia Universidade Católica (RJ) e da Universidade Federal Fluminense.

* Gostaria de registrar aqui meus agradecimentos ao professor Mark St. John, da Universidade da Califórnia, Berkeley, que leu os originais e fez comentários muito valiosos.

** Resolvi deixar o termo “accountability” no original inglês, pois não encontrei um termo que correspondesse exatamente a ele em português, sendo possivelmente o de “responsabilização” o mais aproximado, como já indiquei em trabalho anterior (Lüdke, 1975).

tomos mais precisos para essa disciplina, dentro do domínio da educação, indicando a necessidade do estabelecimento de um acordo entre os que prestam e os que esperam serviços em educação. Assim, a avaliação apareceria como uma espécie de ponte entre os que fazem e os que recebem esses serviços.

Tal como foi proposto nos anos 60, o conceito ("accountability") pareceu muito útil para os profissionais da educação; ele os ajudou a organizar seu trabalho de ensino de maneira mais racional e a orientar seus esforços para um produto mais visível e contável ("accountable"). De maneira especial, a idéia de limitação da responsabilidade em educação estava implícita naquele conceito: ninguém poderia ser apontado como responsável pela falha no atingimento de metas, se as condições necessárias para esse atingimento não fossem suficientemente garantidas. Esta concepção original de "accountability" parecia particularmente adequada para lidar com problemas da educação brasileira. Potencialmente, ela poderia ajudar a aliviar o sentimento de culpa de professores que se sentem falhar no cumprimento de metas que consideram importantes, mas para o qual não dispõem das condições mínimas necessárias. Num país onde as prioridades educacionais raramente coincidem com as orçamentárias, a frustração entre os professores atinge níveis muito elevados. Qualquer fator que possa ajudar a aliviar esse sentimento negativo é extremamente bem-vindo, como seria essa visão do conceito de "accountability".

A evolução subsequente desse conceito nos Estados Unidos, durante a década de 70, produziu, entretanto, alguns efeitos negativos, talvez devido à sua aplicação abusiva. A intenção de organizar o ensino de modo mais racional freqüentemente veio a ser interpretada de maneira muito estreita, concentrando-se na questão dos objetivos mais específicos e concretos. Nas palavras de Scriven: "Uma falácia comum associada à "accountability" é supor que a justiça requer a formulação de metas e objetivos precisos, se alguma "accountability" deve haver; mas de fato uma pessoa pode ser apontada como responsável ("accountable") pelo que faz, mesmo dentro da concepção mais geral de trabalho profissional, como, por exemplo, ensinar estudos sociais no 2º grau", ainda que não formule objetivos precisos (Scriven, 1980).

A interpretação mais estreita de "accountability" logo se alastrou e dominou o cenário educacional americano. Sentia-se que para se tornar creditável ("accountable"), o trabalho do professor deveria ser especificado em termos de unidades quantificáveis, tais como a percentagem de alunos que atingem os escores mínimos pré-especificados de um teste, dentro de uma escala de tempo determinada, em função de detalhados objetivos instrucionais. Nem todos os aspectos do processo de ensino/aprendizagem podem, entretanto, ser tão prontamente quantificáveis; portanto, aqueles aspectos de mais difícil mensuração sofrem um risco mais alto de serem ignorados, a despeito de sua importância educacional. Aliás, a própria quantificação de fenômenos educacionais implica numa questão muito mais profunda, relativa ao quadro conceitual onde se insere a possibilidade mesma de mensuração desses fenômenos. Deve ser observado também que nem todo evento no processo educacional pode ser previsto com precisão, de tal forma que seria necessário um esforço gigantesco, com métodos muito cuidadosos para tentar abarcar de maneira justa todo o trabalho do professor. Muito do que acontece dentro de uma sala de aula escapa à possibilidade de medida por um instrumento, ou mesmo por vários.

Se o conceito de "accountability" for aplicado estritamente, importantes aspectos da educação serão certamente sacrificados. Para os professores haverá o risco de confinar o escopo do ensino dentro das dimensões "oficiais" esperadas, aquelas incluídas nos critérios avaliativos do produto final. Atualmente, o espaço reduzido do ensino de arte no currículo das escolas elementares americanas é apenas um exemplo das importantes modificações introduzidas nos últimos anos, sob a pressão na direção de produtos mais contabilizáveis, usualmente identificados aos "três R's" (ler, escrever e contar).

A insistência na "accountability" pode também levar a uma polarização, com professores de um lado e administradores (e avaliadores) de outro. Os dois grupos, sujeitos a um modelo estabelecido para obter maior eficiência na educação, "crescentemente vêm-se a si próprios como gerentes de um lado e trabalhadores de outro" (Eisner, 1979), com todas as conseqüências dessa oposição.

Uma observação geral da evolução do currículo nas escolas elementares americanas sugere

que uma certa distorção tende a se infiltrar nas expectativas educacionais, através da adesão ao conceito de "accountability" tal como foi descrito: professores e alunos tendem a concentrar seus esforços nas metas pré-especificadas e contáveis, porque elas são as únicas que contam, oficialmente. Assim, procura-se evitar, naturalmente, a possibilidade de críticas oriundas de instâncias superiores, mas, conseqüentemente, deixa-se de procurar desenvolver um trabalho criativo com as crianças.

Baseada em minha percepção do que sucede no momento nos Estados Unidos, acho que posso antecipar que o uso do conceito de "accountability" no Brasil seria suplantado pelo seu abuso. A energia natural dos professores seria concentrada na perseguição de metas artificialmente prescritas e um processo reducionista se seguiria rapidamente. Uma nova força constrangedora se acrescentaria aos já numerosos problemas encarados por professores, alunos e administradores.

A concepção original desse conceito, ligada a uma justa atribuição de responsabilidade em educação, ainda parece muito atraente. Mas a sua evolução naquele país mostra que sua adoção para o nosso sistema educacional demandaria um cuidadoso exame crítico, especialmente das lições negativas lá aprendidas a partir de sua aplicação.

2. MENSURAÇÃO

Este é um outro item importante da bagagem avaliativa. Muito desenvolvido nos Estados Unidos, este conceito tem sido ligado de tal modo ao de avaliação a ponto de se identificarem, durante um certo período.

A contribuição da mensuração à avaliação é enorme, em termos de recursos teóricos, conceitos, instrumentos e treinamento prático. Tradicionalmente, a preparação de avaliadores educacionais tem dependido em grande parte do seu domínio de habilidades de mensuração e de análises estatísticas. Uma grande quantidade de testes e instrumentos de mensuração vem sendo desenvolvido e difundido no sistema educacional americano. Tudo isso exerceu, a seu tempo, um efeito muito positivo sobre a avaliação, chamando a atenção dos especialistas para a necessidade de medidas válidas e precisas, adequadas para um melhor conhecimento da realidade educacional.

O sucesso do movimento de mensuração foi tão grande e tão atraente a esperança de se conseguir medidas válidas para tudo em educação, que capitalizaram grande parte do esforço e da energia dos avaliadores, por um considerável período de tempo. Mas, certamente, isso não haveria de fornecer a resposta completa que procuravam, pois não há exata coincidência entre avaliação e mensuração. Ainda que muito importante, a parte estava, indevidamente, sendo tomada pelo todo.

Avaliação requer mais que mensuração. A natureza do processo educacional em si tem dimensões que não poderiam ser alcançadas pela mensuração imediata. Como poderiam ser medidos, através da vida de um indivíduo, os efeitos a longo termo da escolaridade? Mesmo efeitos a médio termo são difíceis de medir, devido à complexidade de fatores sempre envolvidos na composição do processo de educação.

A excelente preparação oferecida pelos cursos sobre mensuração, no treinamento de especialistas em avaliação, torna-os muito alertas para as dificuldades relacionadas ao desenvolvimento de bons instrumentos de medidas educacionais. Quanto melhor informados eles estão, tanto mais conscientes das exigências para a construção de um bom instrumento e mais inclinados a lidar com conteúdos que possam mais facilmente responder a essas exigências.

Como conseqüência, numerosos instrumentos têm sido desenvolvidos para a mensuração daqueles assuntos que têm um conteúdo simples e bem delimitado, enquanto outros, cuja natureza é mais delicada, menos tangível, portanto menos redutível à uma boa mensuração, têm sido deixados a descoberto, o que é compreensível. Se a avaliação basear-se estritamente em bons instrumentos de mensuração, seu foco será também reduzido àquelas áreas em que elas são mais exequíveis e disponíveis. Essa tendência, aliás, já se acha impressa na avaliação, desde algum tempo, como resultado da combinação do conceito de mensuração com o movimento pela "accountability". Particularmente esse é o caso dos Estados Unidos. Por um lado, a demanda por serviços educacionais contabilizáveis, traduzidos por uma unidade que seja confiável e capaz de exibir o lucro do esforço; por outro, o desejo de desenvolver instrumentos tão acurados quanto pos-

sível, para assegurar medidas estáveis e seguras dos resultados.

Esta discussão evoca uma outra ainda muito mais profunda, sobre os próprios pressupostos da mensuração em ciências humanas, e que não cabe nos limites deste trabalho. Até que ponto pode-se admitir em educação o exercício da análise, que decompõe o fenômeno focalizado em partes, para fins de construção de instrumentos de medida, sem comprometer a integridade de seu todo? O esforço de mensuração tem certamente obtido resultados extremamente úteis para responder a muitas questões, mas não a todas. Fica aqui apenas levantado esse importante ponto, cuja discussão eu gostaria de aprofundar em outro trabalho.

Os dois movimentos juntos, o de mensuração e o de "accountability", acarretaram inevitavelmente à avaliação um alto risco de reducionismo. Metas e objetivos passaram a ser considerados mais em função de sua possibilidade de acesso por instrumentos de medida do que de sua importância. Metas difíceis de se avaliar passaram a correr o risco de serem evitadas pelos planos educacionais. Por serem difíceis de contabilizar, elas podem deixar de contar . . .

Mesmo a noção de metas, tão amplamente difundida na avaliação americana, e também na nossa, merece um questionamento, que aliás já vem sendo feito há algum tempo. Ela própria faz parte, possivelmente, de toda uma inclinação para a segurança, garantida por um trabalho pré-fixado e planejado; portanto, acima de críticas fáceis. Trata-se de objetivos propostos, refletidos, que merecem por isso consideração! Trata-se também de idéias, que apontam o caminho a seguir, mas que relação guardam efetivamente com a realidade? Não é possível, igualmente, desenvolver plenamente este ponto aqui. Ao menos fique assinalado aqui o importante questionamento a respeito, feito por um avaliador americano de aguçado senso crítico, Michael Scriven, e que se tornou muito conhecido sob a expressão "Goal free evaluation" * (Scriven, 1977).

Há muitos outros aspectos importantes relacionados com a questão de mensuração, cuja discussão é válida, mas, infelizmente, não cabe aqui. Entre eles se encontram os benefícios de uma boa preparação na área de medidas, o poder de bons e confiáveis instrumentos e também algumas atitudes alternativas sobre seu uso em avaliação. Penso que é possível se desenvolver, dentro da educação, uma posição em que possamos fazer o melhor uso de boas mensurações, sem sermos apanhados por suas limitações. Um senso de equilíbrio entre o ideal e o exequível deveria ser buscado, sem comprometer a boa qualidade dos instrumentos nem a amplitude dos fins em educação. Dever-se-ia pensar numa estratégia que permitisse a aplicação da boa prática avaliativa para a melhoria da educação, nunca para seu estorvo, como rezava a admoestação clássica de Cronbach (1963), um dos principais fundadores da disciplina.

Talvez a solução esteja numa posição situada entre os dois extremos. Ela poderia combinar a medida de certas metas intermediárias, por exemplo, as de implementação, com a esperança de que as de longo alcance sejam atingidas. Deveríamos ser capazes de aceitar certo grau de incerteza, desde que nem tudo pode ser representado por medidas exatas. E, sobretudo, deveríamos ser capazes de manter a crença de que estamos atingindo importantes metas em educação, mesmo se não pudermos provar isso com números.

3. TESTES

Aqui está um sub-produto do esforço de mensuração que cresceu suficientemente para merecer uma consideração especial. Sua importância é tão grande que dificilmente poderíamos imaginar a vida dentro das escolas, hoje em dia, sem eles, pelo menos nas escolas americanas. O comportamento de crianças e professores tem sido modelado de acordo com eles nas últimas duas ou três décadas.

Criado para ser um recurso útil, nas mãos de professores, para averiguar e relatar o aproveitamento das crianças, nas diferentes áreas do currículo, ele se tornou um fator determinante

* Avaliação independente de objetivos pré-fixados.

desse aproveitamento, por vezes mesmo o fator determinante. Quando se observa que professores e alunos orientam seu comportamento primordialmente para o objetivo de obterem sucesso no teste, sua conotação perniciosa torna-se óbvia. Uma espécie de mimetismo irresistível faz com que o comportamento da classe escolar reproduza as expectativas do teste, o que é compreensível, já que o teste é o mais importante sancionamento daquele comportamento. Numa civilização altamente competitiva, ele encontrou ambiente extremamente favorável para seu desenvolvimento.

Os testes tiveram seu papel inflacionado no cenário educacional, não apenas nos Estados Unidos, mas também em outros países. No Brasil também já se constatou essa síndrome, embora em grau menor, especialmente a nível do vestibular, com más conseqüências para as escolas de 2º grau.

O exame vestibular, com seu sistema de testes, agiu sobre as escolas de 2º grau de uma maneira marcante, transformando o currículo destas num espelho daquele exame. O objetivo primordial dessas escolas, explícita ou implicitamente, se tornou preparar os estudantes para "passar" nos testes daquele exame, isto é, desenvolver em cada estudante a habilidade quase mecânica de passar em testes específicos. Não importava, aos planejadores de currículo dessas escolas, se muitos objetivos educacionais importantes foram ignorados ou mesmo negados por esta busca quase exclusiva de um automatismo. A visão deles estava conformada, é preciso reconhecer, a um modelo de testagem muito em moda há quinze ou vinte anos atrás. A esta altura já se está procurando minimizar os efeitos desse tipo de testagem com a introdução de provas de tipo discursivo.

A lição negativa da supertestagem, extraída do exemplo indicado, poderia ajudar a prevenir os níveis anteriores de ensino contra a sua invasão. Neles, os efeitos seriam extremamente danosos, pois atingiriam uma população estudantil muito mais numerosa. Certamente isso iria enfatizar ainda mais o já agudo caráter seletivo do nosso sistema educacional.

Testes, mensuração e "accountability" combinados têm trabalhado na direção de uma imagem muito mecânica da educação. O reforço mútuo que se deram quase acabou por desenvolver uma caricatura do que na realidade a educação deveria ser. É preciso reconhecer, não entanto, que todos esses esforços começaram com as melhores intenções e que se tivessem sido usados com discrição, dentro dos seus limites naturais, talvez a educação não apresentasse, hoje, essa imagem desfigurada que às vezes até nos assusta.

4. A AVALIAÇÃO COMO PROFISSÃO

Embora haja, nos Estados Unidos, profissionais especificamente dedicados à avaliação educacional, não acredito que no Brasil o mesmo possa suceder. Eu diria até que o próprio educador, em nosso país, está longe de atingir um nível de trabalho nitidamente profissional, enquanto educador.

Nos Estados Unidos, o avaliador tem um papel específico, ao lado de outros especialistas em educação e recebe um treinamento acadêmico bem planejado e altamente complexo, dentro das Faculdades de Educação. No âmbito da prática ele é geralmente responsável pelo planejamento da avaliação, escolha ou confecção dos seus instrumentos e muito frequentemente pela inteira execução do plano.

Esta posição específica do avaliador faz com que ele apareça, aos olhos dos educadores engajados na prática, como a única pessoa responsável pela função avaliativa, em todos os seus estágios. Ele é o especialista, ao mesmo tempo respeitado e temido. Absorvendo as funções avaliativas, sendo "o profissional responsável", o avaliador, às vezes, toma dos professores, diretores e outros membros do processo educacional suas próprias responsabilidades avaliadoras. Todos eles passam a esperar que o especialista proporcione orientação profissional e se inclinam a seguir seu conselho, ainda quando, eventualmente, este não coincide com a prática ou as expectativas existentes antes da chegada do especialista.

A esse respeito, o aspecto mais sério a considerar é a possibilidade de se estabelecer um

fosso entre o avaliador e os outros profissionais da escola; em alguns casos até com o aparecimento de um sentimento de temor em relação ao primeiro. O problema básico aqui é que a presença do especialista pode causar a cristalização das funções avaliativas. Corre-se o risco de um enriquecimento dessas funções de um modo que é totalmente oposto ao ideal da avaliação, que é ajudar o processo educacional, tentando melhorá-lo e permanecendo bastante flexível para responder às suas reais necessidades.

Assumindo uma posição superpoderosa, o avaliador atrai para si a imagem de um controlador, da pessoa responsável pela realização de um certo número de coisas, de uma determinada maneira — em outras palavras, pela manutenção de uma certa ordem. Naturalmente, essa pessoa tem mais chance de inspirar sentimentos de medo do que de identificação.

Seria diferente se a avaliação fosse percebida como uma função própria de todos e de cada um dos profissionais que trabalham numa escola ou num programa educacional, se cada um assumisse a responsabilidade pelo processo de avaliação, como parte do próprio processo educacional.

O medo da avaliação é, aliás, um estranho fenômeno, cuja presença é surpreendentemente vívida no sistema educacional americano. Se considerarmos que há mais de vinte anos Cronbach (1963) escreveu seu memorável artigo, propondo um curso de ação para a avaliação totalmente a serviço da melhoria da educação, precisamos reconhecer que ainda há muito a ser feito para converter aquele propósito em ação. Às vezes, tem-se a impressão de que o verdadeiro é o inverso, isto é, a educação existe para preencher as exigências da avaliação.

Algo desse sentimento negativo vem do conflito de papéis desempenhados pela avaliação e pelo avaliador, ora ajudando a construir o programa, a desenvolver o currículo, na avaliação formativa, ora julgando-os e classificando-os ao final, na avaliação somativa. Isto é, algumas vezes, o avaliador assume com o pessoal do programa a posição de membro, e outras vezes se coloca do lado de fora, como juiz.

Acho que, ao menos, parte desse sentimento negativo poderia ser evitada se a avaliação fosse realmente assumida como uma função integrada na vida da escola ou do programa. E, sobretudo, se o avaliador fosse considerado como um a mais entre os outros membros do grupo. Talvez o avaliador pudesse atuar como um sensibilizador de todo o grupo para certos problemas específicos que ele distingue melhor e para cuja solução ele pode sugerir estratégias próprias. No entanto, caberia ao próprio grupo envolvido assumir realmente a responsabilidade da solução desses problemas.

A observação do que se passa a este respeito nos Estados Unidos, com a educação e a avaliação localizadas nem sempre do mesmo lado da cerca, suscita uma reflexão crítica a propósito da oportunidade de uma especialização em avaliação. Além do possível e já mencionado fosso afastando os diferentes profissionais, há também o risco de se estimular o desenvolvimento de um novo especialista, mais preocupado com os estritos interesses de sua própria especialidade do que com os problemas da educação em si. A preparação desse profissional, tal como é feita nos Estados Unidos, pode, aliás, tomar estranhos caminhos, verdadeiros desvios, que acabam, às vezes, por desembocar em posições bastante distantes das preocupações reais da prática educacional. Estou pensando, particularmente, em algumas das exigências para a realização de certo tipo de pesquisa avaliativa.

5. NOVAS ABORDAGENS EM AVALIAÇÃO EDUCACIONAL

Essas abordagens representam tentativas férteis para encontrar novas maneiras de lidar com as questões de avaliação. São abordagens paralelas e complementares às tradicionais, que se apresentam, em geral, grandemente dependentes de métodos quantitativos. As novas abordagens estudam os mesmos problemas educacionais de diferentes modos, esperando alcançar aspectos não revelados pelas abordagens tradicionais. Com esta intenção, alguns caminhos inteiramente novos têm sido explorados, assim como alguns outros já percorridos por outras disciplinas.

Comentarei aqui apenas algumas dessas novas metodologias, dentre as que vêm sendo

desenvolvidas pelos interessados em avaliação nos Estados Unidos. A mais popular delas é, provavelmente, a abordagem etnográfica de avaliação. Apoiando-se num recurso há muito tempo usado e aprovado pelos antropólogos, a avaliação vem tentando, na última década especialmente, olhar os seus objetos de estudo da maneira a mais próxima possível da visão que deles têm os participantes do fenômeno que está sendo estudado. Observando e participando dessa realidade, o avaliador deveria ser capaz de fazer um relato que desse conta de toda a situação, mesmo que essa imagem fosse limitada ao foco de apenas um dos grupos participantes. Não seria possível descrever a realidade tal como é vista por um e por todos os participantes, mas, ainda que restrita à visão de apenas um deles, a meta é tentar comunicar a totalidade do fenômeno.

Esse traço dá à avaliação uma possibilidade nitidamente nova: a de integrar num quadro total os pedaços e fragmentos recolhidos pelo investigador. Além dos instrumentos de análise, tão bem desenvolvidos pela abordagem tradicional, o avaliador dispõe, agora, da ajuda dessa nova maneira de buscar uma imagem integrada da totalidade. Isso, certamente, será útil no esforço de entregar aos usuários um relato avaliativo que faça sentido para eles, que seja congruente com a realidade em que vivem e do qual eles possam retirar elementos para melhorar essa realidade, ou mesmo para conhecê-la. . . Ao invés de explicações pelas relações de causa e efeito, entre as partes decompostas pela análise, esta nova abordagem procura antes a reconstrução de um todo que possa ser entendido.

Há um perigo que deveria ser logo assinalado, quando se trata de novas metodologias e que é representado pela ameaça de exaustão e morte prematura, característica dos modismos de sucesso rápido. Seria altamente recomendável o estudo cuidadoso da bibliografia já disponível, antes de uso indiscriminado e possivelmente comprometedor de uma nova metodologia. Em inglês, Ray Rist (1980) preparou uma exaustiva e útil bibliografia anotada e, entre nós, já temos atualmente publicado o artigo de Marli André (1978).

Particularmente importante e promissor entre as novas metodologias é o estudo de caso, que fez sua entrada triunfal na avaliação educacional americana especialmente pelas mãos de Robert Stake e seus associados na Universidade de Illinois, em Urbana/Champaign. Com uma longa e respeitada reputação entre os antropólogos, o estudo de caso inicia o que poderá ser uma brilhante carreira na educação. Pela leitura de alguns dos estudos de caso em educação científica, produzidos pela C.I.R.C.E.*, pode-se sentir um ar fresco soprando sobre o campo da avaliação. Certamente há muitas questões a serem esclarecidas antes que ele se torne um instrumento de largo uso na avaliação. Essas questões estão ligadas principalmente a aspectos metodológicos, tais como o tempo básico mínimo para a realização de um estudo de caso, a forma adequada de redigir seu relato, o número mínimo de pessoas a serem entrevistadas antes de fechar o caso ou como decidir sobre o tamanho da unidade a ser estudada. Estas e muitas outras questões ainda esperam por respostas adequadas, que virão provavelmente do próprio exercício de estudos de caso. Somente uma considerável prática acumulada a respeito, acompanhada da necessária reflexão, pode indicar os caminhos a serem seguidos.

É ainda cedo para o estabelecimento de regras, mas os exemplos já disponíveis asseguram, para um futuro próximo, a contribuição efetiva dessa abordagem muito útil para os estudos de avaliação. De minha parte, estou disposta a estimular sua aplicação aos nossos próprios proble-

* Center for Instructional Research and Curriculum Evaluation, University of Illinois, Urbana/Champaign, U.S.A.

mas educacionais, seja através de meus próprios estudos, seja pelos dos orientandos dos programas de pós-graduação onde trabalho*.

Outro sopro de ar fresco sobre a avaliação se acerca através da abordagem desenvolvida por Elliot Eisner, da Universidade de Stanford. Seus dois conceitos interrelacionados, o de "connoisseurship" e o de criticismo, aplicados à educação, podem ser muito úteis aos avaliadores, em seus esforços para encarar os processos educacionais de maneira mais realística. Para ser um "connoisseur", uma pessoa precisa estar profundamente envolvida em seu objeto de estudo, experimentando-o, direta e pessoalmente, por um período de tempo suficiente. Para ser um crítico, a pessoa deveria ser capaz de traduzir em palavras o conhecimento adquirido pelo envolvimento e pela experiência com o objeto de estudo. E a tradução deveria comunicar aquele conhecimento a outras pessoas que não têm acesso ao mesmo tipo de experiência, ou não estão preparadas para captar e aproveitar todas as mensagens vindas dela. Assim como os críticos de arte podem ajudar os não iniciados, admiradores da arte, a apreciar e usufruir mais plenamente das obras de arte que observam, assim o crítico educacional deveria ser capaz de tornar mais clara e compreensível, para um educador menos experiente ou menos informado, uma específica situação educacional ou um problema.

A vida numa sala de aula, com toda a sua complexa teia de relações e multiplicidade de reações simultâneas, aparentando, por vezes, uma imagem de caos, pode se tornar transparente, através da visão de um criticismo cuidadoso, que busca as razões situadas atrás dos movimentos "brownianos" da superfície. Esta qualidade, isto é, a habilidade de ver através de uma confusão aparentemente opaca e de chegar a explicações que fazem sentido, e que fazem crescer a compreensão a respeito, é o fruto de uma longa preparação na área específica, teórica e praticamente, ao lado do cultivo de uma fina sensibilidade, sintonizada especificamente em questões educacionais. Evidentemente, o desenvolvimento dessa qualidade não é fácil e certamente nada rápida. Talvez ela chegue, como uma demorada conquista, a ser adquirida em etapas sucessivas, cada uma delas dependendo da apreensão de um grau mais elevado de consciência e capacidade. Alguns chegarão até os últimos estágios, outros pararão antes disso, mas todos crescerão dentro do processo. Para mais ampla discussão destes conceitos, veja-se o livro de Eisner (1979), cujo título já é bem sugestivo: "A imaginação educacional".

Gostaria de mencionar, rapidamente, algumas outras tentativas de inovação em matéria de técnicas avaliativas ora em curso nos Estados Unidos. Todas representam maneiras originais e até mesmo inesperadas de lidar com questões de avaliação, permitindo uma certa liberdade em seu modo de abordagem, de onde podem surgir interessantes e reveladoras surpresas. Esses são os casos, entre outros, do uso de metáforas, de fotografias ou de jornalismo investigativo em avaliação. Todas essas sugestivas e tentadoras propostas estão sendo cuidadosamente estudadas em centros de estudos de avaliação, nos Estados Unidos, especialmente no Northwestern Laboratory, sob a direção de Nick Smith (Smith, 1978).

Falando de maneira geral, as novas abordagens aqui mencionadas, e outras não mencionadas, representam tentativas de enriquecer a avaliação com novos recursos não exatamente antecipados pela abordagem tradicional já estabelecida. Elas quebram a uniformidade das

*No momento da publicação deste artigo, já se registram alguns eventos importantes nesse sentido. Um deles foi a participação de Robert Stake no *Seminário Internacional de Avaliação*, realizado na PUC/RJ, em agosto de 1982, onde apresentou dois trabalhos originais sobre estudo de caso e sobre a abordagem qualitativa em avaliação, publicados, posteriormente, em *Educação e Seleção* nº 7, jan-jun, 1983. Outro, é a elaboração de uma dissertação de mestrado, por Tânia de Martino Salim, defendida em fevereiro de 1984, na PUC/RJ, sob o título "Alfabetização: ponto de partida ou ponto final?". Nesse estudo, a autora procurou observar cuidadosamente as exigências de um estudo de caso na busca da construção de um quadro geral dos problemas que cercam a alfabetização numa escola de periferia. Ainda outro evento merece registro a respeito: a realização de um *Simpósio sobre metodologia qualitativa em pesquisa educacional*, no *II Seminário Regional de Pesquisa em Educação da Região Sudeste*, em Belo Horizonte, em outubro de 1983. Os artigos ali apresentados serão publicados, proximamente, nos *Cadernos de Pesquisa*, da Fundação Carlos Chagas.

soluções quantitativas usualmente aceitas e utilizadas, liberando uma quantidade considerável de energia criativa, que estava bloqueada pela imposição do velho "paradigma da botânica e agricultura", como diz uma expressiva metáfora, que caracteriza bem um certo período. Permitindo-se buscar novas respostas, vindo de uma variedade de fontes, os avaliadores podem, agora, usar, livremente, sua imaginação para descobrir as melhores soluções para os problemas que enfrentam. Essas soluções podem vir de uma interpretação pessoal, baseada num estudo de caso ou numa crítica educacional, assim como de uma análise estatística estrita, se isso for mais apropriado. Elas podem vir, aliás, de uma combinação desses dois tipos de abordagens, se o problema em questão o exigir.

Essa liberdade de movimentos para o avaliador, essa liberação de limites estritos estabelecidos muito tempo atrás, essa valiosa possibilidade de inovação representa talvez o evento mais importante da história recente da avaliação educacional, como ficou tão bem assinalado, há já algum tempo, no excelente. "*Beyond the numbers game*", de Hamilton *et alii* (1977). Que isso não sirva, entretanto, para uma falsa radicalização, numa posição extrema de escolha entre métodos qualitativos ou quantitativos, como se eles fossem necessariamente opostos e mutuamente exclusivos. A liberdade do avaliador consiste justamente em poder escolher as estratégias que quiser, que achar mais apropriadas ao seu objeto e propósito de estudo, sempre atento às exigências de rigor científico, que devem presidir seu trabalho de pesquisa e que são necessariamente diferentes, de acordo com as diferentes abordagens.

6 – O COMPONENTE POLÍTICO

Praticamente ignorado até há bem pouco tempo, esse assunto começa, gradativamente, a assumir seu papel no cenário educacional dos Estados Unidos. É claro que esse aspecto esteve sempre presente, mas de certa forma seu papel era subestimado, talvez na crença errônea de que bastavam os componentes teóricos e técnicos para que o esquema avaliativo se desenvolvesse. A prática da avaliação tem gradualmente mostrado, entretanto, a importância do aspecto político, e muitos educadores têm experimentado o gosto amargo de verem um projeto bem pensado e bem planejado, do ponto de vista técnico e teórico, fracassar inteiramente pela falta de um bom embasamento político.

A consciência política não tem sido particularmente forte entre os educadores americanos, e os avaliadores, infelizmente, não representam exceção. Há uma grande necessidade de despertar essa consciência e, também, de afastar uma certa resistência existente entre os profissionais contra o tratamento político de fenômenos educacionais em geral e da avaliação em particular. Há quase uma espécie de pudor, ao encarar esse tipo de problema, que faz com que o lado político fique inteiramente na sombra, mesmo em cursos de avaliação de ótimo nível. Assim, os estudantes perdem boas oportunidades para desenvolverem sua atenção e senso crítico a respeito desse aspecto tão importante. Minha impressão, baseada em observações pessoais, é de que, se a função política da educação nos círculos acadêmicos americanos é muito pouco estendida, e ainda menos questionada, o que pode ser esperado com relação às funções da avaliação?

Sem dúvida, esse parece-me ser um aspecto bem mais desenvolvido entre educadores brasileiros. Acho que ainda falta muito para se esclarecer plenamente as principais questões envolvidas, mas já se registra, indubitavelmente, entre nós, a consciência viva da conotação política da avaliação. Há tudo para ser feito e discutido a partir desse ponto, o que não farei aqui, mas já está bem distante o estágio da crença nas soluções puramente educativas, baseadas em decisões apoiadas pela neutralidade científica. As questões de avaliação requerem, também, para sua correta colocação, de conceitos de poder e de controle, que ultrapassam a esfera estritamente educacional.

7 – O CARÁTER JULGADOR

Este é um aspecto negativo, mas inevitável para a avaliação. A própria natureza da avaliação

implica comparar o objeto avaliado a algum critério, pré-explicitado ou não. A operação resultará na classificação do item em estudo como tendo ou não valor. Assim, naturalmente, a avaliação precisa assumir a espinhosa tarefa de julgar, e é ilusório pensar que se pode esconder essa função por detrás de uma neutralidade já desmascarada. Mas este, a meu ver, é apenas o ponto inicial da avaliação educacional, que, se ficar por aí, deixará de prestar sua melhor contribuição. Após julgar, segue-se muito importante função de auxiliar professores e estudantes a encontrarem os melhores caminhos para seu próprio aperfeiçoamento.

Ajudando as pessoas envolvidas a conhecerem melhor a situação geral, em que se localizam os problemas em discussão, o trabalho do avaliador vai permitir que, a partir de uma visão concreta da realidade, os próprios participantes do processo procurem suas melhores soluções. Essa é a verdadeira função da avaliação:— ajudar os educadores a melhorarem seu trabalho, como tão bem expressou Cronbach, e não ameaçar ou amedrotar pessoas já sobrecarregadas com uma grande carga de trabalho. Reconheço que não é fácil resgatar essa função da avaliação, frente às outras tão mais evidenciadas e sofridas, tais como as funções de julgamento e de controle, representadas pela classificação simplista de estudantes, professores, programas etc., em categorias dicotômicas e definitivas, do tipo bons/maus, aprovados/reprovados, promovidos/não promovidos.

Nesse contexto, é muito importante mostrar as outras funções que a avaliação pode e deve desempenhar, como pode ser usada para promover melhores programas, atingindo um maior número de estudantes, com utilização mais racional de recursos. Dentro das circunstâncias de escassez em que se encontra, usualmente, a educação em nosso país, é importante estar informado e capacitado para aplicar qualquer recurso disponível para ajudar a resolver nossos problemas educacionais. Se simplesmente rotularmos a avaliação como instrumento de controle, sistematicamente ignorando suas possíveis contribuições, estaremos nos arriscando a perder o melhor uso dos recursos disponíveis, tarefa que a avaliação está precisamente apta a cumprir.

8. REFERÊNCIAS BIBLIOGRÁFICAS

- ANDRÉ, M. (1978). A abordagem etnográfica: uma nova perspectiva na avaliação educacional. *Tecnologia Educacional* (24), set/out.
- CONBRACH, L.G. (1963). Course improvement through evaluation, *Teacher's College Record*, 64, 672-683.
- EISNER, E.W. (1979). *The Educational Imagination*, Macmillan Publ. Co., New York.
- HAMILTON, D. et alii. (1977). *Beyond the Numbers Game*, McCutchan Publ. Co., Berkeley.
- LÜDKE, H. (1975). O Ginásio Pluricurricular e a Reforma do Ensino: Um Estudo Avaliativo. *Pesquisa e Planejamento*, nº 17, maio.
- RIST, R. (1980). Ver especialmente *On Qualitative Research. Issues of Design, Implementation Analysis Ethics, and Applicability*. A bibliography prepared by Ray Rist. College of Human Ecology, Cornell University.
- SCRIVEN, M. (1977). Dentre vários artigos do autor sobre o tema pode-se ver "Goal Free Evaluation". In: Hamilton, David et alii. (1977). *Beyond the Numbers Game*, McCutchan Publ. Co., Berkeley.
- SCRIVEN, M. (1980). *Evaluation Thesaurus*. Edgepress, Inverness.
- SMITH, N. (1978). *The development of new evaluation methodologies*. Research on Evaluation Program, Northwestern Regional Educational Laboratory, Portland, Oregon.