


Ensaio

Abordagens de avaliação educacional: a constituição do campo teórico no cenário internacional

Educational evaluation approaches: constitution of the theoretical field in the international scenario

Regilson Maciel Borges*¹, José Carlos Rothen**²

*Universidade Federal de Lavras (UFLA), Lavras, MG, Brasil

**Universidade Federal de São Carlos (UFSCar), São Carlos-SP, Brasil

Resumo

Este artigo apresenta uma revisão teórica das principais abordagens de avaliação que marcaram a trajetória da avaliação educacional no cenário internacional. Trata-se de uma pesquisa bibliográfica que, referenciada na literatura da avaliação, busca estudar as variadas dimensões e os diferentes sentidos que constituíram historicamente o campo da avaliação educacional. São descritas sete abordagens de avaliação: a avaliação baseada em objetivos, a avaliação baseada na lógica científica, a avaliação baseada no valor agregado, a avaliação a serviço da decisão, a avaliação orientada para consumidores, a avaliação centrada nos participantes e a avaliação qualitativa. Cada uma dessas abordagens é caracterizada segundo seus protagonistas, objetivos, enfoques conceituais, orientações teóricas decorrentes de pressupostos metodológicos e os estudiosos que deram continuidade nas suas respectivas abordagens. Essa literatura internacional exerceu forte influência sobre os pesquisadores brasileiros da avaliação que, fundamentados nesses referenciais, buscaram sustentar uma área de conhecimento que ainda se encontra em processo de constituição e fortalecimento em nosso país.

Palavras-chave: Avaliação educacional, Abordagens de avaliação, Teorias em avaliação.

Abstract

This article presents a theoretical review of the main evaluation approaches that marked the trajectory of educational evaluation in the international scenario. This is a bibliographical research that, based on the assessment literature, aimed at studying the various dimensions and different meanings that have historically constituted the educational evaluation field. Seven evaluation approaches are described: goal-based evaluation, evaluation based on scientific logic, value-based evaluation, decision-making evaluation, consumer-oriented evaluation, participant-centered evaluation, and qualitative evaluation. Each one of them is characterized according to its protagonists, objectives, conceptual approaches, theoretical orientations deriving from methodological assumptions, and the scholars who continued their respective approaches. Such international literature had a strong influence on Brazilian evaluation researchers who, based on these references, sought to support a field of knowledge that is still in the process of constitution and strengthening in our country.

Keywords: Educational evaluation, Evaluation approaches, Theories on evaluation.

1 Professor do Departamento de Educação da Universidade Federal de Lavras (UFLA). Doutor em Educação pela Universidade Federal de São Carlos (UFSCar) com Pós-Doutorado pela Universidade Estadual de Ponta Grossa (UEPG, Paraná). E-mail: regilsonborges@gmail.com

2 Doutor em Educação pela Universidade Metodista de Piracicaba (UNIMEP) com Pós-Doutorado pela Université de Strasbourg (UNISTRA, França). Professor do Programa de Pós-Graduação em Educação da Universidade Federal de São Carlos (UFSCar). E-mail: jcr3219@yahoo.com.br

Financiamento: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Introdução

Este artigo apresenta uma revisão das principais abordagens que marcaram a trajetória da avaliação educacional no cenário internacional. São analisadas as principais concepções teóricas formuladas por autores que estiveram envolvidos com o desenvolvimento da avaliação, procurando-se apontar os sentidos e as formas que ela adquiriu ao longo do tempo. Trata-se de uma pesquisa bibliográfica que, referenciada na literatura da avaliação, busca estudar as variadas dimensões e os diferentes sentidos que constituíram historicamente esse campo.

A justificativa para um estudo dessa natureza encontra razão no que Stufflebeam e Shinkfield (1987) mencionam ao ressaltarem que nenhuma introdução à avaliação educacional está completa se não prestar devida atenção à evolução histórica do tema. Igualmente, Fernandes (2010, p. 19) entende que “a análise das múltiplas abordagens e das diferentes definições e concepções de avaliação constitui um meio indispensável para se compreender o desenvolvimento e a construção teórica no domínio da avaliação em educação”.

A partir disso buscamos compreender os numerosos elementos que são encontrados nas concepções produzidas por pesquisadores estadunidenses e europeus interessados nas discussões em torno da temática da avaliação educacional. Trata-se de concepções individuais de autores que se esforçam “no sentido de ordenar o conteúdo de um campo novo e parcial em algum tipo de estrutura lógica” (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 237). Para tanto, tomamos como ponto de partida os trabalhos desses estudiosos que analisam diferentes abordagens de avaliação existentes na literatura especializada da área, enfatizando-se as concepções teóricas e os pressupostos metodológicos formulados.

A adoção da designação “abordagens de avaliação” em vez de “modelos de avaliação”, como refere Stufflebeam e Shinkfield (1987), ou “paradigmas da avaliação”, como prefere De Ketele (1993), pareceu-nos mais razoável por considerarmos “que muitas abordagens parecem mais um conjunto de argumentos de persuasão do que construções teóricas sólidas que lhe confirmam o estatuto de *modelo*” (FERNANDES, 2010, p. 18, grifo do autor). Dias Sobrinho (2002) ressalta, por exemplo, que Ernest House prefere usar o termo “enfoque” em vez de “paradigma” para tratar dos tipos básicos de avaliação.

Entre as discussões sobre abordagens de avaliação encontra-se a realizada por Stufflebeam e Shinkfield (1987), que expõem um panorama da avaliação abrangendo oito modelos de trabalho avaliativo. Outro importante trabalho é produzido por De Ketele (1993), que examina como alguns paradigmas de avaliação podem se complementar. Worthen, Sanders e Fitzpatrick (2004) apresentam uma ampla visão de modelos e abordagens de avaliação a partir de uma minuciosa visão sobre suas mais variadas modalidades. Alkin e Christie (2004) elaboram o desenho de uma árvore da teoria da avaliação, cujo tronco é construído sobre uma dupla fundação: a prestação de contas e a investigação social.

Assim, Stufflebeam e Shinkfield (1987) agrupam as ideias e os procedimentos avaliativos em oito “modelos”. São eles:

- 1) avaliação orientada nos objetivos: a tradição de Ralph Tyler;
- 2) Edward A. Suchman e o método científico da avaliação;
- 3) método de planejamento avaliativo de Lee Cronbach;
- 4) Stufflebeam: avaliação orientada para o aperfeiçoamento;
- 5) Stake: método avaliativo centrado no cliente;
- 6) método de contraposição de Thomas R. Owens e Robert L. Wolf;
- 7) avaliação iluminativa: método holístico; e
- 8) método avaliativo de Michael Scriven voltado para o consumidor.³

De Ketele (1993), por sua vez, classifica a avaliação dentro de dez “paradigmas”:

- 1) paradigma da intuição pragmática;
- 2) paradigma docimológico;
- 3) paradigma sociológico;
- 4) paradigma da avaliação centrada nos objetivos;
- 5) paradigma da avaliação formativa no ensino diferenciado;
- 6) paradigma da avaliação a serviço da decisão;
- 7) paradigma da avaliação centrada no consumidor segundo Scriven;
- 8) paradigma da avaliação centrada no cliente ou paradigma da avaliação responsiva segundo Stake;
- 9) paradigma econômico; e
- 10) paradigma da avaliação como processo de regulação.⁴

Worthen, Sanders e Fitzpatrick (2004) classificam as muitas abordagens de avaliação em seis “categorias”:

- 1) abordagem centrada em objetivos;
- 2) abordagem centrada na administração;

3 No original: “La evaluación orientada hacia los objetivos: la tradición tyleriana; Edward A. Suchman y el método científico de evaluación; La planificación evaluativa según Cronbach: una sinopsis; Stufflebeam: la evaluación orientada hacia el perfeccionamiento; Stake: el método evaluativo centrado en el cliente; T. R. Owens y R. L. Wolf: un método contrapuesto de evaluación; La evaluación iluminativa: el método holístico; El método evaluativo de Scriven orientado hacia el consumidor”.

4 No original: “Le paradigme de l’intuition pragmatique; Le paradigme docimologique; Le paradigme sociologique; Le paradigme de l’évaluation centrée sur les objectifs; Le paradigme de l’évaluation formative dans un enseignement différencié; Le paradigme de l’évaluation au service de la décision; Le paradigme de l’évaluation centrée sur le consommateur selon Scriven; Le paradigme de l’évaluation centrée sur le client ou paradigme de l’évaluation répondante selon Stake; Le paradigme économique; Le paradigme de l’évaluation comme processus de régulation”.

- 3) abordagem centrada no consumidor;
- 4) abordagem centrada em especialistas;
- 5) abordagem centrada no adversário; e
- 6) abordagem centrada nos participantes.

Alkin e Christie (2004) colocam no topo de sua árvore da teoria da avaliação três ramos e suas respectivas ênfases teóricas. São elas:

- 1) metodologia: Ralph Tyler, Donald Campbell, Edward Suchman, Thomas Cook, Robert Boruch, Peter Rossi, Huey-Tsyh Chen, Carol Weiss e Lee J. Cronbach;
- 2) utilização: Daniel Stufflebeam, Malcolm Provus, Joseph Wholey, Michael Patton, Marvin Alkin, J. Bradley Cousins, Hallie Preskill, Jean King, John M. Owen e David Fetterman;
- 3) valores: Michael Scriven, Elliot Eisner, Thomas Owens/Robert Wolf, Robert Stake, Barry MacDonald, Ernest House, e Egon Guba e Yvonna Lincoln.

Com base nesses estudos, descreve-se a seguir sete abordagens de avaliação educacional que resultam do cruzamento das informações apresentadas pelos autores mencionados, de onde também são extraídas as designações que nomeiam cada uma das abordagens apresentadas. As sete abordagens são caracterizadas segundo seus protagonistas, objetivos, enfoques conceituais, orientações teóricas decorrentes de pressupostos metodológicos e os estudiosos que deram continuidade nas suas respectivas abordagens.

Abordagem de avaliação baseada em objetivos

Essa abordagem resulta do trabalho realizado por Ralph W. Tyler durante o “Estudo de oito anos”, desenvolvido na Universidade de Ohio, nos Estados Unidos. Nessa abordagem a avaliação consiste em determinar em que medida os objetivos educacionais foram alcançados.

A principal recomendação de Tyler era a de que o professor, o administrador e os que conduziam os currículos emitissem juízos racionais sobre as áreas de programas que estivessem sob sua direção. Tyler considerava que a avaliação deveria ser uma etapa essencial do desenvolvimento do programa. Ponderava que precisaria existir uma série de etapas lógicas que promovessem o processo de avaliação real (STUFFLEBEAM; SHINKFIELD, 1987).

Assim, sua abordagem de avaliação deveria obedecer aos seguintes procedimentos:

1. Estabelecer metas ou objetivos gerais.
2. Classificar as metas ou objetivos.
3. Definir os objetivos em termos comportamentais.
4. Encontrar situações em que é possível mostrar que os objetivos foram alcançados.

5. Criar ou selecionar técnicas de mensurações.
6. Coletar dados relativos ao desempenho.
7. Comparar os dados do desempenho com os objetivos formulados comportamentalmente. (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 130)

De Ketele (1993) observa que a avaliação proposta por Tyler deveria ser realizada no final do processo de formação, a fim de colocar em evidência se os alunos dominaram os objetivos. A informação resultante do processo possibilitaria um duplo atendimento: a certificação dos alunos e a regulação do processo de formação.

Como ampliação do método tyleriano, uma série de trabalhos enfatizou a metodologia centrada nos objetivos. Stufflebeam e Shinkfield (1987) e Worthen, Sanders e Fitzpatrick (2004) incluem nessa abordagem a proposta de Newton S. Metfessel e William B. Michael, que desenvolveram um paradigma avaliativo de oito etapas com o propósito de avaliar a consecução dos objetivos nos currículos escolares. Os oito passos para o processo de avaliação são:

1. Envolver os interessados como facilitadores da avaliação do programa.
2. Formular um modelo coerente de metas e objetivos específicos.
3. Traduzir objetivos específicos de forma comunicável.
4. Selecionar ou construir instrumentos para fornecer medidas que permitam inferências sobre a efetividade do programa.
5. Realizar observações periódicas usando testes de conteúdo válidos, escalas e outras medidas comportamentais.
6. Analisar os dados usando métodos apropriados.
7. Interpretar os dados usando modelos de níveis desejados de desempenho de preferência sobre todas as outras medidas.
8. Fazer recomendações para implementação, modificação e revisão posteriores de metas gerais e objetivos específicos. (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 133-134)

Ainda nessa abordagem, Worthen, Sanders e Fitzpatrick (2004, p. 136) incluem o modelo de avaliação da discrepância desenvolvido por Malcom Provus “para ajudar os administradores a determinar a medida em que o desenvolvimento do programa está avançando no sentido de alcançar os objetivos estabelecidos”. O modelo busca eliminar todas as discrepâncias, de modo que, quando se detecta uma delas, é sugerido um processo cooperativo de resolução do problema, que envolve as seguintes perguntas: “1) por que existe uma discrepância? 2) que providências é possível tomar para corrigi-las? e 3) qual providência corretiva é a mais indicada?” (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 136).

A principal característica dessa abordagem reside no grau em que os objetivos foram alcançados. Alkin e Christie (2004) mencionam outros teóricos que expandiram o trabalho de Tyler, entre os quais se encontram Benjamin Bloom, com suas taxonomias de objetivos educacionais; Robert L. Hammond, que apresenta um

sistema para gerar e analisar objetivos; James Popham, que se concentra na defesa das especificações dos objetivos comportamentais; e Andrés *Steinmetz*, que discute o modelo de avaliação da discrepância.

Abordagem de avaliação baseada na lógica científica

Essa abordagem é marcada pela forte defesa do experimento. Stufflebeam (2001) cita Donald T. Campbell e Julian C. Stanley como os pioneiros no uso da experimentação na avaliação de programas. Seu foco se concentra na tentativa de descartar ameaças e impedir inferências casuais na condução de pesquisas.

Campbell e Stanley (1963) defendem a experiência como o único meio para resolver os problemas relacionados à prática educativa. Os autores creditam à experiência a possibilidade de verificar as melhorias educacionais e indicar os caminhos que pudessem ser introduzidos sem prejudicar a tradição acumulada, evitando o descarte da sabedoria antiga em detrimento de novidades inferiores. Os autores também defendem a formação de pesquisadores educacionais em estatística experimental, o que elevaria a qualidade na experimentação educacional.

A avaliação nessa abordagem enfatiza a experimentação como meio de aperfeiçoar a relevância dos testes, principalmente no processo de seleção. De acordo com Stufflebeam (2001), essa abordagem é marcada por uma estratégia que se inicia com uma série de perguntas, as quais procuram avaliar o mérito e o valor de um programa. Os métodos típicos usados nesse processo são os projetos experimentais e quase-experimentais.

Campbell e Stanley (1963) apontam a distinção entre validade interna e validade externa na realização de um estudo experimental verdadeiro. Para isso apresentam oito diferentes classes de ameaças à validade interna e quatro fatores que comprometem a validade externa do experimento. As variáveis que comprometem a validade interna são:

1. *História*, os eventos específicos que ocorrem entre a primeira e segunda medição para além da variável experimental.
2. *Maturação*, os processos que operam dentro dos respondentes em função da passagem do tempo por si só (não é específica para eventos particulares), incluindo o envelhecimento, o aumento da fome, do cansaço e afins.
3. *Os testes*, os efeitos de fazer um teste sobre a pontuação de um segundo teste.
4. *Instrumentação*, onde as mudanças na calibração de um instrumento de medição ou alterações nos observadores ou marcadores utilizados podem produzir alterações nas medidas obtidas.
5. *Regressão estatística*, operacional em que grupos foram selecionados com base nas suas pontuações extremas.
6. Os vieses, resultando em *seleção* diferencial dos entrevistados para os grupos de comparação.
7. *Mortalidade experimental*, ou perda diferencial dos entrevistados dos grupos de comparação.

8. *Interação seleção-maturação* etc., que em alguns dos projetos quase-experimentais de grupos múltiplos, como o Projeto 10, pode ser confundida, ou seja, tomada erroneamente pelo efeito da variável experimental.⁵ (CAMPBELL; STANLEY, 1963, p. 5, tradução nossa, grifo dos autores)

Já os fatores que comprometem a validade externa, segundo Campbell e Stanley (1963), são:

9. *O efeito reativo ou interativo dos testes*, em que um pré-teste pode aumentar ou diminuir a sensibilidade ou a capacidade de resposta do entrevistado para a variável experimental e, assim, fazer com que os resultados obtidos para uma população pré-testada não sejam representativos dos efeitos da variável experimental para o universo pré-testado a partir do qual os entrevistados experimentais foram selecionados.

10. *Os efeitos da interação de vieses de seleção e a variável experimental*.

11. *Efeitos reativos de arranjos experimentais*, o que impediria generalização do efeito da variável experimental sobre as pessoas sendo expostas a ela em ambientes não experimentais.

12. *Interferência de múltiplos tratamentos*, que pode ocorrer quando vários tratamentos são aplicados aos mesmos respondentes, uma vez que os efeitos de tratamentos anteriores geralmente não podem ser apagados.⁶ (CAMPBELL; STANLEY, 1963, p. 5-6, tradução nossa, grifo dos autores)

Alkin e Christie (2004, p. 20, tradução nossa) ressaltam que as ideias de Campbell e Stanley se encontram presentes na “fundação de quase todos os cursos sobre métodos de pesquisa em ciências sociais”⁷, principalmente as ideias relacionadas à formulação de projetos de pesquisa. A introdução de projetos quase-experimentais na investigação favoreceu um clima maior de aceitação de um pensamento alternativo para as ciências sociais.

Outro autor que defende a aplicação de métodos de pesquisa na avaliação, ou seja, que a avaliação se fundamente na lógica do método científico, é o sociólogo Edward Suchman (STUFFLEBEAM, 2001; ALKIN; CHRISTIE, 2004). Para Suchman, a avaliação é um processo científico e, deste modo, os mesmos procedimentos

5 No original: “1. History, the specific events occurring between the first and second measurement in addition to the experimental variable. 2. Maturation, processes within the respondents operating as a function of the passage of time per se (not specific to the particular events), including growing older, growing hungrier, growing more tired, and the like. 3. Testing, the effects of taking a test upon the scores of a second testing. 4. Instrumentation, in which changes in the calibration of a measuring instrument or changes in the observers or scorers used may produce changes in the obtained measurements. 5. Statistical regression, operating where groups have been selected on the basis of their extreme scores. 6. Biases resulting in differential selection of respondents for the comparison groups. 7. Experimental mortality, or differential loss of respondents from the comparison groups. 8. Selection-maturation interaction, etc., which in certain of the multiple-group quasi-experimental designs, such as Design 10, is confounded with, i.e., might be mistaken for, the effect of the experimental variable”.

6 No original: “9. The reactive or interaction effect of testing, in which a pretest might increase or decrease the respondent’s sensitivity or responsiveness to the experimental variable and thus make the results obtained for a pretested population unrepresentative of the effects of the experimental variable for the unpretested universe from which the experimental respondents were selected. 10. The interaction effects of selection biases and the experimental variable. 11. Reactive effects of experimental arrangements, which would preclude generalization about the effect of the experimental variable upon persons being exposed to it in non experimental settings. 12. Multiple-treatment interference, likely to occur whenever multiple treatments are applied to the same respondents, because the effects of prior treatments are not usually erasable”.

7 No original: “foundation of almost all social science research methods courses”.

utilizados para descobrir o conhecimento poderiam ser usados no processo avaliativo a fim de se obter resultados mais objetivos (STUFFLEBEAM; SHINKFIELD, 1987). Contudo, Suchman distingue a avaliação da pesquisa avaliativa: “a primeira se referia em geral a um ‘processo de emissão de juízos de valor’, enquanto considerava que a pesquisa avaliativa consistia nos ‘procedimentos de coleta e análise de dados’”⁸ (STUFFLEBEAM; SHINKFIELD, 1987, p. 112, tradução nossa).

Nessa abordagem se incluem também outros estudiosos que defendem um método de avaliação baseado na experimentação e centrado na medida e na objetividade, entre os quais se encontram Lee Cronbach, Everett F. Lindquist, Richard E. Snow, Robert F. Boruch, Gene V. Glass, Thomas O. Maguire, David E. Wiley, R. Darrell Bock (STUFFLEBEAM, 2001), *Otto Klineberg*, George James, Elizabeth *Herzog* e Arthur C. Fleck (STUFFLEBEAM; SHINKFIELD, 1987).

Abordagem de avaliação baseada no valor agregado

Essa abordagem foi inicialmente desenvolvida por William L. Sanders e Sandra P. Horn. Seus métodos e teorias foram testados pelos autores no estado norte-americano do Tennessee.

A abordagem de valor agregado tem a intenção de determinar o valor com que cada escola contribui para as conquistas dos alunos atendidos, para em seguida “relatar os resultados para fins de política, de responsabilização e de melhoria”⁹ (STUFFLEBEAM, 2001, p. 23, tradução nossa).

Sanders e Horn (1994) descrevem a metodologia de modelos mistos em avaliação educacional resultantes do Sistema de Avaliação de Valor Agregado do Tennessee (TVAAS). Segundo os autores, o “TVAAS é um processo estatístico que fornece medidas da influência que os sistemas de ensino, a escola e os professores têm sobre os indicadores de aprendizagem dos alunos”¹⁰ (SANDERS; HORN, 1994, p. 301, tradução nossa).

O método requer três componentes:

(...) um processo de teste que produz escalas fortemente relacionadas ao currículo e medições que se prologam acima e abaixo do nível de ensino; a construção e expansão contínua de uma base de dados longitudinal; e um processo estatístico que permite uma análise longitudinal multivariada para produzir estimativas imparciais e eficientes dos efeitos desejados.¹¹ (SANDERS; RIVERS, 1996, p. 1, tradução nossa)

O TVAAS possibilitou a criação de um banco de dados com informações sobre o nível de sistema de ensino do Tennessee (cerca de 3 milhões de registros), fornecendo acesso às medidas individuais dos estudantes sobre matemática, ciência, leitura, linguagem e estudos sociais, que foram organizadas após aplicação de teste de desempenho (SANDERS; RIVERS, 1996). Para Stufflebeam (2001), essa experiência

8 No original: “A la primera se refería por lo general como un ‘proceso de emitir juicios de valor’, mientras que consideraba que la investigación evaluativa consistía en los ‘procedimientos de recopilación y análisis de datos’”.

9 No original: “report the results for policy, accountability, and improvement purposes”.

10 No original: “TVAAS is a statistical process that provides measures of the influence that school systems, schools, and teachers have on indicators of student learning”.

11 No original: “a testing process which produces scales that have a strong relationship to the curriculum and which produces measurement that extends above and below grade level; the construction and ongoing expansion of a longitudinal data base; and a statistical process that enables a multivariate, longitudinal analysis to produce unbiased and efficient estimates of the desired effects”.

do Tennessee poderia levar as Secretarias Estaduais de Educação a coletar, anualmente, dados de todos os estudantes (em uma sucessão de níveis de classificação). Assim, o avaliador teria condições de analisar os dados e fazer comparações entre áreas curriculares, professores, escolas de ensino fundamental e médio, recursos, entre outros aspectos do processo de ensino e aprendizagem.

Stufflebeam (2001, p. 24, tradução nossa) menciona que “a vantagem central do monitoramento de resultados envolvendo avaliações baseadas no valor agregado está na sistematização e na institucionalização de um banco de dados de resultados que pode ser usado ao longo do tempo e de modo padronizado para estudar e encontrar meios de melhorar os resultados”¹².

Ainda de acordo com Stufflebeam (2001), a avaliação baseada no valor agregado trouxe uma vantagem especial para a utilização de testes padronizados para avaliar efeitos de programas e políticas, pois o uso eficiente dos testes, nesse caso, possibilita uma análise das tendências dos diferentes componentes de um sistema de ensino (níveis de sala de aula, grupos de escolas, escolas individuais e professores individuais). Essa abordagem se diferencia do típico programa de testes padronizados pela “ênfase na pontuação de ganhos sofisticada e na análise hierárquica de dados para delinear efeitos dos componentes do sistema e identificar quais devem ser melhorados e quais devem ser elogiados e reforçados”¹³ (STUFFLEBEAM, 2001, p. 23, tradução nossa).

Outros autores que discutem a abordagem avaliativa baseada no valor agregado são William J. Webster, Robert L. Mendro, Ted O. Almaguer, que utilizam o valor agregado para medir o efeito na escola; e Peter Tymms, que propõe a criação de um sistema nacional de valor agregado para o ensino primário na Inglaterra (STUFFLEBEAM, 2001).

Abordagem de avaliação a serviço da decisão

Daniel Stufflebeam propôs uma abordagem de avaliação centrada no processo de tomada de decisões, cujo objetivo é melhorar, isto é, tomar decisões adequadas (DE KETELE, 1993). O autor define a avaliação “como um processo que irá fornecer informações úteis para a tomada de decisões”¹⁴ (STUFFLEBEAM; SHINKFIELD, 1987, p. 179, tradução nossa).

A fim de ajudar administradores e gerentes a tomar decisões acertadas, o autor desenvolveu uma estrutura de avaliação esquematizada na sigla CIPP (Contexto, Insumo, Processo e Produto), que corresponde a um quadro de avaliação baseado em quatro tipos diferentes de decisões, que são assim definidos:

1. *A avaliação do contexto serve para decisões de planejamento.* Determinar que necessidades precisam ser atendidas por um programa ajuda a definir seus objetivos.
2. *A avaliação do insumo serve para decisões de estrutura.* Determinar que

12 No original: “The central advantage of outcome monitoring involving value-added assessment is in the systematization and institutionalization of a database of outcomes that can be used over time and in a standardized way to study and find means to improve outcomes”.

13 No original: “emphasis on sophisticated gain score and hierarchical analysis of data to delineate effects of system components and identify which ones should be improved and which ones should be commended and reinforced”.

14 No original: “como un proceso mediante el cual se proporcionará información útil para la tomada de decisiones”.

recursos estão disponíveis, que estratégias alternativas do programa devem ser consideradas e que plano parece ter o maior potencial para atender as necessidades e facilitar a criação de procedimentos do programa.

3. *A avaliação do processo serve para decisões de implementação.* Qual é o grau de êxito da implementação do programa? Que barreiras ameaçam seu sucesso? Que revisões são necessárias? Depois que essas perguntas foram respondidas, será possível supervisionar, controlar e refinar os procedimentos.

4. *A avaliação do produto serve para decisões de reciclagem.* Que resultados foram obtidos? Até que ponto as necessidades foram reduzidas? O que deverá ser feito com o programa depois que ele chegar a seu termo? Essas perguntas são importantes para julgar o que o programa conseguiu realizar. (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 152-153, grifo dos autores)

De acordo com Stufflebeam e Shinkfield (1987, p. 183, tradução nossa), o modelo CIPP postula três finalidades para a avaliação, que são: “orientar a tomada de decisão, fornecer dados para a prestação de contas e promover a compreensão dos fenômenos envolvidos”¹⁵. Nesse modelo, a avaliação é apresentada não como um julgamento, mas como um processo que inclui três etapas, sendo elas: a identificação, a obtenção e o fornecimento de informações.

Sendo assim, na perspectiva da avaliação a serviço da tomada de decisão, os avaliadores devem projetar um processo, e não um produto (ALKIN; CHRISTIE, 2004). Stufflebeam e Shinkfield (1987, p. 204, tradução nossa) mencionam que “para saber como uma avaliação [...], o avaliador precisa, obviamente, planejar o trabalho que se deve fazer. Isso supõe preparar planos preliminares e, em seguida, modificá-los e explicá-los à medida que o estudo progride”¹⁶. Diante disso, Stufflebeam propôs aos avaliadores os seguintes passos:

A. *Concentrar-se na avaliação;*

B. *Coleta de informações;*

C. *Organização das informações;*

D. *Análise das informações;*

E. *Relatório das informações;* e

F. *Gestão da avaliação.* (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 153-155, grifo dos autores)

Um dos grandes méritos do modelo de Stufflebeam, segundo De Ketele (1993), é ser um modelo global em termos de uma avaliação global de formação. Worthen, Sanders e Fitzpatrick (2004) mencionam outros autores que fizeram contribuições importantes para a abordagem de avaliação orientada para a melhoria, entre os quais se encontram Marvin C. Alkin, Lee Cronbach, Michael Patton, Charles S. Reichardt e Joseph S. Wholey.

15 No original: “servir de guía para la toma de decisiones, proporcionar datos para para la responsabilidad y promover la comprensión de los fenómenos implicados”.

16 No original: “para saber cómo una evaluación [...] el evaluador necesita, obviamente, planificar el trabajo que debe hacer. Esto supone preparar los planes preliminares y luego modificarlos y explicarlos a medida que avance el estudio”.

Abordagem de avaliação orientada para consumidores

Michael Scriven foi um dos pioneiros na abordagem da avaliação para os consumidores. Segundo Scriven, os consumidores são, por um lado, “os que recebem as pessoas formadas e as ‘utilizam’; e, por outro, [...] os beneficiários da formação”¹⁷ (DE KETELE, 1993, p. 71, tradução nossa).

Stufflebeam e Shinkfield (1987) destacam que Scriven contribuiu significativamente para o desenvolvimento do campo profissional da avaliação. Sua definição de avaliação é centrada no problema do valor, considerando que “a avaliação desempenha muitos papéis, mas que tem uma única *meta*: determinar o valor ou mérito de alguma coisa que esteja sendo avaliada” (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 39, grifo dos autores).

Nessa abordagem o avaliador procura ajudar os consumidores a identificar e avaliar o mérito e o valor de programas concorrentes, serviços e produtos (STUFFLEBEAM, 2001). Assim, no modelo de Scriven, o avaliador tem que deixar o espírito livre de qualquer objetivo. Sua primeira tarefa consiste em reunir o máximo de informações sobre todos os componentes do sistema de avaliação, utilizando-se para isso de diferentes instrumentos de coleta. A tarefa seguinte é entregar um juízo de valor e se pronunciar sobre a avaliação realizada (DE KETELE, 1993).

Scriven (1967) também contribuiu com essa abordagem ao destacar as duas principais funções da avaliação: a formativa e a somativa. A avaliação formativa é uma parte integrante do processo de desenvolvimento: ela “fornece informação contínua para ajudar a planejar e, em seguida, produzir um objeto”, enquanto a avaliação somativa “investiga todos os efeitos dos objetos e os examina, comparando-os com as necessidades dos consumidores”¹⁸ (STUFFLEBEAM; SHINKFIELD, 1987, p. 345, tradução nossa). A formativa ajuda no desenvolvimento de programas dando “informações úteis para a melhoria do programa”, e a somativa “leva a decisões relativas à continuidade do programa, a seu encerramento, a sua ampliação, adoção, etc.” (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 48).

Scriven lista alguns critérios que devem ser observados na avaliação de um produto:

- Evidência de ter alcançado importantes objetivos.
- Evidência de haver alcançado importantes objetivos não educacionais (como objetivos sociais).
- Resultados do acompanhamento.
- Efeitos secundários e imprevistos, como efeitos sobre o professor, nos colegas do professor, em outros alunos, administradores, pais, na escola, no contribuinte, e outros efeitos acidentais positivos ou negativos.
- Alcance da utilidade (isto é, para quem o produto vai ser útil).
- Considerações morais (usos injustos de punição ou conteúdo controvertido).

17 No original: “la clientèle ou marche-cible qui recevront les personnes formées et les ‘utiliseront’ d’une part; les vrais consommateurs d’autre part, c’est-à-dire les récipiendaires de la formation”.

18 No original: “Proporciona información continua para ayudar a planificar y luego producir algún objeto”; “investiga todos los efectos de los objetos y los examina comparándolos con las necesidades de los consumidores”.

– Custos. (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 167)

Mais tarde, Worthen, Sanders e Fitzpatrick (2004, p. 167-168) observam que Scriven ampliou os seus critérios com uma “lista de verificação de produtos”. Essa lista inclui os seguintes tópicos:

1. Necessidade;
2. Mercado;
3. Desempenho – experimentos de campo autênticos;
4. Desempenho – consumidor autêntico;
5. Desempenho – comparações críticas;
6. Desempenho – longo prazo;
7. Desempenho – efeitos colaterais;
8. Desempenho – processo;
9. Desempenho – modelo casual;
10. Desempenho – relevância estatística;
11. Desempenho – relevância educacional;
12. Custo-benefício; e
13. Ampliação da base de apoio.

Stufflebeam e Shinkfield (1987, p. 348-351) apresentam dezoito pontos que compõem “a lista de verificação dos indicadores de avaliação”, que sintetizam as ideias de Scriven. São eles:

1. Descrição;
2. Cliente;
3. Antecedentes e contexto;
4. Recursos;
5. Função;
6. Sistema de distribuição;
7. Consumidor;
8. Necessidades e valores dos afetados e potencialmente impactados;
9. Normas;
10. Processo;

11. Resultados;
12. Possibilidade de generalização;
13. Custos;
14. Comparações;
15. Significado;
16. Recomendações;
17. Relatório; e
18. Meta-avaliação.¹⁹

Stufflebeam (2001) menciona que uma das principais vantagens da avaliação orientada para consumidores é sua independência e objetividade no sentido de proteger os consumidores de programas, serviços e produtos de má qualidade. Sua desvantagem reside justamente no seu caráter independente/objetivo, de modo que “ela pode ser tão independente dos profissionais que pode não lhes ajudar em servir melhor os consumidores”²⁰ (STUFFLEBEAM, 2001, p. 60, tradução nossa).

Um autor que, segundo Stufflebeam (2001), contribuiu para os estudos de metodologia dessa abordagem de avaliação orientada para consumidores desenvolvida por Scriven foi Gene V. Glass, pesquisador do Centro de Política Nacional de Educação da Universidade do Colorado. Glass destacou que as avaliações deveriam concentrar-se no julgamento sobre o mérito e o valor de um programa. Outros autores que contribuíram com essa abordagem foram Komoski, Morissett e Stevens, que elaboraram listas de verificação e diretrizes para a análise de currículo (WORTHEN; SANDERS; FITZPATRICK, 2004).

Abordagem de avaliação centrada nos participantes

Assim como Scriven está para a avaliação orientada para os consumidores, Robert E. Stake está para a abordagem da avaliação centrada nos participantes (ou clientes), pois foi ele, segundo Worthen, Sanders e Fitzpatrick (2004, p. 225), “o primeiro teórico a dar impulso significativo a essa abordagem no campo da educação”.

De Ketele (1993) assinala que Stake tentou equilibrar o pensamento de Stufflebeam – ao enfatizar a necessidade de avaliar contexto, insumos, processos e produtos – com o pensamento de Scriven, estando de acordo que os juízos de valores devem estar incluídos na avaliação. Stake (1975a, p. 14, tradução nossa) acredita que não se trata de uma abordagem nova, pois “é uma avaliação baseada no que as pessoas naturalmente fazem para avaliar as coisas: elas observam e reagem”²¹.

Stake (1975a, p. 14, tradução nossa, grifo do autor) assim definiu sua perspectiva de avaliação, denominando-a de avaliação responsiva:

19 No original: “la lista de control de indicadores de la evaluación [...] 1. Descripción; 2. El cliente; 3. Antecedentes y contexto; 4. Recursos; 5. Función; 6. Sistema de distribución; 7. El consumidor; 8. Las necesidades y valores de los afectados y potencialmente impactados; 9. Normas; 10. El proceso; 11. Resultados; 12. Posibilidad de generalización; 13. Costes; 14. Comparaciones; 15. Significado; 16. Recomendaciones; 17. El informe; 18. La metaevaluación”.

20 No original: “it can be so independent from practitioners that it may not assist them to better serve consumers”.

21 No original: “It is evaluation based on what people do naturally to evaluate things: they observe and react”.

Uma avaliação educacional é uma *avaliação responsiva* se for mais diretamente orientada para as atividades do programa do que para suas intenções, se responde às exigências de informação do público e se as diferentes perspectivas de valor das pessoas presentes são referidas ao relatar o sucesso ou o fracasso do programa²².

A estrutura da avaliação responsiva segue os principais componentes do programa educacional. A escolha dos dispositivos de coleta de dados é feita com base na observação da ação do programa e na interação com os diversos grupos interessados. A avaliação responsiva é particularmente útil durante a avaliação formativa e somativa. Seus principais estímulos são aqueles do programa, incluindo as respostas dos alunos e os diálogos posteriores. A abordagem tenta responder às formas naturais com que as pessoas assimilam informações e chegam à compreensão (STAKE, 1975a; 1975b).

Worthen, Sanders e Fitzpatrick (2004, p. 231) salientam que o objetivo fundamental da avaliação responsiva é “a melhoria da comunicação com os interessados”; desse modo, “o objetivo, quadro de referência e foco de uma avaliação responsiva surgem das interações com o público, e essas interações e observações resultam na focalização progressiva de questões”. Trata-se de uma ênfase na comunicação natural, menos dependente de uma comunicação formal (STAKE, 1975a; 1975b).

Nesse processo o avaliador deve trabalhar com o apoio de um grupo de clientes diversificado, incluindo professores, administradores, legisladores, administradores e outros (STUFFLEBEAM, 2001). Stake (1975a, p. 14, tradução nossa) assim descreve o papel do avaliador:

Para fazer uma avaliação responsiva, o avaliador, claro, faz muitas coisas. Ele faz um plano de observações e negociações e o organiza para que várias pessoas observem o programa. Com a ajuda dessas pessoas, ele prepara narrativas breves, retratos, exibições de produtos, gráficos, etc. Ele descobre o que seu público valoriza. Ele reúne expressões de valor de vários indivíduos cujos pontos de vista diferem. Obviamente, ele verifica a qualidade de seus registros. Ele recebe o pessoal do programa para observar sua reação à precisão de seus retratos. Ele recebe as figuras de autoridade para observar sua reação à importância das várias descobertas. Ele recebe os membros do público interessado para observar sua reação à relevância de tais descobertas. Ele faz muitas dessas coisas informalmente, iterando e mantendo um registro de ação e reação. Ele escolhe meios acessíveis às suas audiências para aumentar a probabilidade e fidelidade da comunicação. Ele pode preparar um relatório final escrito – ou não, dependendo do que ele e seus clientes tenham acordado.²³

22 No original: “An educational evaluation is responsive evaluation if it orients more directly to program activities than to program intents, if it responds to audience requirements for information, and if the different value perspectives of the people at hand are referred to in reporting the success and failure of the program”.

23 No original: “To do a responsive evaluation, the evaluator of course does many things. He makes a plan of observations and negotiations. He arranges for various persons to observe the program. With their help he prepares for brief narratives, portrayals, product displays, graphs, etc. He finds out what is of value to his audiences. He gathers expressions of worth from various individuals whose points of view differ. Of course, he checks the quality of his records. He gets program personnel to react to the accuracy of his portrayals. He gets authority figures to react to the importance of various findings. He gets audience members to react to the relevance of his findings. He does much of this informally, iterating, and keeping a record of action and reaction. He chooses media accessible to his audiences to increase the likelihood and fidelity of communication. He might prepare a final written report; he might not – depending on what he and his clients have agreed on”.

Essas perspectivas apontadas pelo autor supõem uma troca contínua entre o avaliador e os clientes. O enfoque está na comunicação contínua entre avaliador e público para descobrir, investigar e tratar dos problemas do programa. Os projetos de avaliações são relativamente informais e abertos, e são construídos ao longo do processo. Assim, dependendo do propósito da avaliação, o avaliador pode empregar diferentes variedades de métodos (STUFFLEBEAM, 2001); por isso “é mais provável que se sintam à vontade trabalhando com o paradigma naturalista ou etnográfico, fundamentando-se principalmente em técnicas qualitativas” (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 232). Entre as técnicas preferidas para o retrato dos programas, estão o *feedback*, o estudo de caso, relatórios de amostragem, relatórios de observação, relatórios descritivos, sociodrama, o relato de histórias, etc.

Essa abordagem foi também desenvolvida e ampliada por outros estudiosos, conforme observa Stufflebeam (2001), dentre os quais se destacam Terry Denny, Barry MacDonald, Malcolm Parlett, David Hamilton, Robert M. Rippey, Louis M. Smith e Paul A. Pohland. Stufflebeam (2001) ressalta ainda a forte influência da avaliação responsiva na proposta construtivista de Egon G. Guba e Yvonna S. Lincoln, cuja ênfase recai sobre a história pessoal e perspectivas valorativas das partes interessadas.

Abordagem de avaliação qualitativa

A preocupação das abordagens de avaliação qualitativa está em como são utilizados os resultados da avaliação e no desenvolvimento de métodos que colocam poucas restrições sobre o avaliador. “Conhecidas como *naturalistas* ou *qualitativas*, essas abordagens dão liberdade para os avaliadores escolherem os métodos usados para coletar, analisar e interpretar seus dados”²⁴ (OWSTON, 2008, p. 607-608, grifo do autor, tradução nossa).

Exemplos da abordagem qualitativa da avaliação são encontrados nos trabalhos de Egon G. Guba e Yvonna S. Lincoln (2011), Barry MacDonald (1982), Ernest House (HOUSE; HOWE, 2000) e Malcolm Parlett e David Hamilton (1982).

Guba e Lincoln (2011) propõem uma abordagem alternativa à avaliação que denominam de avaliação de quarta geração, também conhecida como avaliação construtivista (STUFFLEBEAM, 2001). Seu foco encontra-se na negociação, estando fundamentada em um enfoque responsivo e em uma metodologia construtivista. Essa abordagem entende que “os resultados da avaliação representam construções significativas que atores específicos ou vários grupos de atores formam para ‘atribuir sentido’ às situações nas quais eles se encontram” (GUBA; LINCOLN, 2011, p. 14).

Os autores definem a avaliação de quarta geração como um

(...) modelo de avaliação em que as reivindicações, preocupações e questões dos grupos de interesse ou interessados (*stakeholders*) servem como enfoques organizacionais (princípio para determinar quais informações são essenciais), que são implementados de acordo com os preceitos metodológicos do paradigma de investigação construtivista. (GUBA; LINCOLN, 2011, p. 54, grifo dos autores)

24 No original: “Known as naturalistic or qualitative, these approaches give the evaluator freedom to choose the methods used to collect, analyze, and interpret their data”.

De acordo com seus elaboradores, o avaliador nessa abordagem é responsável pelos seguintes procedimentos na aplicação da metodologia da avaliação de quarta geração:

(1) Formalizar contrato com o cliente ou o patrocinador que encomendou a avaliação; (2) Organizar a avaliação; (3) Identificar os grupos de interesse ou interessados (*stakeholders*); (4) Gerar uma construção conjunta (colaborativa, compartilhada) dentro do grupo de interesse por meio do processo do círculo hermenêutico-dialético, enfatizando especificamente as RPQs [Reivindicações, Preocupações e Questões]; (5) Testar e ampliar as construções internas ao grupo introduzindo informações novas ou complementares e possibilitando que os membros do grupo alcancem níveis de esclarecimento superiores ao lidar com essas informações; (6) Separar as RPQs resolvidas – aquelas nas quais se obteve consenso; (7) Priorizar as RPQs ainda não resolvidas; (8) Coletar informações relacionadas às RPQs não resolvidas; (9) Preparar uma agenda de negociação; (10) Conduzir a negociação; (11) Utilizar estudo de caso como relato – a construção conjunta como produto; (12) Reciclar. (GUBA; LINCOLN, 2011, p. 206, grifo dos autores)

Segundo Guba e Lincoln (2011, p. 283-286), as principais consequências de seguir o modelo da avaliação de quarta geração são que o absolutismo paroquial cede lugar ao relativismo ecumênico; a responsabilidade dá lugar à responsabilidade compartilhada; a espoliação dá lugar à concessão de poder; a ignorância dá lugar à compreensão e à apreciação; e a imobilização dá lugar à ação.

Outro enfoque da abordagem de avaliação qualitativa são as tradições de orientação democrática na avaliação, que tiveram sua origem na formulação de Barry MacDonald (1982) para o campo da educação na Inglaterra e no compromisso com a avaliação democrática e deliberativa desenvolvida nos Estados Unidos por Ernest House (HOUSE; HOWE, 2000).

O modelo de avaliação democrática desenvolvido por MacDonald (1982, p. 17) “é um serviço informativo prestado à comunidade acerca das características de um programa educacional. Ela reconhece a existência de um pluralismo de valores”. Sua metodologia principal é o estudo de caso. Nessa abordagem, o avaliador age como um intermediário cuja principal atividade é o “levantamento das concepções e reações acerca do programa estudado”, com total sigilo (MACDONALD, 1982, p. 17). O relatório avaliativo não apresenta recomendações, seu critério de sucesso é o alcance do estudo avaliativo, medido em termos das audiências beneficiadas pelo estudo. O conceito fundamental que justifica a avaliação democrática é o “direito à informação” (MACDONALD, 1982, p. 17).

A avaliação democrática e deliberativa de House e Howe (2000) defende que a avaliação deve contribuir para o avanço da democracia em uma sociedade democrática. Sua intenção é a construção de conclusões válidas, por meio da incorporação de processos democráticos, para chegar a uma conclusão sobre o programa. Essa abordagem é caracterizada por três princípios básicos: a inclusão, o diálogo e a deliberação. O primeiro princípio considera a inclusão de todos os interesses, valores e opiniões dos atores envolvidos no programa analisado. O segundo supõe o diálogo entre as partes interessadas e os avaliadores durante o processo de avaliação, cabendo ao avaliador a estrutura do diálogo. E o terceiro prevê a deliberação para

chegar de modo reflexivo às conclusões ponderadas, o que inclui a identificação das preferências e dos valores de todos os grupos interessados.

A abordagem qualitativa é também encontrada na proposta de avaliação formulada por Malcolm Parlett e David Hamilton (1982, p. 40), que a denominaram de avaliação iluminativa. Sua principal preocupação é com a descrição e a interpretação, e não com a mensuração e a previsão. Seus objetivos são: estudar o programa inovador; descobrir e documentar em que consiste a participação na experiência inovadora; e discriminar e discutir as características mais importantes da inovação. O processo de avaliação se faz por meio de três etapas, que se superpõem e são inter-relacionadas, a saber: observação, questionamento e explicação. “Neste processo de três etapas, é traçado um perfil de informações, utilizando os dados coletados a partir de quatro fontes: observação, entrevistas, questionários e testes, documentos em geral” (PARLETT; HAMILTON, 1982, p. 42).

Outros usos da abordagem qualitativa de avaliação são apresentados por autores como Edward S. Halpern, que desenvolveu um modelo extensivo para investigações naturalistas de auditoria (WORTHEN; SANDERS; FITZPATRICK, 2004); H. S. Bholá, com sua vasta experiência na aplicação da abordagem construtivista para a avaliação de programas na África; Thomas Schwandt e seus trabalhos sobre os fundamentos filosóficos da avaliação construtivista; e David Fetterman, com sua abordagem de “avaliação de empoderamento” (tradução nossa)²⁵, que está estritamente atrelada à avaliação construtivista (STUFFLEBEAM, 2001).

Fernandes (2010, p. 21) observa que o processo de desenvolvimento das abordagens formais de avaliação “tem permitido, entre outros aspectos, identificar os seus propósitos, as suas epistemologias, os seus métodos preferenciais e os papéis dos avaliadores e dos diferentes intervenientes”. Este é um fator primordial, segundo o autor, sobretudo quando se pensa na construção de teoria em avaliação, de onde se requer compreensão, organização e análise das múltiplas abordagens e das diferentes definições e concepções da avaliação. Isso ajuda, inclusive, a entendermos o próprio funcionamento do campo, tendo em vista que os grupos vão se organizando e se solidificando em torno de determinadas abordagens, de modo que cada abordagem ganha força quando “tem importantes seguidores, apresenta explicitamente alguns fundamentos racionais, aparece nas bibliografias, [e] é posto em prática” (DIAS SOBRINHO, 2002, p. 17).

Considerações finais

Este artigo analisou a constituição teórica do campo da avaliação no cenário internacional a partir da contribuição de estudiosos que sistematizaram diferentes abordagens existentes no campo da avaliação educacional, com destaque para as abordagens de avaliação centradas nos objetivos, na lógica científica, no valor agregado, no serviço para a decisão, nos consumidores, nos participantes e no enfoque qualitativo.

A avaliação baseada em objetivos busca determinar se os objetivos educacionais foram atingidos. Os principais autores dessa abordagem são Tyler, Metfessel, Michael, Bloom, Hammond, Popham e Steinmetz. Suas vantagens são a fácil utilização, a aceitabilidade e o foco nos resultados. As principais críticas a essa abordagem

25 No original: “empowerment evaluation”.

se concentram no seu foco exclusivo nos resultados, no reducionismo que gera e na sua linearidade.

A avaliação baseada na lógica científica se equivale à pesquisa científica e, enquanto tal, segue a mesma lógica do método científico. Os principais autores dessa abordagem são Campbell, Stanley, Suchman, Cronbach, Lindquist, Snow, Boruch, Glass, Maguire, Wiley, Bock, Klineberg, James, Herzog e Fleck. As vantagens dessa abordagem são o foco nos resultados e a credibilidade. Suas desvantagens se direcionam às aplicações da experimentação em assuntos humanos.

A avaliação baseada no valor agregado procura determinar o quanto a escola contribui para o crescimento de seus alunos. Os principais autores dessa abordagem são Sanders, Horn, Webster, Mendro, Almaguer e Tymms. As vantagens dessa abordagem são a criação de banco de dados, o uso eficiente de testes padronizados e o progresso contínuo. Suas desvantagens são a forte dependência do quantitativo, o não detalhamento das entradas e dos processos do programa e a pouca utilização de métodos qualitativos.

A avaliação baseada no serviço para a decisão fornece informações para que os programas melhorem seus serviços. Os principais autores dessa abordagem são Stufflebeam, Alkin, Cronbach, Patton, Reichardt e Wholey. Suas vantagens são o atendimento de necessidades específicas, a ajuda na decisão, a apresentação lógica das informações, o equilíbrio no uso de métodos quantitativos e qualitativos, e seus princípios democráticos. As desvantagens são a perda de independência dos avaliadores, a ênfase excessiva na avaliação formativa e a pouca dedicação à avaliação somativa.

A avaliação baseada nos consumidores é vista como um processo de determinação de mérito e valor. Os principais autores dessa abordagem são Scriven, Glass, Ken Komoski, Morrisett e Stevens. Suas vantagens são a ênfase numa avaliação independente e objetiva, com redução do fator subjetivo, e a alta credibilidade com grupos de consumidores. Suas desvantagens estão relacionadas à sua independência, pois pode intimidar a inovação e sufocar a criatividade.

A avaliação baseada nos participantes visa descobrir, investigar e abordar os problemas de um programa. Os principais autores dessa abordagem são Stake, Denny, MacDonald, Parlett, Hamilton, Rippey, Smith e Pohland. As vantagens dessa abordagem são o uso de métodos qualitativos, a triangulação de diferentes fontes, a participação dos interessados na avaliação e a comunicação eficaz dos resultados. As desvantagens são a falta de credibilidade externa e de independência dos avaliadores e a pouca clareza dos resultados.

A avaliação qualitativa se baseia nas perspectivas valorativas dos interessados. Os principais autores dessa abordagem são Guba, Lincoln, MacDonald, House, Parlett, Hamilton, Halpern, Bhola, Schwandt e Fetterman. As vantagens dessa abordagem são os processos de negociação, as orientações democráticas, a descrição detalhada e a liberdade de escolha dos métodos. Suas desvantagens estão na dificuldade da produção de relatórios devido ao envolvimento e à interação contínua entre avaliador e avaliados.

As abordagens de avaliação descritas no quadro da revisão teórica realizada neste estudo, além de fornecerem uma compreensão do desenvolvimento histórico e conceitual do campo da avaliação no cenário internacional, também oferecem, aos interessados pela avaliação, a possibilidade de escolha daquela abordagem que

julgarem mais apropriada para responder à natureza de determinado problema educacional. Nesse caso, o avaliador precisa, necessariamente, dispor de liberdade para elaborar o planejamento da avaliação, escolher os instrumentos a serem aplicados e participar efetivamente na execução do plano de avaliação – sem contar a necessidade de se observar as diferentes realidades a serem avaliadas e de incentivar o envolvimento dos demais membros no processo avaliativo. Caso contrário, corre-se o risco de que o avaliador pareça o único responsável pela avaliação.

Entendemos que, de alguma maneira, as abordagens de avaliação apresentadas influenciaram e continuam presentes em nossas práticas e estudos de avaliação. No passado, em nosso cenário educacional, essas abordagens marcaram fortemente a produção científica brasileira sobre o tema que, inicialmente focada na ideia tyleriana de avaliação como atividade final para alcançar objetivos, passou, com o tempo, a incorporar outras alternativas metodológicas de cunho mais qualitativo para a avaliação, com destaque para as concepções de Scriven, Stake, Stufflebeam, Parlett e Hamilton, incluindo a dimensão política de MacDonald e democrática de House. Ainda hoje os conceitos de avaliação somativa e formativa de Scriven têm orientado nossos estudos, assim como a proposição de Guba e Lincoln em torno da ideia da avaliação como negociação e a presença do estudo de caso como técnica de pesquisa na avaliação, conforme proposto por Stake, Guba e Lincoln e MacDonald.

De modo geral, compreendemos que essa literatura internacional, traduzida nas abordagens de avaliação expostas, exerceu forte influência sobre os pesquisadores brasileiros da avaliação que, fundamentados nesses referenciais, buscaram sustentar uma área de conhecimento que ainda se encontra em processo de constituição e fortalecimento em nosso país. A adoção dessas fontes marcou um novo momento para os estudiosos brasileiros da avaliação que, munidos por essa literatura, passaram a desenvolver referenciais teóricos alternativos ao modelo tecnicista de avaliação predominante até então. Com isso criaram espaço para o surgimento, em nosso meio, da chamada avaliação emancipatória.

Referências

- ALKIN, Marvin C.; CHRISTIE, Christina A. An evaluation theory tree. In: ALKIN, Marvin C. (Ed.). **Evaluation roots: tracing theorists' views and influences**. London: Sage, 2004. p. 381-392.
- CAMPBELL, Donald T.; STANLEY, Julian C. Experimental and quasi-experimental designs for research on teaching. In: GAGE, N. L. (Ed.). **Handbook of research on training**. Chicago: Rand McNally, 1963. p. 1-84.
- DE KETELE, Jean-Marie. L'évaluation conjugue en paradigmes. **Revue Française de Pédagogie**, Lyon, n. 103, p. 59-80, 1993.
- DIAS SOBRINHO, José. Campo e caminhos da avaliação: a avaliação da educação superior no Brasil. In: FREITAS, Luiz Carlos (Org.). **Avaliação: construindo o campo e a crítica**. Florianópolis: Insular, 2002. p. 13-62.
- FERNANDES, Domingos. Acerca da articulação de perspectivas e da construção teórica em avaliação educacional. In: ESTEBAN, Maria Teresa; AFONSO, Almerindo Janela (Orgs.). **Olhares e interfaces: reflexões críticas sobre a avaliação**. São Paulo: Cortez, 2010. p. 15-44.
- GUBA, Egon G.; LINCOLN, Yvonna S. **Avaliação de quarta geração**. Trad. Beth Honorato. Campinas: Editora da Unicamp, 2011.
- HOUSE, Ernest; HOWE, Kenneth R. Deliberative democratic evaluation in practice. In: STUFFLEBEAM, Daniel L.; MADAUS, George F.; KELLAGHAN, Thomas (Eds.). **Evaluation models: viewpoints on educational and human services evaluation**. Boston: Kluwer Academic Publishers, 2000. p. 409-422.
- MACDONALD, Barry. Uma classificação política dos estudos avaliativos. In: GOLDBERG, Maria Amélia Azevedo; SOUSA, Clarilza Prado de (Orgs.). **Avaliação de programas educacionais: vicissitudes, controvérsias, desafios**. São Paulo: EPU, 1982. p. 16-17.

OWSTON, Ron. Models and Methods for Evaluation. In: SPECTOR, J. Michael et al. (Eds.). **Handbook of Research on Educational Communications and Technology**. New York: Lawrence Erlbaum Associates, 2008. p. 605-618.

PARLETT, Malcolm; HAMILTON, David. Avaliação Iluminativa: uma nova abordagem no estudo de programas inovadores. In: GOLDBERG, Maria Amélia Azevedo; SOUSA, Clarilza Prado de (Orgs.). **Avaliação de programas educacionais: vicissitudes, controvérsias, desafios**. São Paulo: EPU, 1982. p. 38-45.

SANDERS, William L.; HORN, Sandra P. The Tennessee value-added assessment system (TVAAS): mixed model methodology in educational assessment. **Journal of Personnel Evaluation in Education**, New York, v. 8, n. 3, p. 299-311, 1994.

SANDERS, William L.; RIVERS, June C. **Cumulative and residual effects of teachers on future student academic achievement**. Knoxville: University of Tennessee Value-Added Research and Assessment Center, 1996.

STAKE, Robert E. **Evaluation the arts in education: a responsive approach**. Columbus: Merrill, 1975a.

STAKE, Robert E. **Program evaluation, particularly responsive evaluation**. Kalamazoo: Western Michigan University Evaluation Center, 1975b. (Occasional paper n. 5).

STUFFLEBEAM, Daniel L. Evaluation models. **New directions for evaluation**, San Francisco, n. 89, p. 7-98, 2001.

STUFFLEBEAM, Daniel L.; SHINKFIELD, Anthony L. **Evaluación sistemática: guía teórica y práctica**. Trad. Carlos Losilla. Barcelona: Paidós, 1987.

WORTHEN, Blaine R.; SANDERS, James S.; FITZPATRICK, Jody L. **Avaliação de programas: concepções e práticas**. Trad. Dinah de Abreu Azevedo. São Paulo: Gente, 2004.

Enviado em: 18/outubro/2017

Aprovado em: 25/outubro/2018

Ahead of print em: 10/janeiro/2019