
REVOLVENDO O PASSADO DA AVALIAÇÃO EDUCACIONAL E ALGUMAS REPERCUSSÕES NA ESCOLA

Sheyla Maria Fontenele Macedo^(*)
Marcos Antonio Martins Lima^(**)

Nesse caso, terás de te julgar a ti próprio, disse o rei. Isso é o mais difícil de tudo. É muito mais difícil julgares-te a ti próprio do que julgar outros. Se fores bem sucedido ao julgares-te a ti próprio, terás encontrado a verdadeira sabedoria.
(EXUPERY, 1999, p. 21).

INTRODUÇÃO

O vocábulo avaliação vem do latim, “avalere”, que significa “dar valor a”. Dar valor exige tomar decisões, escolher caminhos, emitir opiniões frente a um fato, objeto ou fenômeno. A todo instante emitem-se julgamentos de valor sobre fatos, circunstâncias, falas, pessoas. Deste ponto de vista, entende-se que a avaliação se mistura na rotina dos indivíduos, em seu cotidiano. Esta atitude é eminentemente humana. Mas quando se trata da instituição escolar, do *locus* que sistematiza a avaliação e a transforma em ato pedagógico, o que é de fato avaliar? Para compreender esta dimensão, é necessário retomar o passado e revisar as histórias construídas e que ocuparam e ocupam espaço no aqui e no agora da escola, e esta é a finalidade desta pesquisa de cunho qualitativo e bibliográfico. As incursões históricas sempre remetem a reflexões sobre acertos e desacertos, e permitem a possibilidade em redimensionar novos caminhos.

Os subsídios teóricos ora levantados permitem a construção mental sobre o panorama da avaliação, na perspectiva de demarcar limites e possibilitar reflexões quanto a influência dos diferentes modelos avaliativos no *modus fazendi* da instituição escolar.

A história da avaliação principia em didáticas difusas que se pronunciam em práticas desde a Antiguidade. Entretanto, o seu marco existencial é datado no início do século XX, a partir dos testes docimológicos, sendo o termo “avaliação educacional” utilizado pela primeira vez por Tyler, então considerado “pai da avaliação”, na década de 1930. O período pós Tyler é caracterizado por inúmeras pesquisas no campo, destacando-se os trabalhos de Daniel L. Stufflebeam (1960), Cronbach (1963), Scriven (1967) e Stake (1967).

^(*) Professora Assistente II do Departamento de Educação (DE) da Universidade do Estado do Rio Grande do Norte (Uern). E-mail: sheylafontenele@yahoo.com.br.

^(**) Professor Adjunto da Faculdade de Educação (Faced) da Universidade Federal do Ceará (UFC). E-mail: marcos.a.lima@terra.com.br.

Num segundo momento, procurou-se analisar a avaliação frente aos desafios impostos pelo advento das ciências sociais, que caracterizam a avaliação enquanto fenômeno social, associando-a, por vezes, ao fracasso escolar de crianças, jovens e adultos das classes populares. O caráter meritocrático, centrado no rendimento escolar, escamotearia a exclusão escolar, ao mesmo tempo em que, de acordo com as teorias reprodutivistas, manteriam a hegemonia dos valores da classe dominante.

Os estudos em avaliação educacional, no Brasil da década de 1990, alavancaram pesquisas e estratégias no sentido de empreender a avaliação de massa, na finalidade de manter o controle e a regulamentação dos processos educativos do estado, em especial, quanto aos fenômenos da repetência, evasão e desnivelamento série/idade.

A avaliação no século XXI toma novos contornos, e a escola passa a contemplar discursos acerca do “ato de avaliar” enquanto emancipatório, dialógico e mediador.

O referido trabalho tem como cerne, na verdade, fazer o exercício do lembrar, o que significa apropriar-se dos passos dados para que outros caminhos sejam abertos, e que os erros sejam realmente superados.

A avaliação na escola é ainda “bicho de sete cabeças”, mas é preciso construir a consciência do feito, para que se aquilatem valores ao ato de avaliar, tornando-o mais equânime e afim com sua função social: favorecer a aprendizagem significativa.

O PASSADO DA AVALIAÇÃO DIFUSA NA ANTIGUIDADE À RACIONALIZAÇÃO CIENTÍFICA DO SÉCULO XX

A história da avaliação escolar e as suas práticas pedagógicas estão registradas desde a Antiguidade, que se configurava numa didática, por vezes enfadonha e que priorizava a memória como método fundamental na aprendizagem:

Sabe-se que na Antiguidade, antes da difusão da escrita, a memória era o único e indispensável meio de aprendizagem, e para isto serviam o verso e a música, ou a cantilena [...]. Outra razão obrigava ao uso da memória; a sacralidade e o caráter esotérico dos ensinamentos, reservados em geral a um grupo fechado, a uma casta sacerdotal. (MANACORDA, 1989, p. 92).

É neste período que a avaliação passou a ter como principais instrumentos, as arguições, questionários, provas, em especial as orais eram tomadas como uma das principais técnicas de ensino, desde a avaliação com crianças até alunos das universidades, e a defesa oral era exigência nas aprovações das teses de seus diplomados.

Já no século XII o rei de Castela e Leão, D. Afonso, o Sábio (1252-1284), estabeleceu no denominado código das “Siete Partidas”, livro das Leis, no seu artigo no. 10, como deveriam ser realizados os exames na academia:

O estudante que quer ter a honra de ser mestre, primeiro deve ser discípulo: quando este acabou de aprender a ciência, deve apresentar-se perante os ‘mayorales’ do estudo, que têm o poder de conferir-lhe licença. Estes (após verificado se é pessoa de bom nome e de bons costumes) deve lhe propor a leitura de livros das disciplinas de que quer ser mestre e se ele tem boa compreensão do texto e do glossário daquela ciência, bom método e linguagem fluente para expô-la, e responde bem às questões e perguntas que lhe fazem, devem dar-lhe publicamente a honra de ser mestre, fazendo-o jurar para expor correta e lealmente sua ciência (MANACORDA, 1989, p. 152).

Rituais e cerimônias de passagem tanto para criança como para o adulto, também eram comuns.

A escola tradicional foi influenciada, em suas práticas cotidianas, pelo Método de Ensino Mútuo, criado pelo escocês Andrew Bell (1753-1832), cujo objetivo era o de proporcionar educação a um grande número de alunos com baixos custos (SILVA, 2008). Este método mais adiante foi denominado de Método Lancaster, em referência ao jovem professor Joseph Lancaster (1778-1838) que cria em Londres uma escola destinada a crianças pobres, cujo número de alunos aumenta rapidamente e que sem condições para contratar professores, decide adotar a metodologia de Bell, e que se destaca pela rigorosa disciplina “[...] e a distribuição hierarquizada dos alunos sentados em bancos dispostos num salão único e bem amplo [...]” (SAVIANI, 2007, apud SILVA, 2008, p. 8). Este *viés* disciplinar repercute no sistema “[...] avaliando continuamente o aproveitamento e o comportamento dos alunos, esse método erigia a competição em princípio ativo do funcionamento da escola” (SAVIANI, 2007, apud SILVA, 2008, p. 8). Ainda no que se refere às situações relativas à avaliação propriamente ditas:

[...] se um aluno comete um erro, cede o lugar ao que está depois dele, que o corrige, e dessa forma o mais hábil chega a colocar-se no primeiro lugar. Desse modo, quem avança passa sempre para frente da divisão e quem regride passa para trás. A criança que por algum tempo ficou ocupando o primeiro lugar da classe passa para o último da classe superior. Assim também quem não faz progressos suficientes desce para o primeiro lugar à classe inferior (HAMEL apud MANACORDA, 1989, p. 260).

O método em questão evidencia o erro, ao mesmo tempo em que gera, entre os alunos, um espírito de competição e rivalidade.

A avaliação caminhou nesta linha tradicional até o final do século XIX, quando o advento da Psicologia trouxe, como resultado de seus estudos, os primeiros testes psicométricos, cujo interesse é despertado pela intenção da medição científica da inteligência humana. Os referidos testes sofreram influências do positivismo das ciências físico-naturais, sustentados nas exigências do rigor científico; de pesquisas tais como os trabalhos de Darwin, que se alicerçava no recurso da medição para delinear categorias dos indivíduos e suas diferenças e no desenvolvimento dos métodos estatísticos. Surge assim, em 1922, na França, a ciência “docimológica” ou a “docimologia”, que significa “medida por exame” (ou em latim *docimê*), caracterizando a era dos testes em avaliação.

A Primeira Guerra Mundial obriga ao aperfeiçoamento dos testes, frente à necessidade de novos recrutamentos e, neste sentido, é crescente a preocupação com a validade dos exames prestados e, estendendo-se a mesma, aos critérios estabelecidos nos concursos públicos realizados.

Os testes docimológicos visavam a eliminar quaisquer visões de ordem pessoal por parte dos examinadores. Partiam da ideia de que o exame é uma medida, cujos resultados devem ser aferidos com critérios claros e preestabelecidos. Eram constituídos de questionários de caráter eminentemente objetivos, nos quais buscava-se o rigor metodológico visando evitar discrepâncias nos resultados, que muitas das vezes, nos casos dos testes de concursos, redundavam em injustiças díspares. Por isso, o termo avaliação até hoje é confundido por “medida”.

O início do século XX é marcado pela mudança do paradigma socioeconômico agrário para o industrial, o que mobiliza a necessidade em se repensar a relação dos conhecimentos transmitidos pela escola e de forjar mecanismos para o aprimoramento e seleção de alunos. Ganham volume as preocupações associadas à institucionalização da educação de massas, que exigia a extensão da educação escolarizada em níveis cada vez mais altos a segmentos cada vez maiores da população. O pensamento surge em especial nos Estados Unidos e na Inglaterra, cujo crescente processo de industrialização e urbanização gerou a necessidade de manutenção da identidade nacional, resultado das sucessivas ondas imigratórias.

É com Ralph Tyler, educador norte americano, que se registra o momento histórico em que a avaliação se constitui enquanto área de estudos e pesquisas, sendo o termo “avaliação da

aprendizagem” a ele atribuído a partir de 1930. Os pressupostos de Tyler foram construídos com base em influências, tais como Alfred Binet¹ em seus estudos sobre testes de inteligência em 1905.

Os resultados dos chamados testes de inteligência eram utilizados na interpretação do desempenho dos estudantes nos testes de escolaridade, sobretudo quando se tratava de um baixo desempenho. A ideia de que qualquer pessoa poderia aprender começou a ser posta à prova e, em consequência, houve mudanças nos testes de rendimento, que eram referenciados a critério e o desempenho do aluno era acompanhado a um valor absoluto, passando a testes do tipo referenciados a normas, em que o desempenho do estudante é relacionado ao comportamento do seu grupo (VIANNA, 2000, p. 48).

Dentre as obras de Tyler, destaca-se o ensaio sob o título de “General Statement on Evaluation” em 1942, em que apresenta que a avaliação por objetivos permite a escola verificar até que ponto cumpre satisfatoriamente sua função enquanto instituição educacional.

Tyler apresenta ao cenário da avaliação um enfoque atrelado aos objetivos educacionais. Suas pesquisas tinham a finalidade de promover a eficácia da escola e de seus métodos educativos, e tomou como base a análise dos resultados de programas educativos nos Estados Unidos. A avaliação é, portanto, no modelo de Tyler, em 1942, consignada ao paradigma da racionalização científica, até então, utilizado nos processos industriais. Em linhas gerais, entendia que os resultados da educação deveriam provocar mudanças comportamentais nos padrões preestabelecidos pela sociedade. Ou seja, os objetivos propostos precisariam estar em estreita relação com os resultados esperados, pois só assim se concretizaria a mudança de novos valores sociais e culturais. Neste sentido, a avaliação por objetivos também promoveria a revisão do currículo na construção de novas habilidades específicas. A avaliação passaria a ser, portanto, a ligação entre a escola e a sociedade, apontando problemas e interferindo nas possíveis soluções. É um modelo que promove a tomada de decisão.

É relevante em sua pesquisa a premissa teórica construída sobre a relação avaliação-professor/aluno-currículo. Entende que o professor deve deixar claro sobre os resultados a serem alcançados com seus alunos e expor seus objetivos através das diferentes práticas curriculares. São pontos importantes de sua obra, sintetizados por Vianna (2000, p. 52):

¹ Alfred Binet (1857-1911): psicólogo francês que criou o primeiro teste de inteligência no ano de 1905. O teste era estandardizado, de forma que o último dos vários grupos de teste que fossem completados com sucesso, por uma criança, indicava a sua “idade mental”

A educação enquanto processo que visa a criar padrões de conduta, ou a modificar padrões anteriores, nos indivíduos;

Os padrões de conduta na escola são, na realidade, os objetivos educacionais;

O êxito de um programa educacional, verificado através da avaliação, depende da concretização desses objetivos;

A avaliação deve incidir sobre o aluno, como um todo, nos seus conhecimentos, habilidades, modos de pensar, atitudes e interesses, sem se concentrar em apenas elementos isolados, como, na realidade acontece nos dias fluentes;

A avaliação pressupõe diversidade de instrumental para avaliar múltiplos comportamentos, não devendo ficar restrita, apenas, a exames escritos, como geralmente ocorre;

A avaliação não se concentra apenas no estudante [...] não é um ato isolado, mas um trabalho solidário que deve envolver, além do aluno, claro, os professores, administradores e, sem sombra de dúvida, os próprios pais, que devem ter voz ativa no processo.

A teoria de Tyler vinculada à construção de objetivos comportamentais teve reações e críticas, em especial no que se refere: às posições epistemológicas adotadas frente aos padrões comportamentalistas selecionados, a influência behaviorista e a metodologia de caráter quantitativo. A crítica acirrada, segundo Vianna (2000) deu-se no confronto com teóricos sociais, em torno do caráter metodológico apontado na orientação individualista, que desconsideraria os aspectos sociais da avaliação. Críticas neste sentido foram inclusive tecidas, alguns anos depois, por Habermans (1982) e Guba (1990).

Por outro lado, a era Tyler, de 1930 a 1945, ao manter sua tônica na pedagogia por objetivos, permitiu que se revisasse o paradigma docimológico, centrado nas medidas, para aquisição de habilidades, propiciando reflexões quanto ao significado e a instrumentalização sobre o ato de avaliar.

As premissas teóricas de Tyler influenciaram outros estudos, dentre eles, a teoria da taxonomia dos objetivos, de Benjamin Bloom, na década de 1950, que prevê uma estrutura de organização hierárquica para os objetivos educacionais e que também classifica os tipos de aprendizagem em cognitivo, afetivo e psicomotor, criando subdomínios para estes campos. Mais tarde outros modelos se delinearam, tais como a *accountability* (ou prestação de contas), metodologia utilizada pelo poder público na prestação de contas à população dos serviços que esta oferece, assim como por agências financiadoras.

Nas demais décadas proliferam-se a literatura neste campo e o avanço de testes e provas, em especial no período pós Segunda Guerra Mundial. Crescem as instituições educacionais e com eles, os testes padronizados, estandardizados, especialmente mal vistos pelo movimento “escolanovista”.

REVOLVENDO INFLUÊNCIAS DAS CIÊNCIAS SOCIOLÓGICAS. A AVALIAÇÃO DE MASSA DOS SÉCULOS XX E XXI

O período de 1958-1972 foi influenciado por estudos nos Estados Unidos, que instituiu a avaliação obrigatória da educação, por conta da necessidade em se realizar diagnósticos frente às hipóteses acerca das reais causas que assolavam os baixos rendimentos escolares das camadas populacionais mais pobres. Neste sentido, a avaliação amplia seu raio de ação, sofrendo forte influência dos estudos sociológicos, que avançaram em pesquisas quanto aos resultados desiguais produzidos pelo sistema educacional, tendo especial preocupação com o fracasso escolar de crianças, jovens e adultos da classe popular nos Estados Unidos (SILVA, 1999, p. 65). A partir deste marco, escolas, professores, alunos, conteúdos, metodologias passam a ser o foco da avaliação, que desloca o seu olhar do caráter meritocrático para o sociológico. A avaliação é então orientada tanto pela tendência positivista, de caráter quantitativo, quanto pela fenomenológica, naturalista, de cunho qualitativo.

Destacam-se, neste momento histórico da avaliação, os trabalhos de Cronbach (1963), Scriven (1967), Stake (1967) e Daniel Stufflebeam (1960).

Cronbach (1963) não teve necessariamente a intenção de apresentar um modelo para a realização de avaliações, na verdade seu embate é no campo das ideias que dão sustentação a reflexões mais amplas no campo da avaliação. Em sua concepção, a avaliação não tem por finalidade fazer um julgamento final, o que significaria limitar os objetivos. Sua função central seria a de fornecer subsídios para o aprimoramento dos currículos. Suas ideias contribuíram para os trabalhos de Scriven (1967) e Stake (1967).

A avaliação, no seu sentido mais amplo, pode ser definida como um processo que visa à coleta e ao uso de informações que permitam decisões sobre um programa educacional. A avaliação, portanto, deve ser entendida como uma atividade diversificada, que exige a tomada de vários tipos de decisões e o uso de grande número de diferentes informações. A avaliação com vistas ao aprimoramento de currículos, não deve ser confundida, como muitos o fazem, com a construção de instrumentos de medida e a obtenção de escores de fidedignidade, processos que, eventualmente, podem entrar no

contexto da avaliação, mas que não são indispensáveis para que ela possa atingir seus objetivos. (CRONBACH,1963 apud VIANNA, 2000, p. 68).

Do apresentado, destaca-se que Cronbach desloca o eixo da avaliação dos objetivos educacionais para a “tomada de decisão”. Neste aspecto, enfatiza Vianna (2000, p. 71) que um dos problemas destacados por Cronbach (1963) diz respeito ao tipo de decisão que se deve tomar e quais instrumentos utilizarem para tal fim. O instrumental deve cumprir todos os pré-requisitos psicométricos desejáveis, em especial em avaliações individuais. Para a avaliação de cursos, sugere a técnica de amostragem de itens, por oferecer julgamentos mais precisos.

Em suas premissas teóricas Cronbach (1963) realiza estudos do tipo *follow-up*, que consiste em observações realizadas de perto e até o fim de um plano ou execução de algo. Nesta perspectiva, discute técnicas para as medidas de atitudes:

A mensuração de atitudes pode ser feita de diferentes modos: entrevistas, questionários e outras técnicas semelhantes, sem dúvida, valiosas, mas cujos resultados, como adverte Cronbach (1963), devem ser aceitos com as devidas cautelas, porque estão sujeitos a distorções. (VIANNA, 2000, p. 74).

Em seus trabalhos encontram-se ainda discussões no campo do planejamento da avaliação, assim como discute o papel político da avaliação nos diferentes campos em que essa se constitui.

Finalmente, apresenta que o avaliador possui função comunicacional ao investigar o fenômeno que é de interesse público, o que consiste, necessariamente, na responsabilidade ética ao difundir e divulgar resultados de suas pesquisas, que devem estar assentadas no vocábulo “credibilidade”.

Este período de expansão do paradigma da avaliação é marcado pelas discussões em torno da avaliação formativa e somativa, com Scriven, em 1967.

As premissas teóricas de Tyler consolidam muitas compreensões no campo da avaliação, dentre elas de que é atividade finalística, isto é, supõe fins a atingir. Logo, ao estabelecer uma relação íntima entre a avaliação e a formulação de objetivos, que viriam a nortear os resultados e os impactos frente ao processo ensino-aprendizagem, indicando o quê e como julgar, ou seja, o quê e como avaliar, a função da avaliação enquanto somativa vai se consolidando nas pesquisas de Scriven.

Michael Scriven (1967) em sua obra “Methodology of Evaluation”, é o teórico que se debruça no sentido de diferenciar os papéis somativo e formativo da avaliação. Ratifica as

construções em torno da função somativa, delineando seu caráter classificatório e finalístico. Demonstra que a função somativa consiste em classificar os resultados de acordo com os níveis de aproveitamento previamente estabelecidos, sendo função norteadora nas tomadas de decisão.

A avaliação formativa, também possui função de controle, é processual, com o intuito de verificar se os objetivos previstos estão sendo atingidos, isto é, se os resultados foram alcançados durante o desenvolvimento de cada uma das etapas do processo. É através da avaliação formativa que se identificam os erros e acertos e de onde se extraem estímulos para um estudo sistematizado. Por isso, a avaliação formativa é fonte de motivação para o estabelecimento adequado no momento da tomada de decisão.

Scriven considera imprescindíveis os dois tipos de avaliação. Neste sentido, aponta Vianna (2000, p. 87):

Os dois tipos de avaliação-formativa e somativa - são fundamentais, considerando que várias decisões devem ser tomadas no decorrer das diversas fases do desenvolvimento do programa, e que, ao término dever-se-á determinar o mérito desse programa e empreender outras decisões, inclusive em relação ao futuro do programa, se deve continuar ou ser encerrado definitivamente. Apesar da importância desses dois tipos de avaliação, observamos que, habitualmente, os avaliadores concentram seus esforços em avaliações somativas. É preciso que o avaliador tenha cautela, pois, sem a avaliação formativa, o processo de desenvolvimento, seja de um programa, projeto ou material, resultará incompleto e inteiramente ineficiente. A realização de uma avaliação apenas ao final – somativa, portanto, - somente tem possibilidade de, muitas vezes, constatar o fracasso de um projeto.

Desta forma, é notório em sua teoria que os dois tipos de avaliação são imprescindíveis e complementares e podem acontecer de forma combinada, ou mesmo alternada.

Scriven (1973) ainda discute a dimensão da avaliação por objetivos forjada por Tyler, que denomina de *goal-free*. Explicita que os objetivos formulados numa avaliação geram objetivos colaterais. E que estes nem sempre são também avaliados, pois coexistem na subjetividade do processo. O modelo *goal-free* parte do princípio que os objetivos estabelecidos numa avaliação devem passar pelo crivo da avaliação, ou seja, para que não se criem objetivos de caráter colateral, que nem sempre são identificados pelo sujeito que avalia.

Os estudos de Scriven (1973) estendem-se ainda quanto ao sentido do papel do avaliador formativo, da necessidade do estabelecimento de critérios claros na avaliação, dos modelos e aspectos metodológicos frente ao ato de avaliar, do momento da formação do julgamento de valor, que demanda a tomada de decisão, além da necessidade de que se empreendam estudos avaliativos sobre a própria avaliação, revelando mais uma área de pesquisas, a da meta-avaliação.

Daniel Stufflebeam na década de 1960, ao viver a experiência na avaliação do sistema público elementar e secundário em escolas do distrito de Columbus, Ohio nos Estados Unidos, cria o modelo CIPP *Context, Input, Process and Product*, que significa: contexto (C), Insumo (I), processo (P) e produto (P).

O referido modelo revela os diferentes estágios de desenvolvimento de uma avaliação e tem como foco uma análise ampla sobre o processo de tomada de decisões. Em outras palavras, o modelo CIPP consiste em recolher informações, permitindo aos avaliadores reunirem dados para decidirem o melhor caminho frente a um processo decisório subsequente.

Em linhas gerais, é um modelo orientador, sendo também conhecido como modelo das quatro fases, conforme explicitação a seguir.

A avaliação do contexto tem como premissa básica o foco nas necessidades da avaliação (identificação dos problemas e deficiências), delineando para tal, objetivos e amostragem do objeto a ser considerado no estudo. É o tipo mais comum de avaliação.

O estágio do insumo possui o foco no detalhamento dos procedimentos, estratégias da avaliação, enfim, dos recursos humanos e físicos, com vistas a fornecer informações precisas, considerando inclusive a relação custo-benefício, para melhor planificar ações, programas, etc.

A avaliação de processo tem como eixo central a sistematização do processo decisório, por meio do fornecimento de informações para que as ações necessárias sejam planificadas e tomadas. Tem ainda função de realimentar, periódica e continuamente, os responsáveis pelo programa em todas as fases de seu desenvolvimento, no sentido de detectar, no processo, as possíveis deficiências de planejamento ou implementação, para que sejam sanadas e corrigidas em tempo hábil.

A avaliação de produto configura-se na verificação, medida e interpretação dos resultados obtidos na avaliação como um todo, quer seja em momentos pré-definidos ou no ato conclusivo, com vistas a determinar as discrepâncias entre os objetivos traçados e o real. Fornece informações para prosseguimento ou alteração dos objetivos.

Com base nas definições traçadas, Stufflebeam apud Vianna (2002), destaca três variáveis para melhor compreensão de seu modelo:

A avaliação enquanto processo sistemático e contínuo;

O processo de avaliação pressupõe três momentos da maior importância:

esboçar as questões a serem respondidas,

obter informações que sejam relevantes para responder às questões propostas;

proporcionar aos responsáveis pela tomada de decisões todas as informações necessárias.

A avaliação serve para a tomada de decisões. (STUFFLEBEAM apud VIANNA, 2000, p. 103).

O modelo apresentado está assentado na definição da avaliação “como um processo para descrever, obter e proporcionar informação útil para julgar decisões alternativas”. (STUFFLEBEAM apud VIANNA, 2000, p. 103).

O modelo de Stufflebeam promove avanços no campo da avaliação institucional, do uso dos resultados nos processos de tomada de decisões, entretanto, não considera dimensões relacionadas ao campo da subjetividade e interações socioculturais.

Na esteira do avanço das pesquisas em avaliação, Robert Stake, em 1967, escreve “The Countenance of Educational Evaluation”, em que apresenta o modelo da “avaliação responsiva” que provoca uma ampla discussão nos diversos meios educacionais (VIANNA, 2000, p. 126).

A avaliação responsiva ou *countenance*, demanda a necessidade do aprofundamento do conhecimento de todos os elementos que constituem a cultura institucional a ser avaliada e dos posicionamentos ao que ela deve chegar a ser. A finalidade da avaliação é de caráter responsivo, ou seja, dar resposta. Logo a coleta da informação tem como objetivo promover a troca, o diálogo entre os sujeitos e sobre a avaliação em todas as fases de seu processo, para a solução de problemas. Pressupõe este tipo de avaliação uma mudança de atitudes, crenças, comportamentos, o que demanda tomar como base metodológica a coleta da informação em caráter descritivo, para o posterior julgamento dos resultados. Desta forma, a avaliação é um fenômeno que ora necessita da precisão de dados quantitativos, ora dados qualitativos “A avaliação qualitativa se resume a um pequeno número de casos e um grande número de variáveis e a avaliação quantitativa a um grande número de casos e um pequeno número de variáveis” (STAKE, 1967, p. 524).

A avaliação responsiva não se vê obrigada, necessariamente, a aplicar testes de caráter taxológicos, toma, como uma das metodologias, as entrevistas, de modo que envolvam todos os sujeitos da avaliação, manter como critério ético a confidencialidade dos dados na elaboração do relatório final.

A avaliação responsiva, portanto, encaminha reflexões quanto ao papel da avaliação frente aos paradigmas quantitativo e qualitativo, tão discutidos nos dias de hoje:

As generalizações a que as avaliações e/ou pesquisas quantitativas e qualitativas chegam não são as mesmas [...] considerando que possuem diferentes conceituações dos dados e diferentes bases para validação e interpretação, por isso, acreditamos que ambos os tipos de avaliação/pesquisa se completem para a estruturação de novos conhecimentos, a partir de um mesmo fenômeno visto por intermédio de óticas diferentes (VIANA, 2000, p. 124).

Stake mais adiante, no ano de 1994, na obra “The Purpose of case study is not to represent the world but to represent the Case”, toma o estudo de caso como objeto de investigação sob a perspectiva da avaliação, que nos dias de hoje ainda não foi devidamente explorado. Destaca, neste trabalho, que tomar o estudo de caso como metodologia avaliativa implica, necessariamente, ter que considerar diferentes aspectos, dentre eles: a conceituação do objeto de estudo, a definição da prioridade das questões avaliativas, o estabelecimento de padrões frente aos dados, a necessidade de metodologia adequada às observações, a escolha de interpretações que possam realmente subsidiar as investigações e o desenvolvimento de afirmativas sobre o caso.

A década de 1980 é marcada pelos estudos das teorias reprodutivistas, que tiveram como expoentes Louis Althusser (teoria dos aparelhos ideológicos do estado), Bourdieu e Passeron (a reprodução), Baudelot e Establet (a escola dualista). Estas premissas teóricas favoreceram inúmeros recortes epistemológicos em diversos campos do conhecimento, em especial, da avaliação educacional. A avaliação passa a ser refletida como reprodutora do “status quo”, ou, em outras palavras, reproduz as estruturas de classe. As respectivas teorias denunciam que as práticas meritocráticas ao premiar o “merecimento”, ou seja, “dar boas notas a quem merece”, cria lacunas frente às análises das questões relativas às disparidades sociais existentes na sociedade e provocadas no interior da escola e que, por vezes, põe os alunos de classes sociais distintas em “pé” de desigualdade, configurando-se o mérito em requisito fundamental para a exclusão social.

Pulverizam-se, nas décadas de 80 e 90 do século XX os estudos sobre avaliação, que passa a ser tomada como objeto de estudo no campo das ciências humanas e sociais. As questões discutidas

sobre avaliação quantitativa e qualitativa são retomadas a partir deste panorama, desta vez, buscando-se a complementaridade entre as diferentes modalidades de avaliação.

O período das reformas neoliberais empreendidas no Brasil, em especial na década de 1990, origina, para o campo da educação, estudos que considerem o desafio de buscar instrumentais para realizar a avaliação de massa no País. Este cenário foi organizado no governo de Fernando Henrique Cardoso (FHC), cujas medidas foram empreendidas com empenho pelo Ministério da Educação e Cultura (MEC), com objetivos de analisar o rendimento escolar dos alunos, compreender as variáveis de construção da desigualdade social e escolar, assim como os fenômenos da evasão e repetência, mas de também e, especialmente, a avaliação foi tomada como cerne na promoção do controle e regulamentação dos processos educativos. Neste sentido, foram implementados programas de avaliação, assim como o Sistema de Avaliação da Educação Básica (SAEB) a nível federal, Exame Nacional de Cursos, ou “o Provão” -como é conhecido popularmente - e programas de avaliação no âmbito estadual, um exemplo a ser citado é o Sistema Permanente de Avaliação da Educação Básica do estado do Ceará (SPACE). Nesta rota, é ainda implementado no âmbito do ensino superior, o Sistema Nacional de Avaliação da Educação Superior (SINAES) é implantado pela lei 10.861 de 14 de abril de 2004, cujo método de classificar as universidades por resultados alcançados, termina por ser uma nova forma de produzir rankings entre as diferentes instituições do país.

No bojo dos programas e sistemas implementados no Brasil com vistas ao controle e administração do sistema educativo a avaliação consolida a dicotomia entre os paradigmas emancipatório *versus* somativo. Este último tem focado sua atenção em instrumentos sofisticados de controle, enquanto o setor público “patina” no sentido de assimilar as mudanças necessárias para que de fato esta avaliação se institucionalize na prática.

No Brasil, a concepção de avaliação, a partir do final do século XX, adquire novas formas e contextos, originam-se discussões em torno da avaliação enquanto emancipatória, dialógica, mediadora, dentre outros modelos.

A avaliação emancipatória, de acordo com Saul (1995), organiza-se com base no processo de crítica consciente da realidade por parte dos sujeitos “aprendentes”, a fim de que estes possam se emancipar, escrever suas próprias histórias, por vias da libertação que se dá pela interação com o conhecimento, revelando o caráter eminentemente político pedagógico da avaliação.

A avaliação dialógica, discutida na obra de Romão (1998), dá-se na perspectiva de que a produção do conhecimento é um processo coletivo de descoberta, deixando a avaliação de ocupar o

espaço de processo finalístico, de cobrança, de mecanismo de coerção para se transformar em espaços em que a dialética esteja centrada no vocábulo ‘aprendizagem’.

Hoffmann (2003), por sua vez, apresenta a concepção da avaliação enquanto mediadora, ou seja, a ação educativa realiza-se com base na busca de alternativas de integração aluno-escola, com vistas a que, mediante questões desafiadoras, o educando desenvolva a sua autonomia nos diversos campos de sua formação, seja este moral, afetivo, intelectual ou social.

As pesquisas no campo da avaliação que despontaram a partir da década de 1990 e até os dias de hoje, vêm redimensionando o contexto conservador das diferentes funções delegadas à avaliação educacional (a diagnóstica, a de controle e a de classificação). A função diagnóstica, tradicionalmente realizada nas fases iniciais do “fazer avaliativo”, é compreendida como processo e que acontece em vários momentos da ação educacional. Deixa também de estender seu foco somente ao aluno, e passa a debruçar o seu olhar sob a análise criteriosa dos conteúdos ou habilidades a serem desenvolvidas, revê objetivos traçados, estratégias metodológicas, as condições da estrutura organizacional da escola, dentre os inúmeros aspectos que interpenetram o cotidiano escolar e que possam influenciar na aprendizagem.

A função de controle, de caráter formativo, busca a formulação de novas hipóteses sobre os aspectos afetivos, cognitivos e relacionais, estendendo-se não somente ao âmbito da sala de aula, mas abrindo novos espaços epistemológicos para compreender de que forma é produzida a assimilação do conhecimento, ou em outras palavras, como o aluno aprende.

E, finalmente, a função somativa, cujo desempenho sempre fora classificatório, certificativo, aponta que os resultados do rendimento escolar devem nortear na construção de novos caminhos no processo ensino-aprendizagem, enfim, a avaliação tradicional, historicamente “ensimesmada”, realiza contrapontos com outros espaços e áreas afins.

CONSIDERAÇÕES FINAIS

As incursões sobre a temática da avaliação iniciam seu percurso desde a Antiguidade, configurando-se, entretanto, como área de pesquisa somente a partir das investigações realizadas por Tyler, a partir da década de 1930. É notório que a avaliação por objetivos de Tyler continua sendo útil nos momentos de julgamento de valor, mas é preciso compreendê-la enquanto processual, sistemática, contínua e formativa, tendências estas que necessitam se consolidar na escola.

As contribuições advindas das ciências psicológicas e sociológicas para as pesquisas em avaliação promoveram considerações importantes para revelá-la enquanto reprodutora do “status quo” e a desvendar a realidade socioeconômica dos alunos frente aos resultados do rendimento escolar, permitindo que se abrissem novos sulcos na busca por uma avaliação equânime e humanizada. Do exposto, entretanto, pode-se concluir que a avaliação é, como área de estudos, “terra virgem e fértil”, em detrimento de seu curto percurso histórico.

É fato que, no Brasil, inicia-se uma caminhada, em que a “duras penas” se busca construir uma cultura avaliativa. É corriqueiro identificar as tensões psicológicas daqueles que empreendem ou mesmo participam de processos desta natureza. Comum é deparar com professores que manifestam sentirem-se inseguros sobre como proceder em suas avaliações, alunos que se desestimulam frente aos diferentes instrumentos avaliativos (tais com provas de final de curso), das instituições, que receiam as avaliações de desempenho e dos próprios sistemas educacionais, que vivem inúmeras situações conflituosas para definir os rumos da avaliação de grande porte. O fato é que, a avaliação não pode ser a promotora em processos de *ranking*, consolidando-se em “avaliação mercadológica” que investe no controle e na regulamentação e que despreza o caráter formativo, instalando mecanismos, por vezes de caráter permanente, de exclusão social, o que não significa que a escola exclua o “mérito” de seu espaço, mas que as formas como este vem ocorrendo devem ser reestruturadas.

Não restam dúvidas que a influência dos estudos sobre a avaliação tem repercutido significativamente no cotidiano da sala de aula, permitindo que a escola reconsidere seu *modus fazendi*, suas práticas, em especial a partir do início do século XXI, em que o paradigma da complexidade sugere a revisão de novos modelos epistemológicos, forjando, por sua vez, a necessidade da criação de uma nova geração de profissionais com formação e pesquisas específicas na área. Logo, é imprescindível formar profissionais especialistas no campo.

Finalmente, avaliar é condição *sine qua non* das instituições escolares, e, neste ponto, a referência se faz a todos os níveis de ensino, que para tal desiderato, deverão compreendê-la como uma construção reflexiva e coletiva sobre o significado das relações entre ensino e aprendizagem.

A avaliação deverá, portanto, preencher todos os espaços da vida escolar, assim como o “polvo” que busca, com seus inúmeros tentáculos, expandir-se com mansidão em seu *habitat* natural, sobrepujando mesmo as pedras mais rochosas do fundo do mar, buscando

[...] decorar a sua habitação com belas pedras, conchas e pedaços de vidro colorido, esconder-se, resolver labirintos e problemas, distinguir padrões e formas, praticar aprendizagem por observação e mesmo brincar atirando

garrafas ou objetos contra a corrente e apanhando-os de seguida [...] (SOARES, 2006).

Para o exercício deste bem viver, não deverá, entretanto, sentir-se sufocado, ou ameaçado, ou do contrário, manchará com “jatos de tinta” este mesmo habitat na intenção de escapar de seus “predadores”, deixando-os por vezes sem visão.

É chegada a hora em que a escola deve analisar minuciosamente todas as instâncias em que este processo de avaliação se cria, produz e reproduz a fim de permitir que a avaliação continue seu percurso natural e cumpra, assim como o “polvo” o pleno desenvolvimento de sua inteligência, com “capacidade sensorial para aprender, [...] e utilizar conhecimentos adquiridos para resolver problemas [...]” (SOARES, 2006) inerentes à vida.

REFERÊNCIAS

HOFFMANN, Jussara. *Avaliação mediadora: uma prática em construção da pré-escola à universidade*. 20. ed. Porto Alegre: Mediação, 2003.

MANACORDA, Mario Alighiero. *História da Educação: da antiguidade até nossos dias*. São Paulo: Cortez, 1989.

ROMÃO, J. E. *Avaliação dialógica: desafios e perspectivas*. São Paulo: Cortez, 1998.

SAUL, A. M. *Avaliação emancipatória: desafios à teoria e à prática de avaliação e reformulação de currículo*. 3.ed. São Paulo: Cortez, 1995.

SILVA, Andréa Villela Mafra. Mostrem as lousas aos mestres. O método lancasteriano de ensino. Dialogando com alguns autores sobre as escolas de primeiras letras. Rio de Janeiro: 2008. Disponível em: <http://br.librosintinta.in/biblioteca/ver-doc/www.histedbr.fae.unicamp.br/acer_histedbr/jornada/jornada8/txt_compl/Andrea%2520Silva.doc.htm>. Acesso em: 13 jun. 2011.

SILVA, Tomás Tadeu. *Documentos de Identidade: uma introdução as teorias do currículo*. São Paulo: Autêntica, 1999.

SOARES, Ana. *A inteligência do polvo*. Portugal: 2006. Disponível em: <<http://www.centrovegetariano.org/Article-420-A%2Bintelig%25Eancia%2Bdo%2Bpolvo.html>>. Acesso em: 13 jun. 2011.

STAKE, R. E. *The Countenance of Educational Evaluation*. Teachers College Record, New York, n. 68, p. 523-540, 1967.

VIANNA, Heraldo Marelim. *Avaliação Educacional: Teoria, Planejamento, Modelos*. São Paulo: Ibrasa, 2000.

RESUMO

O artigo apresenta um breve histórico da trajetória da avaliação educacional, a fim de promover reflexões sobre o acontecer do "ato de avaliar" na escola. A pesquisa é qualitativa, bibliográfica. Denota que a avaliação educacional iniciou seu percurso na Antiguidade, sendo inaugurada como campo epistemológico por Ralph Tyler na década de 1930, até chegar as discussões sobre a avaliação de massa nos dias de hoje. Como referencial teórico: Manacorda (1989), pelas contribuições dadas à história da educação e Vianna (2000), pelos constructos delineados no campo da avaliação (histórico e modelos avaliativos). A pesquisa torna-se relevante para o campo da educação por permitir vasculhar de forma analítica e reflexiva a história particular e pouco discutida da avaliação educacional.

Palavras-chave: Avaliação. História. Educação.

ABSTRACT

The article presents a brief history of the trajectory of educational evaluation in order to promote a reflection on the happening of the act "to evaluate" in school. The research is of qualitative and literary character. It denotes that the educational evaluation began its journey in antiquity, being inaugurated as the epistemological field by Ralph Tyler in the 1930s, until the discussions on the present mass evaluation. As theoretical: Manacorda (1989), given by the contributions to the history of education and Vianna (2000), the constructs outlined in the field of evaluation (history and evaluation models). The research is relevant to the field of education by allowing scour analytically and reflective personal story and little discussion of the educational evaluation.

Key Words: Evaluation. History. Education.

*Recebido em julho de 2013
Aprovado em agosto de 2013*