

Referenciais teóricos e metodológicos da avaliação

Nos estudos sobre avaliação da educação no ensino básico constata-se a predominância de artigos que dizem respeito a discussões de idéias, teorias ou metodologias acerca da avaliação da aprendizagem e suas tendências contemporâneas. Nessa categoria enquadram-se textos que abordam aspectos da avaliação sob o prisma conceitual, histórico, filosófico, sociopolítico, psicopedagógico e educacional propriamente dito, discutindo seus pressupostos, finalidades, teorias, modelos e até mesmo dimensões técnico-metodológicas dos estudos na área. Essa discussão é geralmente feita em textos ensaísticos, não tendo sido considerados, na categoria, os estudos empíricos que eventualmente pudessem dar suporte a alguns deles.

Tais artigos constituem a categoria mais numerosa dos textos sobre a avaliação no ensino básico, reunindo 70 títulos, o que indica haver na tradição brasileira uma preocupação evidente dos estudiosos com o esclarecimento dos significados da avaliação educacional no nosso contexto. A proporção em que esses textos aparecem no conjunto sugere também que mais se discursa sobre a avaliação no País do que propriamente se faz intervenções intencionais na área, como já haviam apontado Candau e Oswald (1995) em relação aos anos 80.

Nesta categoria, bastante heterogênea por sinal, podem ser encontrados vários subgrupos de estudos, de acordo com a abordagem adotada. Aí foram enquadrados artigos realizados na perspectiva dos estados do conhecimento, que revisam a bibliografia proveniente de diferentes fontes e períodos com o fito de analisar tendências e identificar lacunas da produção brasileira na área. Outros estudos procuram mapear, numa perspectiva histórica, as mudanças pelas quais têm passado as variadas concepções de avaliação.

Há textos que se dedicam aos aspectos filosóficos, pedagógicos, sociais, ideológicos e de poder envolvidos na avaliação da educação e na sua relação com a sociedade mais ampla, muitas vezes utilizando argumentos de várias vertentes. Aí podem ser incluídos os numerosos artigos que se propõem elucidar conceitos e pressupostos dos modelos de avaliação mais difundidos nas escolas ou na literatura brasileira sobre o assunto. Outros tratam de questões ligadas à abordagem das práticas avaliativas no contexto da organização da escola, ocupando-se de contribuições de caráter predominantemente teórico ao enfrentamento dos problemas suscitados pelo fracasso escolar.

Há também um subgrupo constituído de artigos que tratam da avaliação escolar tendo em conta a contribuição da Psicologia para o seu entendimento.

Outro subgrupo detém-se na explicitação de modelos de avaliação em larga escala, contendo artigos que propõem abordagens e estudos que permitem traçar a trajetória escolar de diferentes coortes de alunos e o seu perfil cognitivo. São igualmente mencionados modelos alternativos de monitoramento das redes de ensino utilizados fora do País, assim como é discutido o modelo de avaliação de resultados de aprendizagem, que configura o que chamamos, neste estudo, de avaliação de monitoramento, introduzido no Brasil com a montagem do Sistema Nacional de Avaliação da Educação Básica (Saeb) e de seus congêneres nos estados. Dentro da mesma matriz

positivista, encontra-se ainda a explicitação dos pressupostos básicos da Teoria da Resposta ao Item, que procura superar pretensas limitações dos paradigmas de avaliação de aprendizagem da Teoria Clássica de Medida e que passou a informar a metodologia adotada pelos sistemas de avaliação do ensino básico no País. Foram também arrolados nesta subcategoria os artigos que se ocupam do exame dos aspectos técnicos da construção de instrumentos de avaliação.

OS ESTADOS DO CONHECIMENTO E A RECUPERAÇÃO HISTÓRICA DAS CONCEPÇÕES DE AVALIAÇÃO

Cinco artigos reportam-se a estudos cujo objetivo foi o de fazer um balanço da produção acadêmica sobre a avaliação educacional no País na perspectiva dos estados da arte. Três deles são da autoria de Sandra Maria Zákia Sousa, pesquisadora que concentra suas investigações sobre a temática, prestando contribuição original ao conhecimento do que tem sido pesquisado entre nós e às condições em que têm sido produzidas e divulgadas tais pesquisas.

À maneira de vários outros autores, Sousa (1994) procura traçar a trajetória dos estudos sobre avaliação no Brasil. Em *Ensaio*, publica uma análise de pesquisas realizadas no âmbito da pós-graduação em educação entre 1980 e 1990, cujo propósito foi o de conhecer a destinação dada ao produzido nessas investigações sobre a avaliação educacional. Para tanto, colheu depoimento dos pesquisadores versando sobre suas trajetórias na área da avaliação da aprendizagem, a divulgação das pesquisas realizadas e o intercâmbio entre eles.

Entre os 35 pesquisadores identificados, apenas 14 responderam aos questionários, mas, ainda que o retorno tenha sido restrito, os achados do estudo são bastante sugestivos. A maioria dos autores não havia realizado anteriormente pesquisa sistemática sobre o tema, e a tendência era a de não darem prosseguimento às investigações, visto que grande parte deles concluiu tese ou dissertação no final da carreira e se aposentou. A divulgação dos trabalhos se fez de maneira esporádica e predominantemente para professores de escolas públicas de primeiro e segundo graus, indicando que a avaliação da aprendizagem tende a ser mais intensamente problematizada pela escola básica do que pela universidade. O espaço da disseminação por intermédio de palestras, conferências e cursos de curta duração tem sido sobretudo o do município ou do estado, sendo muito poucos os autores que publicam os resultados de seus trabalhos e, em número mais restrito ainda, os que o fazem em veículos com circulação em âmbito nacional ou que chegam a divulgá-los fora do País. A interlocução entre os pesquisadores do tema mostrou-se precária, tendo ocorrido, em geral, de forma esporádica e mediante contatos informais.

Em face do questionamento suscitado pelo conhecimento produzido nessas circunstâncias, Sousa faz uma série de indicações sobre a necessidade de criação de mecanismos de intercâmbio entre os pesquisadores e de desenvolvimento coletivo das investigações, assim como sugere a manutenção de canais de apoio à sua disseminação, uma vez que a difusão do conhecimento é componente da relevância que ele pode ter.

O trabalho de Sandra Sousa (1995a) que abrange período mais amplo de tempo e maior diversidade de fontes – publicado em *Cadernos de Pesquisa* – tenta rastrear a presença da avaliação no conjunto das pesquisas realizadas no País a partir do final da década de 30, com a criação do Instituto Nacional de Estudos Pedagógicos, que, desde 1972, passou a ter a atual denominação – Instituto Nacional de Estudos e Pesquisas Educacionais –, mas conservando a sigla original (Inep).

Sousa demarca a influência da Psicologia na avaliação educacional até os anos 50. Segundo ela, a análise da problemática educacional no período é feita na perspectiva individual, e as diferenças de desempenho são explicadas no plano biopsicológico. A avaliação da aprendizagem é entendida como mensuração de capacidades e características individuais por meio de testes. A influência do pensamento americano é notória, apresentando entretanto uma defasagem de cerca de vinte anos em termos das orientações correntes nos Estados Unidos.

Em meados da década de 50, com a criação dos Centros Regionais de Pesquisas Educacionais, desloca-se o foco das pesquisas para o estudo das relações entre o sistema escolar e a sociedade. O acento é colocado na determinação social da educação, sendo que a avaliação da aprendizagem não é temática privilegiada nas investigações. Desse modo não há proposições que se contraponham à avaliação no sentido de verificação do desempenho escolar.

Na metade da década seguinte, sob a influência da teoria do capital humano e do tecnicismo, passa-se a afirmar que a educação vai mal porque mal administrada, o que justifica a mudança de foco para o planejamento voltado para uma racionalização do trabalho capaz de enfrentar a ineficiência e ineficácia do sistema educacional. A orientação das pesquisas em geral dirige-se a uma visão interna da escola. A avaliação por objetivos, nos moldes preconizados por Tyler, marca a década de 70, que também absorve a contribuição de vários outros autores americanos sobre a avaliação da aprendizagem. A concepção da avaliação como mensuração dá lugar à dimensão tecnológica da avaliação, que coloca ênfase no seu caráter cientificista e nos métodos e procedimentos operacionais.

Nos anos 80 as teorias crítico-reprodutivistas ampliam a compreensão do fenômeno educacional, recuperando a sua dimensão social que extrapola inclusive a própria escola. A visão persistentemente esquemática e fortemente ideologizada da realidade que decorre desse paradigma teórico faz emergir, em contrapartida, estudos que destacam o caráter contraditório do Estado e dão ênfase ao exame das variáveis internas da escola. Agora a perspectiva é explicitamente assumida como política e expressamente comprometida com os segmentos majoritários da população.

Conforme a autora, prevalece, a despeito da tendência dominante, a concepção tecnicista no trato das questões da avaliação. A cultura avaliativa dominante não se realiza porém sem conflitos, e é possível identificar um movimento na direção da construção de referenciais que contextualizem a avaliação no sistema escolar e social, sobretudo a partir da segunda metade da década. A análise da avaliação em diferentes ambientes escolares dá pistas para a compreensão de suas finalidades mais amplas; estudos sobre intervenções na escola propõem transformações das práticas vigentes, cujo eixo é a reflexão sobre seus princípios. Ganha sentido a compreensão da realidade escolar com o fito de subsidiar decisões, que coloca como desafio a construção de uma sistemática da avaliação da escola como um todo.

Retomando o estudo das teses e dissertações realizadas entre 1980 e 1990, Sandra Sousa (1996) propõe-se fazer a sua caracterização e análise. Em primeiro lugar constata que a temática continua não sendo privilegiada, uma vez que os 37 estudos identificados entre os 40 indexados representam apenas 1,3% da produção da área da educação no período. Constata também que os estudos sobre avaliação foram realizados em apenas 13 das 34 universidades do País, estando concentrados particularmente em quatro delas: Pontifícia Universidade Católica de São Paulo, Pontifícia Universidade Católica do Rio de Janeiro, Universidade Federal do Rio de Janeiro e Universidade Federal do Espírito Santo.

Não foram encontradas evidências de articulação em torno de princípios ou eixos temáticos que revelassem movimento de continuidade ou aprofundamento entre as pesquisas, sendo que, na maioria dos trabalhos, não se verificou uma posição explícita de diálogo com as investigações já realizadas até mesmo na própria instituição, seja para corroborar conclusões, seja para opor-se a elas. Tais constatações vão ao encontro de afirmações de estudiosos que se referem à ausência, no âmbito dos programas de pós-graduação, de princípios norteadores e articuladores das pesquisas e mesmo de mecanismos que estimulem e garantam a interlocução de pesquisadores dentro das próprias instituições. Possíveis convergências nos estudos seriam reflexos de movimentos mais amplos, direcionadores de determinadas abordagens teórico-metodológicas da investigação científica, muito mais do que fruto de um empenho no interior das instituições no sentido da consolidação de linhas ou grupos de pesquisa em torno de um campo temático. A despeito das limitações apontadas, as universidades onde há maior concentração de trabalhos são consideradas pólos virtuais para o florescimento e consolidação da investigação na temática.

A avaliação no ensino de primeiro grau constitui objeto do maior número dos estudos realizados nesse período. Quanto ao enfoque, os textos detêm-se sobretudo na caracterização e análise das práticas avaliativas na escola, mas podem ocupar-se da elaboração ou testagem de instrumentos ou propostas de avaliação ou, ainda, da análise da literatura ou legislação sobre o tema. Aproximadamente 50% dos estudos adotam uma concepção eminentemente técnica da avaliação, proveniente da matriz positivista, mas por volta de 40% das pesquisas restantes problematizam essa visão, recorrendo predominantemente, para tanto, às teorias crítico-reprodutivistas. Nesses casos costumam, contudo, dar maior espaço às denúncias sobre as formas escolares de reprodução das desigualdades sociais do que propriamente às contradições que explorem os elos de sua transformação.

De qualquer modo, Sousa admite que os estudos da década prestam uma contribuição relevante à compreensão da maneira como vem sendo realizada a avaliação nas escolas e apontam a necessidade de consolidar outras formas de vivenciá-la, incorporando a sua dimensão política. Deixam clara também a necessidade de romper com o paradigma classificatório em favor de uma avaliação de caráter diagnóstico ou da investigação do processo educacional, adotando como alvo a organização escolar e não apenas o aluno, pois é ela, com sua dinâmica, que produz o fracasso escolar. Há também indicações sobre a necessidade de rever os aspectos normativos vigentes sobre a avaliação. Nesses termos, observa-se uma mudança na tônica da literatura sobre o tema, que deixa de ser prescritiva para assumir um caráter mais descritivo e interpretativo.

Ainda na perspectiva dos estados do conhecimento, Vera Candau e Maria Luiza Oswald (1995) examinam a produção sobre a avaliação no ensino de primeiro grau, publicada em artigos de onze revistas científicas de âmbito nacional entre 1980 e 1992. A revisão bibliográfica realizada por essas autoras e as referidas pesquisas de Sousa terminam por oferecer uma visão bastante representativa da produção científica no período e se complementam reciprocamente. Com especial atenção à avaliação da ação pedagógica, Candau e Oswald chegam a conclusões interessantes e que vão, por várias vezes, na mesma direção das provenientes de outros estudos.

Em primeiro lugar constata-se a prevalência significativa de trabalhos que discutem idéias sobre os que apresentam pesquisas, o que pode ser considerado um indicador da lacuna no campo dos estudos empíricos. Constatam ainda a falta de divulgação de trabalhos de qualidade na área.

Apesar de reconhecerem a importância dos estudos quantitativos, chegam à conclusão de que a maioria deles avalia programas de ensino e não as práticas pedagógicas. Estas são objeto de estudos qualitativos e aparecem em pequena proporção no conjunto dos textos.

A bibliografia sobre o tema mostra íntima relação entre a compreensão da educação e seu papel social e o enfoque dado aos processos de avaliação. As influências mais freqüentes são a tecnicista – que tem crescido mais do que outras – e a dialético-transformadora. A convivência, no mesmo período, de tendências pedagógicas diversas tem conduzido também a práticas de avaliação com diferentes orientações pedagógicas.

A ênfase no tratamento global da avaliação, unida à análise do seu papel ideológico e de controle social, tem sido feita mediante discurso de tom predominantemente genérico e com o caráter de denúncia, sem que sejam apontados caminhos alternativos. A consciência do que a avaliação não deve ser não é acompanhada de um esforço sistemático de construção de metodologia e instrumental que permitam o desenvolvimento de outra maneira de avaliar. A presença pouco expressiva de teorias específicas, modelos, métodos e técnicas indica lacuna importante na produção de conhecimento no País; falta, em particular, instrumental que permita tratar aspectos técnico-pedagógicos do processo de avaliação sob a postura crítico-transformadora.

Algumas pesquisas que abordam a relação fracasso/sucesso escolar oferecem elementos importantes para retrabalhar tais questões, especialmente em relação ao ciclo básico de alfabetização.

Poucos artigos articulam avaliação com desempenho e formação docente ou focalizam a avaliação do docente. A formação do professor sobre o tema é insuficiente e faltam-lhe critérios claros para orientar o processo de avaliação e permitir a escolha dos instrumentos mais adequados.

A insuficiência de conhecimento sobre os conteúdos curriculares por parte do docente incide negativamente sobre o manejo da avaliação: não se ensina o que não se sabe.

Heraldo Marelim Vianna (1992c) oferece outra contribuição aos estudos realizados na linha do levantamento e análise sistemática da produção sobre a avaliação, fazendo um balanço das publicações sobre o tema nos vinte anos de *Cadernos de Pesquisa*: de 1971 a 1991.

Constata a carência de uma teoria geral de avaliação entre os textos e identifica e comenta o tratamento das seguintes dimensões abordadas nos artigos: a centrada no estudante, a que focaliza o professor, a que tem em mira o material instrucional ou os cursos e programas, ou, ainda, a centrada na instituição ou no próprio sistema educacional. A análise dos trabalhos publicados mostra que a temática é muito diversificada, expressa grande preocupação metodológica e trata os assuntos com acentuado rigor científico, tendo sido observada a preocupação com a precisão dos conceitos envolvidos em várias das dimensões contempladas.

De início, os artigos concentram-se na questão da eficácia (racionalidade das decisões) e da eficiência (racionalidade da execução) e contam com a contribuição de Maria Amélia Azevedo Goldberg. Outro aspecto destacado é o das metodologias qualitativas, cujos fundamentos teóricos são abordados por André, Lüdke e Thiollent, e que são ilustradas por vários estudos de caso nem sempre bem fundamentados. Observa-se também a ampliação do escopo da avaliação, que não se restringe apenas ao rendimento escolar, passando a abranger a dimensão da escola. Esse processo exige naturalmente a utilização de métodos qualitativos e quantitativos, o que requer modelos complexos de análise. O autor registra ainda a discussão sobre a avaliação participante na revista, mas constata que a avaliação por procedimentos que adotem esse enfoque não figura entre os artigos.

O texto conclui pela necessidade de proceder a uma avaliação da avaliação, reconhecendo que ainda não possuímos uma cultura avaliativa, mas que, por pressão das autoridades e das agências financiadoras, ela vem sendo colocada em primeiro plano. Reconhece que faltam elementos qualificados para o exercício da avaliação, sendo que, para tanto, não basta transpor modelos de outros contextos, e admite que começa a surgir um esforço da prática da avaliação aplicada à realidade brasileira.

Ao lado dos estudos na linha dos estados do conhecimento, há uma série de outros artigos que se dedicam a fazer uma retrospectiva das concepções de avaliação que influenciaram o ideário brasileiro, contribuindo para que se tenha uma visão mais abrangente do campo. Nessa linha foram arrolados oito textos, sendo que não foram considerados nesse caso os artigos que fazem uma retrospectiva muito ligeira do tratamento dado à avaliação para, depois, entrarem no argumento principal. Combinando os dois tipos de análise, é possível obter um panorama acerca da produção sobre o tema, bem como das suas transformações ao longo do tempo. É freqüente que as tentativas de reconstituição da trajetória da avaliação educacional venham acompanhadas de um esforço de contextualização que busca explicitar as orientações teóricas, adotadas no exterior, que têm constituído as matrizes da sua formulação.

O autor que mais se demora na retomada histórica da avaliação é Heraldo Marelim Vianna, que tem diversos textos arrolados neste subgrupo, todos publicados em *Estudos em Avaliação Educacional*. Por força de tanto escrever sobre o tema, repete por vezes certas informações ou análises, mas, de modo geral, acrescenta dados e reflexões importantes para contextualizar nossa tradição de avaliação, seja em relação a tendências internacionais, seja com referência à evolução do tratamento da questão no País.

Vianna (1992d) registra vários projetos de avaliação realizados no Brasil entre as décadas de 60 e 90 e discute questões relacionadas à avaliação da eficiência de professores; à avaliação de currículos, sistemas e programas; à falta de continuidade dos trabalhos realizados; à validade dos instrumentos utilizados; ao fato de não haver um sistema adequado de disseminação das informações obtidas pela avaliação para os professores, levando a questionar o impacto desses estudos sobre a prática docente. Os diversos problemas são levantados pelo autor para evidenciar

que somente a pesquisa sobre a avaliação e a prática constante desta serão capazes de criar uma cultura avaliativa e dar-lhe credibilidade no contexto das atividades educacionais, em benefício de um ensino mais eficiente.

Em 1993, faz uma análise do processo de avaliação e das práticas que vêm sendo utilizadas nas escolas, remarcando que são eminentemente voltadas para os resultados finais, quando o desejável seria que a avaliação se revestisse do caráter formativo, oferecendo oportunidade de discutir os erros substantivos e formular experiências com aquisição de novas aprendizagens. A seguir, menciona propostas de mudanças radicais baseadas em Wood: a avaliação do desempenho do aluno deve ocorrer em relação a ele próprio; deve evitar incidir sobre construtos psicológicos; não precisa ocorrer sob condições controladas e deve ter um sentido construtivo. Ou seja, o modelo propõe a eliminação da comparabilidade entre os alunos, o abandono de uma visão apenas psicométrica, a adoção de uma avaliação individualizada com o emprego de vários recursos e a utilização de um processo interativo aluno/professor (Vianna, 1993b).

No texto de 1995, Vianna apresenta uma alentada síntese histórica da evolução da avaliação educacional, que teve seu crescimento a partir dos anos 60. Inicialmente recupera essa evolução a partir dos primórdios da tradição avaliativa nos Estados Unidos, há quase dois séculos, tradição essa em cujo contexto se desenvolveram numerosos e importantes trabalhos em decorrência da disseminação da influência de Ralph Tyler. Nos anos 60, a avaliação nesse país começou a incidir sobre grandes projetos de currículo financiados com apoio federal, surgindo a figura do avaliador. Os resultados, ainda que realizados com grande rigor técnico, não se revelaram contudo plenamente satisfatórios. No período de Kennedy e posteriormente, aliado à questão das diferenças de oportunidades educacionais para diferentes segmentos da população, surge o conceito de responsabilidade em educação (*accountability*), que revela a preocupação de evitar desperdícios financeiros nos programas educacionais, desencadeados numa perspectiva compensatória. O dispositivo legal de 1965 que obrigava a avaliação de todos os projetos financiados pelo Poder Público para o ensino primário e médio levou a grandes esforços nessa linha. A prática generalizada da avaliação mostrou entretanto que muitos instrumentos e estratégias não eram adequados. Stake, em 1967, e Stufflebeam, em 1971, darão nova dimensão metodológica à avaliação, que se torna um campo profissional definido a partir de 1970. O autor refere-se ainda à grande preocupação atual com a qualidade da avaliação, tendo em vista os padrões fixados pelo Joint Comitee of Standards for Educational Evaluation.

A seguir, menciona algumas experiências de avaliação na Inglaterra. Passa pela contribuição da psicometria (Galton, Pearson, Spearman, Burt) e do ideário econômico de Keynes; da associação do governo com a Fundação Nuffield para reformulação de currículos e formação de professores (1962); pela experiência da National Foundation for Educational Research (1994) e pelas diretrizes estabelecidas para os School Councils, que destacam a participação de professores e não especialistas na avaliação para a tomada de decisões. Chama a atenção para a tensão existente entre a avaliação de tipo individualizado, usualmente feita nas escolas, e a atual prática de avaliação de resultados da aprendizagem do sistema inglês, que entra em choque com a autonomia dos professores.

Depois, procura recuperar as atividades mais representativas de avaliação realizadas no Brasil no período de 1960 a 1995, enumerando-as e ponderando que elas são ainda bastante escassas.

Reportando-se a análises de Gatti a respeito da avaliação de projetos brasileiros, registra que: nem sempre é possível adotar modelo experimental que possibilite traçar diferenças de tratamento; as amostras muitas vezes ficam comprometidas pela carência de estatísticas básicas; os procedimentos utilizados nem sempre são fidedignos; os modelos de avaliação muitas vezes não se ajustam ao nosso contexto socioeducacional; faltam pessoas experientes na área.

Em artigo de 1997, Vianna, ao propor-se fazer uma análise crítica da avaliação, cujos produtos nem sempre são satisfatórios mesmo em países avançados, discute a relação entre a

avaliação e a pesquisa, reportando-se às características que as distinguem segundo Glass, Worthen e Sanders. Para tais autores, a motivação do pesquisador é a curiosidade científica, enquanto a do avaliador é a tomada de decisões, sendo que o grau de autonomia do primeiro é maior do que o do segundo; a pesquisa visa ao estabelecimento de conclusões, leis, princípios, à descrição de fenômenos de caráter mais universal, enquanto a avaliação se presta à tomada de decisões e à descrição de fenômenos mais restritos; o julgamento da pesquisa é pautado por critérios de validade interna e externa, enquanto na avaliação agrega-se o critério de utilidade e credibilidade.

Reproduz ainda, no artigo, breve caracterização de paradigmas epistemológicos dos quais derivam diferentes tipos de avaliação, mencionando Popper para o positivismo, Kuhn, cujas formulações permitem ir além do positivismo, Guba e Lincoln para o paradigma naturalista; destaca os modelos prescritivos e descritivos de avaliação e, finalmente, detém-se mais longamente em dois tipos de avaliação qualitativa: a responsiva de Stake e a iluminativa, proposta por Parlett e Hamilton.

Em 1998, Vianna (1998b) retoma as análises sobre o modelo de avaliação por objetivos de Tyler e as proposições críticas de Cronbach, Scriven, Stake e Parlett/Hamilton, todas voltadas para uma avaliação de caráter mais eminentemente qualitativo.

Ainda em 1998, em outro artigo e de forma sistematizada, Vianna (1998a) registra as reflexões acerca das teorias que influenciaram seu posicionamento em relação à prática da avaliação, sem desconsiderar sua própria experiência pessoal ao longo do período. Toma como referência as 95 teses apresentadas por Cronbach: um conjunto de proposições advogando uma nova visão de avaliação, especialmente em relação aos programas sociais.

Ainda na linha dos estudos retrospectivos, Clarilza Prado de Souza (1998) faz uma breve descrição de como tem vivenciado a história da avaliação educacional no Brasil nos últimos 30 anos, passando pela tradição positivista dos anos 60 e 70, pelas teorias críticas dos anos 80, que substituem a base psicologizante pelos referenciais da Filosofia e da Sociologia, até chegar a uma análise contextualizada do papel que a avaliação exerce no cotidiano escolar. Afirma que a década de 90 trouxe como desafio a necessidade de integrar as contribuições críticas das últimas décadas e construir instrumentos que permitam analisar o rendimento escolar e compreender os processos de produção das desigualdades sociais. Essa reorientação da avaliação encontra suporte nas tendências e estudos que procuram compreender as macroestruturas determinantes de uma realidade em conjugação com o desempenho dos atores sociais que a constroem. Nesse sentido, para a autora, é tão importante avaliar como o sistema educacional condiciona a qualidade do ensino oferecido nas escolas quanto analisar como os educadores constroem o nível de ensino que os alunos irão receber.

Depresbiteris (1998), fazendo uma revisão de como o tema da avaliação vem sendo discutido por diferentes autores desde as primeiras décadas do século, no exterior e no Brasil, procura uma interlocução com eles a fim de responder a perguntas como: Qual a diferença entre avaliar o ensino e avaliar a aprendizagem? E entre teste e prova? Que outros instrumentos existem para avaliar o desempenho dos alunos? Como se avaliam os resultados da avaliação da aprendizagem?

REFERENTES FILOSÓFICOS, SOCIAIS, POLÍTICOS E PEDAGÓGICOS

Este grupo de artigos, o mais numeroso da categoria, reúne 38 títulos, que constituem 54% desse conjunto, ou seja, pouco mais da metade, sendo que vários dos autores estão aí representados mais de uma vez. A escassez de teorias sobre avaliação propriamente dita não impede o esforço de elucidar pressupostos da Filosofia e das ciências aos quais se filiam concepções de educação – de que derivam, por sua vez, diferentes concepções de avaliação – ou, ainda, de explicitar as matrizes pedagógicas que subsidiam as práticas na área.

Trata-se, nesta subcategoria, de artigos majoritariamente preocupados com a dimensão qualitativa da avaliação, que abordam muitas vezes questões ligadas ao seu estatuto científico,

mas que se atêm, em variadas ocasiões, à mera reflexão sobre as práticas avaliativas. É freqüente que critiquem o paradigma positivista predominante na tradição de avaliação brasileira, sendo que, a partir de várias vertentes teóricas, nem sempre claramente explicitadas e nunca aprofundadas, acabam por esboçar características de um novo modelo de avaliação apresentado como desejável.

Tais características são bastante convergentes, indicando haver já um amplo consenso na área. Elas passam pela defesa de uma abordagem historicamente situada, que, em relação ao aluno, leve em conta não apenas a dimensão cognitiva, mas a social, afetiva, seus valores, motivações e, até mesmo, a sua própria história de vida. A busca da relação entre avaliação e qualidade do ensino faz recair a ênfase nas variáveis do processo, muito mais que no produto da educação, sendo que a sua natureza deve ser eminentemente dialógica e dialética, voltada para a transformação, tanto no plano pessoal como no social. As mudanças em relação ao indivíduo apontam na direção da autonomia e, em relação ao social, na direção de uma ordenação democrática e, portanto, mais justa da sociedade, que pressupõe, em alguns casos, uma orientação socialista não claramente expressa.

O eixo da avaliação deixa de girar exclusivamente em torno do aluno e da preocupação técnica e passa a centrar as atenções em torno das condições em que é oferecido o ensino, da formação do professor e de suas condições de trabalho, do currículo, da cultura e organização da escola e da postura de seus dirigentes e demais agentes educacionais. Essa mudança de enfoque vem possibilitando a paulatina superação da tradição crítico-reprodutivista. Muito tendo contribuído para elucidar os condicionantes estruturais da educação e o papel da avaliação na manutenção da ordem social vigente, essa tradição gerou, como já assinalaram estudos anteriores, uma produção sobre o tema com acentuada conotação de denúncia, mas pouco subsidiou a proposição de enfoques alternativos que se encaminhassem numa linha propositiva.

Além de textos que trazem o foco da avaliação para a escola e, em menor proporção, para o professor e outros agentes educacionais, foram incluídos nesta subcategoria alguns artigos que se reportam à avaliação institucional. Geralmente, a preocupação com a abordagem institucional tem sido mais diretamente associada, na literatura brasileira, às análises do ensino superior. Não obstante, como nos periódicos examinados alguns artigos referentes à avaliação institucional não faziam menção explícita a nível de ensino ou se referiam a vários deles, decidiu-se por mantê-los neste estudo. Assim como em relação ao aluno, à escola e aos agentes educacionais, a proposta de avaliação institucional enfatiza o diálogo, a participação, a reelaboração coletiva das propostas de mudança, o estabelecimento de relações menos fortemente hierarquizadas e mais democráticas entre as instâncias.

O paradigma emergente de avaliação qualitativa não tem uma densidade teórica própria; ele empresta elementos de várias vertentes de pensamento, constituindo-se uma formulação tipicamente interdisciplinar. Os textos que a ele se referem guardam, na maioria das vezes, tom prescritivo, sendo bem menos freqüentes aqueles que refletem sobre ações desenvolvidas sob essa orientação.

O excerto de alguns dos artigos incluídos nesse grupo pode oferecer uma idéia mais precisa das questões por eles discutidas.

Entre os poucos textos que apresentam uma intersecção explícita com a filosofia das ciências, destaque-se o artigo de Franco (1990). Nele é feita uma caracterização sumária das matrizes epistemológicas do modelo de ciência que privilegia o objeto (racionalismo, funcionalismo, positivismo) e do modelo que prioriza o sujeito (idealismo). Transpostos para a avaliação educacional, tais modelos teriam representado um avanço e nortearam a geração de paradigmas mais complexos e abrangentes. No entanto, como em ambos os casos a visão do indivíduo se apresenta de forma abstrata, a autora dá ênfase à necessidade de que a avaliação educacional tenha como unidade de análise o vínculo indivíduo/sociedade numa dimensão histórica. Daí ser fundamental para a avaliação da aprendizagem, em todos os níveis, o entendimento da atividade humana, da ação prática dos homens, o que pressupõe a análise do motivo e da finalidade dessa ação.

Em outro artigo, Franco (1994) aponta para a complexidade do conceito de qualidade do ensino, considerando-o como um produto histórico e social, orientado por diferentes expectativas que incorporam demandas diversificadas e mutáveis ao longo do tempo, e que reflete um posicionamento político e ideológico. Ao discutir critérios e avaliação dos indicadores de qualidade, questiona a sua validade, uma vez que são pensados em função do mito da modernidade, que os elege sobretudo pelas exigências da produção.

Luckesi (1996), em artigo de cunho eminentemente filosófico, ocupa-se da análise do juízo de valor, componente essencial do conceito de avaliação. Esclarece que o juízo de valor, como expressão da qualidade do objeto, pode ser emitido a partir das propriedades físicas desse objeto ou dos fins a que este se destina (valor de uso). Conclui que, nesse momento de grande influência das técnicas, os processos de avaliação não têm ido além do valor do objeto, deixando de contemplar o valor do fim a que servirá o produto obtido. Termina admitindo a necessidade de qualificar o objeto não em função dele mesmo, mas do fim a que se destina.

Um conjunto representativo de textos, ao acentuar os aspectos políticos e histórico-sociais da avaliação, apóia-se predominantemente nos referentes de Habermas ou, mais amplamente, da Escola de Frankfurt, para propor a avaliação do tipo emancipatório dentro de um paradigma dialético. Entre eles estão os escritos de Pedro Demo (1990 e 1995), Tigiboy (1995) e Granato (1995). Saul (1992), dentre outros, também adota conceitos afins, sem entretanto reportar-se explicitamente a essa matriz teórica.

Tendo em conta o fato de que os aspectos políticos da educação têm a ver com o exercício da cidadania, esses textos propõem uma estratégia de formação de sujeitos sociais, dotados de capacidade para se autodefinirem, bem como para determinarem o seu destino histórico, de modo tal que logrem conceber e construir sua emancipação política e econômica, assim como o seu projeto de desenvolvimento como povo.

Pedro Demo (1990), após admitir que a quantidade e a qualidade são partes integrantes e indissociáveis da educação, na tentativa de definir conceitos e critérios de avaliação da qualidade, distingue a qualidade formal (ligada ao domínio tecnológico) da qualidade política (voltada para a cidadania), afirmando que uma não pode ser entendida sem a outra nem, tampouco, substituída pela outra. A qualidade política, referindo-se a conteúdos históricos, é de teor prático e inevitavelmente ideológico e, portanto, de tessitura metodológica distinta daquela que a ciência clássica costuma abarcar em seus cânones. Toca a face social dos valores e compromissos, da consciência social crítica que é capaz de revelar sujeitos autônomos, que não são presa fácil de manipulações e que estão aptos a fazer opções quanto ao uso alternativo das instrumentações tecnológicas.

Para o autor, a conotação política da qualidade da escola está menos na sua característica pública do que no controle democrático sobre ela exercido pela comunidade e pela sociedade em geral.

Julga necessário também recompor a pesquisa como princípio educativo, uma vez que ela não implica a simples aplicação de métodos científicos de descoberta da realidade, mas o questionamento e a transformação dessa realidade. Como princípio educativo, o conhecimento (advindo da pesquisa) está na base do processo emancipatório, que sempre começa com a tomada de consciência crítica e a capacidade de dizer não, ato que inaugura o processo político questionador e que jamais se conclui. O confronto de idéias, o embate entre posições, o reconhecimento do conflito e a constatação da desigualdade são fundamentais para a organização política dos desiguais no sentido da emancipação. A importância da escola estaria na sua possibilidade de realizar a dupla face da pesquisa: apetrechar tecnicamente para fundar competência tecnológica e fazer pensar e se repensar na linha da transformação.

Em outro artigo, Demo (1995) busca caracterizar o compromisso educativo de todo o processo de avaliação escolar e a define como um processo permanente de acompanhamento do aluno. Utilizando como referência o paradigma da comunicação de Habermas, coloca a discutibilidade como critério essencial da cientificidade, que tem a ver com a avaliação, tanto no que diz respeito

aos procedimentos que fazem parte da sua lógica intrínseca, quanto no que se refere à apreciação da sua qualidade.

Granato (1995) propõe a realização do paradigma dialético no âmbito dos sistemas educacionais. Para que aí se estabeleça a avaliação emancipatória, é necessária uma dinâmica participativa, não só no sentido de associar procedimentos de avaliação externa com a auto-avaliação institucional, como no sentido de criar instâncias administrativas articuladas entre si para evitar a excessiva centralização e os programas prontos que descaracterizam o processo. Dentro do paradigma, é importante a valorização do espaço de tensão dialética entre os procedimentos propostos pelo sistema e as áreas de resistência que emergem das instituições, para que o agir comunicativo tenha condições de se efetivar.

Tigiboy (1995) analisa aspectos teórico-práticos de uma estratégia de mudança que combina elementos de uma avaliação dialética e de uma avaliação responsiva, que, na prática, tem se mostrado efetiva na formação de professores. A capacitação de docentes em serviço seria a base da decolagem para transformações do sistema de ensino tradicional, mediante a utilização de uma avaliação que desse ênfase ao diálogo para encontrar respostas meritórias e relevantes. A estratégia apoia-se nos diversos saberes, desde o saber teórico até o saber fazer, passando pelo saber ser, e é fundamentada na teoria do conhecimento de Habermas e na teoria da liberdade de Fromm. A abordagem baseia-se essencialmente na avaliação fundada nos paradigmas dialético e naturalista, bem como na avaliação responsiva, sugerida por Stake e por Guba e Lincoln. Interpretando algumas sugestões de Proppé sobre o paradigma dialético, afirma que a avaliação deve ser considerada um fenômeno político-social; que a ação é parte inerente da avaliação e que as ações da avaliação devem extrapolar o objeto ou fenômeno avaliado e seu contexto, ampliando a influência político-social para outras áreas. O paradigma naturalista, também chamado hermenêutico ou interpretativo, trabalha recorrendo às realidades múltiplas socialmente construídas.

Ana Maria Saul (1992) critica a cultura avaliativa brasileira, assinalando que esta sofreu intensa influência do pensamento positivista norte-americano. Tendo se tornado instrumento burocrático para os gestores dos sistemas de ensino e punitivo para os alunos, a avaliação se reveste de caráter profundamente autoritário no Brasil.

A autora propõe um paradigma alternativo – o da avaliação emancipatória, que se realiza a partir de três eixos: avaliação democrática, crítica institucional e construção coletiva. Sugere uma classificação política dos estudos avaliativos, visando à construção de uma avaliação democrática, cujos conceitos-chave são: sigilo, negociação e acessibilidade, ou seja, direito do usuário à informação e à utilização dos resultados da avaliação, para melhorar ou redirecionar as próprias atividades. A crítica institucional e a criação coletiva concretizam-se por meio de três momentos: expressão e descrição da realidade, isto é, elaboração de um diagnóstico institucional e descrição dos dados obtidos; crítica do material, expressa pela análise do projeto pedagógico da instituição; criação coletiva, ou seja, delineamento das novas ações da equipe de trabalho. Trata-se de abordagem que envolve pesquisa participante, a qual implica o compromisso do pesquisador com a causa em questão, contribuindo para uma ruptura metodológica profunda. O novo paradigma permitiria ao avaliador assumir o papel de coordenador dos trabalhos de avaliação, de um orientador dessas ações, favorecendo o diálogo, a discussão, a busca e a análise crítica sobre o funcionamento de um programa.

Uma série de outros artigos abordam a avaliação na perspectiva de suas implicações sociais mais amplas e das relações entre a educação e a sociedade abrangente, valendo-se, para tanto, do referencial crítico-reprodutivista, da pedagogia dos conteúdos, das reflexões de autores como Gimeno Sacristán e Philippe Perrenoud e de escritos que privilegiam a instituição escolar e sua cultura como foco de análise.

Marli André (1990) considera que a estrutura de poder vigente nas escolas reflete as relações de poder dos órgãos decisórios; é caracterizada por um poder predominantemente centralizador, exercido, via de regra, pelo diretor da unidade, e tende a reproduzir-se tal e qual na organização do ensino na sala de aula. Para que a prática docente possa ser redimensionada,

tornando-se mais justa e mais democrática, recomenda ela que se implante nas escolas uma sistemática de trabalho menos hierarquizada, mediante a criação de espaços coletivos de reflexão onde professores e orientadores possam analisar conjuntamente o fazer pedagógico, espaços esses que possam ser utilizados para acompanhamento, avaliação e reformulação do processo de ensino-aprendizagem, visando ao seu aprimoramento.

Essa mesma autora, em outro artigo (André, 1996), discute a questão do poder e da função social da avaliação escolar, ressaltando a necessidade de serem analisadas as mediações presentes no ato de avaliar. Toma como ponto de partida o conceito de fabricação da excelência escolar de Perrenoud, segundo o qual as normas e os critérios definidos pela escola como de excelência, a despeito de serem frutos de uma dada construção social, são difundidos como a única forma de conceber a realidade. Trabalhar no sentido de uma avaliação mais democrática implica trabalhar simultaneamente nos campos da Didática, da relação professor-aluno e da organização pedagógica da escola. Envolve também a adoção de uma pedagogia diferenciada que assegure a possibilidade de o aluno ser reconhecido pelo grupo, independentemente de suas competências escolares e de seu nível cultural. A diferenciação envolve o acompanhamento dos processos e dos caminhos de aprendizagem, a busca de estratégias para trabalhar com os alunos mais difíceis, sendo a avaliação formativa a sua principal fonte de informação. Com forte ênfase nos aspectos qualitativos da aprendizagem, o papel desse tipo de avaliação é o de contribuir para esboçar um plano de atuação.

Menga Lüdke (1991) propõe uma abordagem sociológica dos delicados mecanismos de controle exercidos pela avaliação em favor de uma ordem social conservadora, exemplificando-os com a seleção de alunos do segundo grau que irão assumir papéis de liderança no mundo econômico. Villas Boas (1993), nesse mesmo periódico, ao propor uma discussão da articulação da avaliação com a organização do trabalho pedagógico, também chega a conclusões que evidenciam a função seletiva da avaliação, conferindo privilégios a uns e excluindo outros da escola. Mediano (1990) faz indicações sobre a atuação do supervisor junto aos professores, tendo em conta o modo como é feita a avaliação na escola brasileira e as implicações sociais da avaliação.

Falcão Filho (1995) questiona o uso que se faz da avaliação realizada pelo professor como único parâmetro para medir a qualidade da aprendizagem. Esta pressupõe a avaliação do ensino, do professor e de suas condições de trabalho, da organização da escola e de seus dirigentes. Salienta a necessidade de avaliar o ensino ministrado pelo professor com a mesma intensidade com que se avalia a aprendizagem do aluno, pela constatação de que certas concepções de educação e de ensino influenciam negativamente os critérios de avaliação da aprendizagem e pelo questionamento que tem sido feito sobre as condições em que se dá a formação de professores nos cursos de pedagogia e licenciatura.

Em artigo dirigido ao administrador escolar, Lüdke (1995b) parte da discussão crítica dos modelos de avaliação comprometidos com a busca de instrumentos de medida "válidos" e "fidedignos", propugnando uma concepção de avaliação fortemente integrada ao próprio processo de ensino aprendizagem, como base para o desenvolvimento de propostas que venham ao encontro das necessidades de alunos e professores. Menciona formas e tentativas de avaliação associadas a esta última concepção. Refere-se ainda à necessidade de se exigir mais do que o simples reconhecimento e repetição de certos conteúdos por parte dos alunos e à maior confiança que deve ser atribuída à avaliação dos professores e dos próprios alunos. A despeito de admitir a maior carga de trabalho envolvida nesse tipo de avaliação, reconhece que ela constitui fator de crescimento que, contudo, não substitui a avaliação baseada em testagens que permitem a comparação de alunos e de instituições, permitindo a visualização de problemas não apreendidos no nível da sala de aula.

Ressalta ainda a importância dos estabelecimentos escolares como foco dos estudos sociológicos, assinalando a proposta de Nóvoa no sentido de que eles constituiriam o estudo do *meso*, ou seja, o nível intermediário entre as variáveis macro e micro. Destaca, da coletânea organizada pelo autor sobre o tema, a contribuição especial dos sociólogos Derouet, Briand e Chapoulie, bem como de Hutmaker, propondo aos diretores uma reflexão sobre a avaliação que leve em conta

que o que emana dos órgãos centrais não chega aos alunos senão pela via dos estabelecimentos escolares – tudo passa pelo crivo dos agentes das escolas, que ampliam, reduzem ou alteram o escopo das mudanças.

Bordignon (1995) realiza uma interessante discussão na perspectiva da avaliação institucional. O autor afirma que a avaliação é um processo de autoconsciência institucional, tornando-se o instrumento analítico que desvenda as causas condicionantes ou determinantes do sucesso institucional, e situa a análise nos meios, métodos e processos para além da mera medição e comparação de resultados. Assim entendida, a avaliação se constitui o processo de sistematização de informações criticadas, assumindo funções diagnósticas – pré-requisito fundamental para o planejamento e a gestão de organizações educacionais.

Um grupo de autoras, a saber, Mere Abramowicz, (1995a e 1995b), Vera Tavares de Souza (1994), Mírian Grinspun (1994 e 1996) e Vera Werneck (1996), propõe uma avaliação holística do aluno, que não se restrinja aos aspectos cognitivos, procurando contemplar sua condição existencial, sentimentos, motivações, valores. Ela deve manter o caráter dinâmico e interativo do processo. Sob esses pressupostos, Abramowicz faz uma tentativa de construção de um novo paradigma de avaliação da aprendizagem, que denomina crítico-humanista, e que implica estudo, pesquisa e compreensão do fenômeno.

Luckesi (1990a, 1991 e 1992), que muito tem contribuído para a elucidação dos pressupostos das práticas escolares de avaliação no ensino básico, comparece, ainda nesta subcategoria, com mais três artigos publicados. No texto de 1990, aborda a prática de aferição do aproveitamento escolar tendo como matriz os conceitos de verificação e de avaliação. Conclui que a aferição é realizada nas escolas sobretudo como verificação e utilizada para classificar os alunos como aprovados ou reprovados. Propõe que o objetivo primeiro da aferição não seja aprovar ou reprovar o aluno, mas direcionar a aprendizagem e promover o desenvolvimento discente. No artigo de 1991, tece considerações sobre a prática da avaliação nas escolas brasileiras, mostrando que esta, ao conquistar espaço tão amplo nos processos de ensino, condicionou a prática pedagógica ao que ele chama de pedagogia do exame. Na publicação de 1992, discute a relação entre planejamento, avaliação e projeto pedagógico da escola, insistindo na dimensão política que possuem essas atividades e preconizando um trabalho de equipe em que todos decidam o que e como fazer para elaborar um projeto pedagógico coerente com as necessidades da clientela. A avaliação nesse contexto seria apenas uma etapa do percurso, contribuindo para detectar problemas e corrigir rumos.

Alguns autores mencionam mitos que podem ser criados em torno da avaliação e chamam a atenção para procedimentos mistificadores a ela associados. Thereza Penna Firme (1994a) classifica as abordagens qualitativas em três categorias: as pseudo-avaliações, que falsificam informações e procuram favorecer pretensões tendenciosas; as que se assemelham à pesquisa, uma vez que priorizam a verificação de hipóteses, a experimentação e a constatação rigorosa de resultados em vez de julgamentos de valor (quase avaliações); e, finalmente, as que se preocupam com a formulação de juízos de valor, sejam eles de mérito ou de relevância (verdadeira avaliação).

José de Anchieta E. Barreto (1993) trata dos critérios de validade e precisão e das dificuldades que podem comprometer um programa de avaliação, classificando-as em mitos e armadilhas. Partindo do pressuposto de que a avaliação dos sistemas escolares é um valor proclamado na comunidade educacional, de que a avaliação é basicamente o julgamento do valor de uma ação e de que ela não pode ser confundida com o ato ou processo de medir um atributo, o autor discute dois mitos: o da facilidade da avaliação, com o conseqüente descaso na formação de profissionais a respeito do tema, e o da dificuldade, que se constitui um mecanismo de defesa dos que temem a avaliação pelos mais variados motivos. Enquanto os mitos são idéias falsas que orientam o comportamento humano, as armadilhas funcionam como artifícios capazes de comprometer definitivamente o esforço realizado. Destaca as armadilhas referentes aos instrumentos utilizados, à reação individual das pessoas que se sentem ameaçadas, e a junção das armadilhas individuais, que formam as armadilhas corporativas.

Abramowicz (1994), buscando entender o porquê da baixa utilização dos resultados de avaliação na formulação de diretrizes de políticas públicas, menciona a falta de comunicação entre quem avalia e quem decide sobre as políticas, a concepção inadequada dos usos da avaliação e a chamada pseudo-avaliação, entre outros fatores.

Outros artigos enquadrados nesta categoria são os seguintes: Hoffmann (1991), Maia (1995), Raphael (1995), Depresbiteris e Taurino (1996), Depresbiteris (1990), Hoffmann (1994), Denize Oliveira (1990 e 1991) e Zambelli (1997).

REFERENCIAIS PSICOPEDAGÓGICOS DA AVALIAÇÃO

Tratados em separado por apresentarem contornos mais claramente delimitados do que os textos que se reportam a outros referenciais, os artigos com enfoque psicopedagógico sobre o caráter e as formas da avaliação evidenciam as contribuições da Psicologia nesse campo. Dos sete artigos situados neste subgrupo, três abordam as questões da avaliação dos resultados da aprendizagem numa ótica piagetiana.

O primeiro deles, desenvolvido por Davis e Esposito (1990), centra-se na discussão de como devem ser entendidos o papel e a natureza dos erros na abordagem piagetiana, para que a avaliação do aproveitamento escolar se torne mais produtiva e mais conseqüente. Toma como referência a “teoria da equilibração majorante”, que pressupõe dois aspectos centrais: o estrutural, que se refere ao conjunto de esquemas já construídos pela criança, a um “saber fazer” por meio do qual o sujeito assimila os objetos à sua forma de pensar, às suas estruturas de pensamento; e o aspecto processual, que diz respeito às alterações que ações e conteúdos, propiciadas na interação sujeito-objeto, exercem sobre os esquemas, levando a um aprimoramento da forma como antes se configuravam. Considerada dessa forma, a construção do conhecimento implica momentos de equilíbrio e de desequilíbrio, criando-se momentos de conflito, cuja riqueza reside em provocar a busca de um novo estado de equilíbrio, superior e melhor que o precedente.

As autoras discutem, então, o papel e o significado do erro nesse processo. Numa proposta de ensino-aprendizagem derivada da teoria psicogenética, em que, na avaliação do aproveitamento escolar, cabe ao professor discernir os erros construtivos – aqueles que evidenciam progressos na atividade mental – daqueles que não o são, identificam alternativas para um possível erro da criança ao resolver uma tarefa. Consideram importante saber discernir entre os aspectos estruturais e processuais da formação do conhecimento, saber distinguir os erros construtivos, reenquadrando o papel do erro no seio da escola. Ou seja, o erro não deve ser punido, temido, evitado, mas sim encarado como resultado de uma postura de experimentação da criança, sendo esta a contribuição pessoal dos professores para reduzir o fracasso escolar e fazer da avaliação algo mais produtivo.

No segundo artigo deste subgrupo, Macedo (1993) propõe que o professor faça uma reavaliação do seu trabalho, da perspectiva de quem ensina e de quem aprende. Abordando, no contexto piagetiano, o ponto de vista da criança no ato de aprender, classifica as reações que ela apresenta em relação a uma tarefa em cinco categorias, que dizem respeito a: maior ou menor intensidade da reflexão sobre suas respostas; satisfação ou não da expectativa do examinador; utilização de recursos biológicos, sociais, coletivos ou de recursos próprios nas respostas a uma questão estimuladora. Partindo dessas colocações, analisa as perspectivas do professor no ato de ensinar, classificando-as segundo o maior ou menor distanciamento que ele revele quanto a saber o que o aluno pensa a respeito de um determinado assunto, ou como ele faz para resolver um determinado problema. Assim, o professor pode ser um simples observador do pensamento infantil ou encontrar formas de criar situações experimentais em que a criança seja levada a confrontar pontos de vista, a buscar explicações e justificativas para as suas respostas, permitindo a dúvida e não a busca de uma resposta tida como “certa”. Chamando o professor de psicopedagogo, Macedo sublinha a preocupação que esse profissional deve ter com os processos de aprendizagem do aluno e suas dificuldades nas relações com o ato de ensinar.

O terceiro artigo, de Corrêa e Moura (1991), questiona a prática de utilização de provas operatórias piagetianas – tarefas ou situações experimentais usadas para o estudo da gênese da competência operatória – com o objetivo de avaliar as capacidades intelectuais dos alunos na assimilação de determinados conteúdos como diagnóstico e classificação de crianças. Analisando a natureza e o papel das situações de pesquisa desenvolvidas por Piaget em relação à própria teoria e às evidências da literatura existente sobre o tema, as autoras consideram que não é factível nem desejável esta aplicação, apontando três argumentos contrários a elas. Primeiro, porque aplicá-las significa atribuir-lhes uma finalidade não pretendida por quem as desenvolveu. Segundo, porque tal uso confere às “provas” um *status* de teste psicométrico que não possuem, ou seja, além das suas limitações intrínsecas, observam-se nelas as falhas que vêm sendo imputadas aos próprios testes: as tentativas de correlacionar diferentes tarefas piagetianas, de um determinado período de desenvolvimento, mostram uma disparidade no desempenho entre diversos domínios. Finalmente, as autoras consideram que utilizar tais provas nessa situação de diagnóstico significa ignorar que, longe de se constituírem instrumento neutro “capaz de captar a lógica intrínseca da criança”, representam contextos específicos de interação, e assim devem ser consideradas.

Nos três artigos citados, observa-se que a utilização da teoria psicogenética para analisar a avaliação dos resultados da aprendizagem escolar ocorre dentro de uma perspectiva mais ampla do que em décadas anteriores, na medida em que ela é tomada como um instrumento de análise desse fenômeno. A grande popularidade alcançada, nas décadas anteriores, pelas provas operatórias como instrumento de diagnóstico psicopedagógico é colocada em cheque. Em vez de se tomar esta ou aquela parte ou conceito da teoria e aplicar à avaliação da aprendizagem escolar, apresenta-se um quadro teórico para a análise do processo de aquisição de conhecimentos constituído pela aprendizagem escolar, de modo a identificar os problemas mais significativos e formulá-los com vista a encontrar soluções relevantes e adequadas. Num sentido mais amplo, parece que, nesta década, a psicologia e a epistemologia genéticas vêm sendo utilizadas muito mais como instrumento de análise dos problemas educacionais do que consideradas como uma fonte virtual de soluções acabadas para resolver os males que assolam a educação.

Outros referenciais da Psicologia também sugerem novos enfoques para a avaliação. É o caso da Teoria das Inteligências Múltiplas, preconizada por Gardner e que Gama (1994) introduz como uma teoria que abre novas possibilidades em educação, propondo perspectivas originais para a avaliação de indivíduos e oferecendo opções que permitem aos educandos realizações pessoais em várias áreas do saber. Para a autora, essa teoria apresenta alternativas para algumas práticas educacionais, especialmente quanto a métodos de avaliação que considerem as inteligências avaliadas, a cultura dos alunos e o desempenho adulto a que se quer chegar. Nesse sentido, seria possível mudar o foco nas escolas: de testagem, para avaliação; da classificação de alunos, para a colaboração com eles na construção de habilidades intelectuais, com um aproveitamento ideal da variedade de oportunidades educacionais. A autora propõe, como decorrência da necessidade, apontada por Gardner, de avaliar as diferentes inteligências em termos de suas manifestações culturais e ocupações adultas específicas, que o levantamento de dados sobre o desempenho dos alunos não vise apenas ao conhecimento do que foi aprendido, mas como foi aprendido e de que maneira ele fará uso de tal aprendizagem.

A ênfase do artigo de Elliot e Filipecki (1995) reside na mudança de enfoque do ensino e da avaliação na formação de alunos críticos e autônomos, capazes de fazer uso de habilidades de investigação nas diversas áreas curriculares. As autoras apresentam o conceito de pensamento crítico, explicitando-o por meio da metáfora da figura da roda em movimento, desenvolvida por Clarke e Biddle em 1993, segundo a qual a aprendizagem é um processo contínuo, que pode ser corrigido em seu percurso, que admite aprofundamentos e ampliações, mas que não pode ser finalizado. O ensino do pensamento crítico baseia-se então num movimento circular, sugerindo um contínuo de diversos tipos de pensamento, dispostos em quatro fases ou quadrantes da “roda”. Com a utilização de estratégias metacognitivas de ensino voltadas para o desenvolvimento do pensamento crítico, a

aprendizagem passa a ser entendida como um processo pessoal de construção ativa de significados, e a avaliação é o instrumento indispensável de acompanhamento desse processo. Colocada a serviço do desenvolvimento do pensamento crítico, a avaliação apresenta características específicas: privilegia a auto-avaliação, além de adquirir uma função formativa, auxiliando o aluno a refletir sobre seu próprio processo de aprendizagem na construção seletiva de conhecimento. Nessa perspectiva, as autoras sugerem a construção, por professores e alunos, de alguns instrumentos de apoio à avaliação, explicitando os critérios ou níveis de desempenho para atribuição ou registro de um determinado conceito.

Mettrau e Almeida (1996) afirmam que a avaliação da aprendizagem é uma via de mão dupla onde transitam todos os que passam pelo processo de ensino-aprendizagem com suas respectivas competências (professores, alunos, famílias e instituições). A nova perspectiva de avaliação é mais abrangente e centrada no que os estudantes produzem, demonstram e expressam em diferentes momentos e deve valer-se de instrumentos cuidadosamente elaborados. A avaliação qualitativa tem o objetivo de captar a experiência dos alunos e deve ser apoiada em critérios estabelecidos conjuntamente por alunos e professores. Aspecto essencial a ser considerado diz respeito aos mecanismos executivos e de controle, usualmente chamados metacognição, que capacitam o aluno a estudar e acompanhar todos os passos dados no sentido da aprendizagem.

Por fim, o artigo de Carelli (1995) apresenta uma escala padronizada para interpretação de testes psicológicos e outros preditores, como sugestão para neutralizar os complexos problemas associados aos instrumentos de medidas psicológicas e educacionais.

Em síntese, tal como os demais referenciais das outras áreas de conhecimento, os estudos psicopedagógicos contribuem para a formulação do novo paradigma de avaliação qualitativa, conferindo especial ênfase à avaliação formativa, com insistência num olhar mais descritivo acerca de como operam os mecanismos de aprendizagem e a construção de conhecimentos pelo aluno. O eixo desloca-se do produto para o processo da aprendizagem, dando elementos para entender e trabalhar o papel do erro na escola; acentua-se o caráter diagnóstico da avaliação; a auto-reflexão do aluno sobre como aprende e o que aprende; a consideração de outras dimensões da avaliação que não a exclusivamente cognitiva; a interatividade no processo avaliativo.

Ainda como importante contribuição dos trabalhos nessa década, ressalte-se o destaque dado por alguns autores ao fato de que a compreensão de "como o aluno aprende" deve ser necessariamente acompanhada por uma intervenção qualificada do professor.

Outra vertente de estudos psicopedagógicos, a que traz nova contribuição à descrição das habilidades de aprendizagem a serem adquiridas e vem subsidiando a avaliação do rendimento dos alunos na linha do desenvolvimento de competências, não está representada nos textos selecionados.

O NÚCLEO DURO DA AVALIAÇÃO DOS RESULTADOS DA APRENDIZAGEM

Importantes indícios de uma prática recente, mas em franca expansão no Brasil, são registrados neste grupo de artigos. Eles representam, dentro do paradigma positivista, novas formas de abordagem da avaliação educacional, sempre dimensionadas para larga escala, e, no mais das vezes, animadas por uma particular lógica de implementação das políticas públicas na área.

Os artigos que introduzem modelos e metodologias de avaliação em larga escala nos meios educacionais brasileiros geralmente o fazem a partir de uma discussão prévia da importância da informação na linha da prestação de contas que o serviço público deve à população (*accountability*). Consideram-na também um instrumento possante para orientar as ações voltadas para os sistemas de ensino no que diz respeito à implementação, manutenção e/ou revisão de medidas adotadas.

Fletcher (1995b) parte do pressuposto de que se gasta mal em educação, de que a função principal da escola é ensinar e de que o resultado a ser esperado, avaliado e cobrado é a aprendizagem do aluno. Chega à necessidade de criação de sistemas extensivos de aferição de

resultados de aprendizagem, bem como afirma a importância de um corpo administrativo que faça pleno uso dessas informações, admitindo que as burocracias dos sistemas de ensino têm pouca experiência em sistemas de monitoramento.

Estabelece, de início, distinção entre pesquisa e avaliação de base escolar e domiciliar, indicando que a primeira parte de amostragem para tirar inferências sobre universo desconhecido é convencional, tópica, intermitente e despreocupada com a finalidade imediata de solucionar questões, enquanto a avaliação se caracteriza como atividade permanente e recorrente, não opera somente na base da amostragem e envolve compromisso de monitoramento a longo prazo.

Diferencia pesquisas de base escolar (que geram resultados agregados, permitindo a caracterização escolar da amostra) das pesquisas de base domiciliar (que ignoram as características internas da escola, fixando-se em informações sobre a população escolar). Nessa segunda modalidade, menciona o pro-fluxo: metodologia de pesquisa sobre o fluxo de alunos no sistema escolar, que se contrapõe às tradicionais análises estatísticas baseadas nas informações provenientes das unidades escolares, optando por valer-se de dados educacionais presentes nos censos demográficos. Mediante sofisticadas técnicas de análise, é possível reconstituir os detalhes da trajetória escolar de populações que freqüentam as escolas públicas e privadas, ou mesmo que estão fora da escola, determinando as taxas de promoção, repetência e evasão que compõem a matriz de transição de uma série do ensino para outra, em diferentes segmentos da população. Ao trabalhar com cortes de alunos, o pro-fluxo possibilita também identificar, nesses segmentos populacionais, a transição de um estágio cognitivo para outro.

Fletcher expõe em seguida, à guisa de ilustração, três modelos de avaliação de monitoramento, utilizados nos Estados Unidos, que se reportam a diferentes aspectos do sistema educacional e têm referenciais teóricos e objetivos diferentes.

O primeiro deles se propõe avaliar *insumos* e procura verificar a conformidade da escola com padrões técnicos especificados pelas agências públicas, a fim de assegurar um padrão básico de qualidade a toda a rede de escolas (empregado em New Jersey). O segundo modelo se propõe avaliar *processos*, tendo como foco a dinâmica do processo ensino-aprendizagem. Objetiva a melhoria progressiva do aluno mediante diagnoses vinculadas a procedimentos de recuperação e utiliza como instrumento diagnóstico de preferência a prova referenciada a critério, articulada com o conteúdo curricular da escola. Tem sido aplicado com grande sucesso no Estado de Minnesota. O terceiro modelo dedica-se à avaliação do *rendimento escolar* do aluno, priorizando o uso de provas cognitivas. Os resultados são agregados por escola ou distrito escolar antes de publicados. Esse sistema de monitoramento, aplicado no Estado da Califórnia, ganhou novo impulso nos anos 80 como parte do movimento em prol da *accountability*. Tem como pressuposto que a demonstração de rendimento das escolas bem-sucedidas serve de estímulo às demais.

A seguir, o autor coloca as limitações e controvérsias em torno dos três modelos, sendo que se destacará neste estudo apenas a argumentação referente ao monitoramento do rendimento do aluno, visto tratar-se do modelo que vem sendo introduzido no Brasil. Esse tipo de monitoramento tenta aplicar princípios da teoria econômica a organizações do setor público. O objetivo de promover a melhoria do desempenho acadêmico entre as unidades do sistema seria alcançado promovendo-se a concorrência entre as escolas. Na maior parte das vezes, espera-se utilizar a publicidade negativa para estimular administradores e professores a dedicar maiores esforços ao ensino. É cogitado o uso de sanções econômicas, políticas e regulamentares como incentivos.

De acordo com Fletcher, uma das grandes dificuldades da implementação desse sistema é que ele depende de enorme melhoria na capacidade de processamento da informação. Trata-se, portanto, de sistema a ser evitado se a administração não dispuser de fôlego financeiro e competência técnica para geri-lo com os novos recursos da informática. Uma segunda limitação consiste no eventual estreitamento do currículo por iniciativa das escolas, que, ao tentarem se adaptar às novas exigências, podem eliminar os conteúdos que não se relacionem diretamente com os referenciais curriculares básicos. Além disso, nem todas as unidades escolares têm condição de responder aos

sinais de mercado preconizado pelo sistema de monitoramento. A circunscrição das matrículas às zonas próximas da escola reduz a capacidade de escolha de melhores escolas pelos pais de alunos. Como lembra o autor, nas circunstâncias de monopólio e oligopólio, princípios de livre mercado não são aplicáveis, sendo duvidoso que escolas deficientes tenham iguais oportunidades para responder a incentivos e sanções. Conseqüência possível de políticas dessa natureza pode ser a exacerbação de diferenças cognitivas entre as comunidades ricas e pobres, urbanas e rurais, reforçando as disparidades sociais existentes.

O modelo do pro-fluxo, também discutido por Klein e Fontanive (1995), deu margem a uma série de estudos de grande impacto no País, mostrando que o grande problema do sistema educacional brasileiro é a repetência.

Na linha da pesquisa de base domiciliar, Fletcher (1991) apresenta ainda projeto de avaliação do perfil cognitivo da população em âmbito nacional, com vista a estabelecer os níveis cognitivos médios existentes nas diversas ocupações da força de trabalho, refletindo preocupação com a competência da mão-de-obra para utilizar tecnologia moderna na sociedade brasileira.

Quanto à utilização do sistema de monitoramento baseado na avaliação do rendimento do aluno, para além deste Estado do Conhecimento, a literatura internacional e mesmo uma série de textos publicados no Brasil indicam que ela está sendo ampliada não só nos Estados Unidos, como também vem sendo adotada em âmbito nacional em diversos países europeus. A sistematização do modelo proposta pela reforma educacional de 1988 na Inglaterra é a que tem sido mais conhecida e divulgada por meio dos organismos multilaterais, em grande parte responsáveis pela adoção de medidas homogeneizadoras em diferentes sistemas escolares do mundo.

Na América Latina, vários países também implantaram seus sistemas nacionais de monitoramento do rendimento escolar, inclusive o Brasil, que o fez a partir de 1990. O monitoramento dos sistemas de ensino básico, por via de regra, tem dado maior ênfase aos aspectos positivos do modelo, ou seja, ao suprimento, pelo Estado, de necessidades detectadas, do que à exacerbação da concorrência entre as escolas ou de eventuais sanções àquelas de baixo desempenho.

Em anos recentes vem sendo introduzida uma metodologia de operação desses sistemas de monitoramento no País, o que permite um grande salto em termos da natureza das informações educacionais produzidas nesse âmbito. Essa nova modalidade de processamento das informações deriva da Teoria da Resposta ao Item (TRI), que permite a comparação dos resultados da avaliação de aprendizagem de populações distintas em espaços e em tempos diferentes. Esclarecimentos metodológicos sobre essa teoria estão em três artigos de Fletcher (1994, 1995a e 1995b) e em outros dois, respectivamente, de Klein e Fontanive (1995) e Andrade e Valle (1998).

Os testes feitos sob o princípio da Teoria Clássica de Medidas não permitiam construir escalas de escolaridade para os diferentes assuntos sobre os quais os alunos eram avaliados; não havia como comparar avaliações feitas em diferentes ocasiões por diferentes professores. A Teoria de Resposta ao Item possibilita informar o que a população ou subpopulações de alunos em diferentes séries sabem ou são capazes de fazer em determinado momento e acompanhar a evolução desses alunos ao longo dos anos. Ela tem como elemento central os itens, e não a prova como um todo, e precisa de um grande número de itens de uma área de conhecimento que devem proporcionar boa cobertura da proposta curricular que, em princípio, deverá ter sido desenvolvida com os alunos. Com a cobertura mais abrangente dos conteúdos curriculares e dos processos cognitivos abordados, é possível evitar a tendência de adaptar o ensino à prova.

A teoria supõe que o desempenho do aluno em um teste pode ser explicado por características ou variáveis latentes (não observáveis diretamente), variáveis essas chamadas Proficiências ou habilidades. Procura-se reunir itens para os quais se supõe que uma certa Proficiência ou habilidade é dominante. Os itens são organizados em blocos incompletos e eqüitativos, com base em amostragem matricial, e os cadernos de blocos são distribuídos em espiral, garantindo que a aplicação seja aleatória e que alunos da mesma turma respondam, em geral, a cadernos diferentes, embora cada caderno de itens seja aplicado, aproximadamente, ao mesmo número de alunos. A TRI

opera com um conjunto de modelos onde a probabilidade de resposta a um item é modelada como função da Proficiência do aluno (variável não-observável) e de parâmetros do item (que expressam as propriedades deste). Uma importante propriedade da teoria é a da invariância dos parâmetros; por sua vez, quanto maior a Proficiência, maior a probabilidade de acerto. Utilizando uma escala comum, é possível estimar a distribuição de habilidades cognitivas ou Proficiências para a população e subpopulações de alunos e compará-las, bem como estimar relações entre as habilidades e variáveis socioeconômicas, culturais e do ambiente escolar. Graças às propriedades da TRI, é possível comparar alunos, mesmo que eles tenham respondido a itens diferentes em diferentes momentos.

Apesar de formulada há cerca de 50 anos, em virtude da complexidade e ausência de programas computacionais eficientes, apenas nos últimos quinze anos a Teoria da Resposta ao Item vem sendo aplicada em larga escala nas principais avaliações educacionais em diferentes países e, a partir de 1995, no Brasil (Andrade, Valle, 1998).

Diante dessa realidade, Vianna (1998c) discute oportunamente questões relativas à implantação da avaliação de sistemas educacionais, considerando que ela deve pautar-se em teoria que oriente o desenvolvimento das investigações e levar em conta variáveis relativas ao contexto da escola e aquelas não diretamente ligadas a ele, mas que têm importância para a educação.

Finalmente, seis artigos que não dizem necessariamente respeito ao núcleo duro da avaliação, foram enquadrados neste subgrupo por se referirem a técnicas de construção e elaboração de instrumentos, incluindo uma matriz de especificação de itens de provas. Por via de regra, esses instrumentos estão associados à abordagem positivista da avaliação, mas isso não impede que alguns deles possam prestar-se a análises de diferentes orientações teóricas. Desses seis artigos, três foram publicados nos *Estudos em Avaliação Educacional* (Depresbiteris, 1991; Costa, 1995; Buchweitz, 1996), dois em *Tecnologia Educacional* (Castilhos, 1990; Mário Silva, 1992) e um no *Em Aberto* (Rocha, Campos, 1993).