

AVALIAÇÃO EDUCACIONAL: CONCEPÇÕES E EMBATES TEÓRICOS

ASSIS LEÃO DA SILVA^I

ALFREDO MACEDO GOMES^{II}

RESUMO

O artigo objetiva realizar uma discussão teórica ao revisar, sistematizar e refletir a respeito de distintas concepções de avaliação e de suas tendências no campo da avaliação educacional. A construção histórica dessas concepções de avaliação sinaliza, em suas origens, determinadas demandas que alteram a ênfase da avaliação educacional em cada momento do campo da avaliação educacional. Destaca-se que, no contexto atual, a compreensão a respeito da avaliação educacional está associada ao hibridismo de distintas concepções de avaliação. Esse hibridismo, materializado nas políticas de avaliação, tem posto em evidência um momento de transição nas tendências das concepções de avaliação, com avanços e recuos importantes no alcance, objetos e lógica da avaliação educacional, ao combinar harmonização ou conflitos das/entre concepções de avaliação.

^I Instituto Federal de Pernambuco (IFPE), Recife-PE, Brasil; assisleao33@gmail.com

^{II} Universidade Federal de Pernambuco (UFPE), Recife-PE, Brasil; alfredomgomes@gmail.com

PALAVRAS-CHAVE CONCEPÇÕES DE AVALIAÇÃO • TENDÊNCIAS DE AVALIAÇÃO • POLÍTICAS DE AVALIAÇÃO • AVALIAÇÃO EDUCACIONAL.

EVALUACIÓN EDUCACIONAL: CONCEPCIONES Y EMBATES TEÓRICOS

RESUMEN

El artículo tiene el objetivo de realizar una discusión teórica al revisitar, sistematizar y reflexionar sobre distintas concepciones de evaluación y sus tendencias en el ámbito de la evaluación educacional. La construcción histórica de tales concepciones de evaluación señala, en sus orígenes, determinadas demandas que alteran el énfasis de la evaluación educacional en cada momento del campo de la evaluación educacional. Se subraya que, en el actual contexto, la comprensión acerca de la evaluación educacional está asociada al hibridismo de distintas concepciones de evaluación. Dicho hibridismo, materializado en las políticas de evaluación, ha puesto en evidencia un momento de transición en las tendencias de las concepciones de evaluación, con importantes avances y retrocesos en lo que se refiere al alcance, objetos y lógica de la evaluación educacional, al combinar armonización o conflictos de las/jentre concepciones de evaluación.

PALABRAS CLAVE CONCEPCIONES DE EVALUACIÓN • TENDENCIAS DE EVALUACIÓN • POLÍTICAS DE EVALUACIÓN • EVALUACIÓN EDUCACIONAL.

EDUCATIONAL EVALUATION: THEORETICAL CONCEPTS AND CONFLICTS

ABSTRACT

This paper aims to conduct a theoretical discussion by revisiting, systematizing, and reflecting upon different concepts and trends in the field of educational evaluation. The historical construction of these concepts of evaluation indicates specific social and educational demands, that change the focus of educational evaluation at different times, in the field of educational evaluation. It should be stressed that, in the current context, the understanding of educational evaluation is associated with the hybridization of different concepts of evaluation. This hybridism, seen in the evaluation policies, has evidenced a moment of transition in the trends of concepts of evaluation, with important advances and retreats in the reach, objects and logic of educational evaluation. It does so by resolving agreement or conflict between the concepts of evaluation.

KEYWORDS EVALUATION CONCEPTS • EVALUATION TRENDS • EVALUATION POLICIES • EDUCATIONAL EVALUATION.

INTRODUÇÃO

A avaliação tornou-se campo do conhecimento e prática social funcional ao modo de gestão estatal nas sociedades capitalistas avançadas nos últimos 60 anos. Como tal, a avaliação caracteriza-se pela vitalidade, dispersão e aparente desordem, aspectos que trataremos ao longo do texto. Tais características colocam grandes desafios à sua compreensão e tendem a dissimular suas bases epistemológicas. Problematizar e compreender o domínio exercido pela avaliação em seu papel de avaliar, regular, justificar ou desacreditar programas e políticas públicas tornam-se tarefas relevantes e inadiáveis, especialmente no campo da educação.

O patrocínio estatal massivo criou, e continua a criar, problemas nunca antes vistos no campo da avaliação educacional, o que inclui, dialeticamente, o ritmo, o conteúdo e as formas de mudança desse campo de conhecimento. Entendemos que se trata de mudanças estruturais na concepção e posicionamento da avaliação, expressando seu predicado polissêmico e os múltiplos interesses em torno de suas práticas. Compreendemos, em contrapartida, que, no contexto de

sua construção histórica, há lutas – em algumas formações sociais mais visíveis que em outras – pela democratização das decisões públicas, o que impacta diretamente as políticas de avaliação da educação em todos os níveis. No entanto, as políticas de avaliação padecem de inúmeras dificuldades para cumprir essa aspiração social, ou seja, sua abertura ao escrutínio e à deliberação pública.

A luta pela democratização das decisões públicas desvela a abordagem política da política de avaliação e, concomitantemente, o posicionamento mais central, crescente e hegemônico da concepção de avaliação “dirigida pela negociação” na agenda das políticas públicas de avaliação. Essa disputa está cimentada no complexo jogo político que as sociedades contemporâneas e suas instituições de educação atravessam, e está fundamentada em visões distintas de sociedade e de gestão pública da educação que se ampara e se propaga a partir do paradigma gerencialista (GUBA; LINCOLN, 2011) por um lado, e, por outro, do paradigma democrático (MACDONALD, 1995), impactando na definição dos objetivos, finalidades, instrumentos, modelos e usos da avaliação no sistema de educação, o que inclui, obviamente, o caso brasileiro.

O presente artigo, escrito na forma de ensaio teórico, discute as concepções de avaliação historicamente construídas, de modo a caracterizar e problematizar suas contribuições e limitações teóricas e defender a tese de determinado “desenvolvimento” teórico-metodológico da avaliação, o que implica reconhecer o desenvolvimento e a consolidação da teorização sobre a avaliação. Para tanto, é um artigo que se baseia em literatura selecionada.

CONSTRUÇÃO HISTÓRICA DAS CONCEPÇÕES DE AVALIAÇÃO EDUCACIONAL

Arredondo e Diago (2009, p. 29, grifos nossos) afirmam que:

[...] a avaliação foi constituída em uma disciplina científica que serve como elemento de motivação e de ordenação intrínseca da aprendizagem. A esse respeito, House (1993) considera que “a avaliação passou de uma atividade

marginal, desenvolvida em tempo parcial por acadêmicos, a uma **pequena indústria profissionalizada**". Essa mudança na forma de conceber e aplicar a avaliação representou importantes transformações, tanto em sua concepção quanto em sua prática, embora processos de mudança possam ser numerosos e abarcar diversos âmbitos do sistema educacional. Todos esses fatores estão nos levando a uma **"cultura da avaliação"**, que não se limita [ao campo educacional], mas que se estendem às demais atividades sociais, o que levou a maioria dos países, cientes dessa realidade, a fornecer recursos econômicos, materiais e humanos, dadas as expectativas que esse fenômeno gerou.

Interpretar essa **"pequena indústria profissionalizada"** ou **"cultura de avaliação"** no cenário descrito pelos autores se faz necessário para compreender determinadas concepções de avaliação no campo educacional. Diversos autores buscaram periodizar a trajetória histórica e o desenvolvimento das concepções da avaliação, caracterizando-os ao longo do século XX. Dentre eles encontram-se Stufflebeam e Shinkfield (1987), Arredondo e Diago (2009) e Guba e Lincoln (2011). Evidentemente que qualquer fixação de datas e caracterização de acontecimentos históricos no contexto da dinâmica acadêmica e social global e nacional é tarefa complexa. No entanto, o esforço de interpretação empreendido aqui busca estabelecer uma aproximação com a produção teórica do campo da avaliação. O Quadro 1 apresenta a síntese do desenvolvimento do campo da avaliação educacional, tendo como foco concepções de avaliação que entendemos merecedoras de caracterização e escrutínio crítico no contexto do presente ensaio, a saber: avaliação concebida como mensuração, avaliação orientada por objetivos, avaliação concebida como juízo de valor e avaliação concebida como negociação.

QUADRO 1 - Periodização do campo disciplinar da avaliação

PERÍODOS	DEMANDAS AO CAMPO DA AVALIAÇÃO EDUCACIONAL	CONCEPÇÃO DE AVALIAÇÃO	TERMOS-CHAVE
1890-1930	Testes mentais para o sistema educacional e as forças armadas norte-americanas	Mensuração	Medição Objetividade
1930-1957	Reformas curriculares e a massificação do sistema educacional nos países desenvolvidos	Objetivos	Objetivos educacionais Descrição Grau de consecução dos objetivos Totalidade do sistema educacional
1957-1973	Críticas à efetividade do sistema educacional norte-americano; governo John Kennedy e as políticas de avaliação	Juízo de valor	Juízo de valor <i>Accountability</i>
1973-dias atuais	Políticas de avaliação e os regimes de regulação burocrático-profissional e pós-burocráticos	Negociação	Avaliação de mudança ocorrida no aluno Negociação Quantitativa/Qualitativa Formativa/Diferenciada/ Integradora

Fonte: Elaboração dos autores a partir de dados de Stufflebeam e Shinkfield (1987), Maroy e Dupriez (2000), Arredondo e Diago (2009) e Guba e Lincoln (2011).

Da síntese, pode-se observar um panorama das concepções, demandas e limites e a inserção da avaliação no campo educacional em dado contexto histórico ao longo do século XX. Os termos-chave, na última coluna do Quadro 1, apresentam a síntese da visão e do esforço dos autores citados para desenhar tais concepções de avaliação educacional.

De modo geral, é possível identificar que as demandas ou imposições de ordem social, econômica e política postas ao campo educacional impactam as concepções e usos das modalidades de avaliação nas políticas educacionais, ampliando ou reconfigurando seus objetivos, finalidades e domínios conceituais. Passamos a explorar, a seguir, de forma mais detalhada, tais concepções.

A MENSURAÇÃO COMO CONCEPÇÃO DE AVALIAÇÃO

Na primeira fase, que compreende o período aproximado entre 1890 e 1930, pode-se dizer que a avaliação foi construída essencialmente como mensuração, fundada numa visão positivista de ciência e de mundo, como instrumentação eminentemente técnica, consistindo de testes de verificação, mensuração e quantificação da aprendizagem de objetos referidos como tais ou aproximados. Stufflebeam e Shinkfield (1987)

denominaram essa primeira fase de “pré-Tyler”, o que prenunciava a relevância de Ralph Tyler para a fase subsequente.

Do ponto de vista disciplinar, a concepção de avaliação como mensuração/medição é produto fundamental do campo da psicologia. Segundo Arredondo e Diago (2009), a palavra avaliação sequer era mencionada nos livros e manuais à época. A principal preocupação dos cientistas centrava-se na mensuração de atributos de crianças e jovens em idade escolar e no recrutamento de soldados para as forças armadas, no caso dos Estados Unidos. A utilização de exames ou testes mentais dominava o conteúdo de diversos cursos, com ênfase nos testes de memória por meio de exames orais e/ou dissertativos. O principal objetivo era ensinar o que se reconhecia como “certo”, medindo aptidões ou aprendizagens, quantificando-as, comparando-as e ordenando-as em escala.

No fim do século XIX, desenvolveram-se as medições psicofísicas amparadas em outras áreas do conhecimento. O primeiro trabalho publicado sobre esses testes em escolas foi o de Joseph Mayer Rice – *The futility of the spelling grind* –, em 1897, resultado dos testes comparativos aplicados em escolas dos Estados Unidos com pretensões de fazer uma avaliação objetiva da educação escolar. A preocupação de Rice era o ensino de conceitos básicos, com a proposta de tornar as escolas mais competentes ou científicas.

Na França, o psicólogo Alfred Binet elaborou testes para realizar triagem de jovens então diagnosticados com retardamento mental. Utilizou, para isso, técnicas psicométricas aperfeiçoadas na Inglaterra por Francis Galton e Karl Pearson, as quais foram revistas ou redefinidas por Wilhelm Wundt, na Alemanha. Binet percebeu que as técnicas elaboradas em ambos os países eram ineficientes e elaborou uma nova abordagem, formulando o conceito de “idade mental” no início do século XX, com objetivo de determinar o “quociente de inteligência” (QI) (STERNBERG, 2000).

Apropriando-se dessa problematização, Edward Thorndike (1931¹ apud LEFRANÇOIS, 2015, p. 76) desenvolveu instrumentos escolares de medida, uso de testes e métodos estatísticos na psicologia e na educação para defender que “a aprendizagem consiste na formação de vínculos entre

¹ THORNDIKE, E. *Educational psychology: the psychology of learning*. New York: Teachers College, 1931.

estímulos e repostas”. Esse autor publicou, em 1903, a obra *Education psychology* e, em 1904, *Introduction to the theory of mental and social measurement*, exercendo notável liderança intelectual nesse campo, criando a psicologia educacional. Foi o precursor do movimento da “gestão científica” no setor educacional.

No que concerne à gestão científica no campo educacional, Hypolito (2007, p. 3) define como modelo de gestão

[...] baseado na lógica racional-técnico-instrumental, orientada pelos pressupostos da administração científica e que acompanhou os modelos de gerência do taylorismo, fayolismo e fordismo.

A gestão científica propôs/propõe uma visão burocrática e tecnicista de escola, com ênfase numa escola supostamente neutra e objetiva, para o controle racional, com foco na estrutura organizacional.

Em 1916, Lewis Terman reviu o teste de Binet-Simon, denominando-o posteriormente de *Stanford-Binet Intelligence Scale*, que se tornou um dos mais tradicionais testes de inteligência, amplamente utilizado no sistema educacional e nas forças armadas norte-americanas. No caso das forças armadas, os comandantes contaram com apoio da American Psychological Association. Terman contou com a expressiva participação da Universidade de Columbia na classificação e seleção de alunos para ingresso na educação superior.

Dessa forma, a mensuração/medição inseriu-se, a partir de uma perspectiva da psicologia, como elemento-chave da avaliação, isso porque a aprendizagem era, à época, a problemática central das pesquisas no campo da psicologia, pois se acreditava em poder mensurá-la e quantificá-la por meio de método científico. A inserção do método científico no campo da psicologia, concomitantemente ao desenvolvimento do campo da avaliação, foi resultado da expressiva ascensão das ciências sociais nos Estados Unidos, Inglaterra, Alemanha e França, tanto nos sistemas educacionais como nos de saúde (STUFFLEBEAM; SHINKFIELD, 1987). Essa ascensão estava diretamente relacionada à quantificação como fundamentação científica e elemento de credibilidade dos estudos sociais e humanos.

Segundo Guba e Lincoln (2011), na década de 1920, floresceu, no campo da administração, a gestão científica, sobretudo para as indústrias e empresas. Esse movimento baseou-se intensamente em pesquisas quanto ao tempo e ao movimento dos trabalhadores para determinar os métodos mais bem-sucedidos de produção, situar as taxas salariais por empreitada e “incentivar” a classe trabalhadora. Esse momento foi marcado pelo trabalho de Taylor (1990). No campo educacional, a introdução da gestão científica possibilitou um novo enfoque de gestão educacional, no qual as escolas passavam a ser concebidas como equivalentes às organizações empresariais. Nesse modelo, os alunos eram vistos como matérias-primas a serem processadas na fábrica-escola. E, nesse cenário, os testes desempenharam função decisiva na especificação do nível dos alunos conforme as expectativas esperadas, sobretudo os referentes à preparação para o ensino superior.

Em síntese, a função do avaliador era técnica, baseada no arsenal de instrumentos disponíveis para mensurar qualquer variável. As funções da avaliação eram classificar, selecionar e certificar o conhecimento, tomado como o principal objetivo da avaliação. A avaliação era descontextualizada e referida como uma norma ou padrão. Essa concepção de avaliação como mensuração e suas características ainda tem destacado peso nos sistemas educacionais atuais, materializados nas práticas de avaliação de testes e exames standardizados como instrumentos das políticas de avaliação para “mediar as relações mais amplas da cultura com o Estado, constituindo-se num vínculo objetivo entre o saber da sociedade civil e o saber do Estado” e

[...] tornar-se, assim, fundamentalmente uma técnica de certificação para medir com pretensa objetividade (e atestar juridicamente) um nível determinado de qualificação. (AFONSO, 2005, p. 30)

CONCEPÇÃO DE AVALIAÇÃO ORIENTADA POR OBJETIVOS

Nos anos de 1930, o desenvolvimento das práticas avaliativas na perspectiva da mensuração demonstrava considerável esgotamento e limitações de ordem teórica,

metodológica e prática. Tornaram-se evidentes as limitações da mensuração dos rendimentos dos alunos face à crescente preocupação com as dimensões mais amplas do processo de ensino-aprendizagem, entre as quais o currículo e suas coerências entre a avaliação e os objetivos preestabelecidos, regulando o conhecimento e as maneiras de adquiri-lo. A abordagem da avaliação-mensuração não oferecia nada além de dados sobre os estudantes e não servia aos propósitos das novas demandas sociais postas à avaliação, entre os quais conceber a avaliação como processo sistemático destinado a produzir mudanças na conduta dos alunos por meio da instrução, do desenvolvimento do currículo e do grau de consecução dos objetivos propostos, distinguindo-se do momento anterior centrado no estabelecimento das diferenças individuais para fins de seleção, pouco havendo com os objetivos dos programas desenvolvidos nas escolas (GUBA; LINCOLN, 2011).

A ideia de que qualquer pessoa poderia aprender começou a ser posta à prova, havendo uma mudança gradativa nos testes de rendimento, que culminou com o início da segunda fase, amparada na concepção de educação como realização de “objetivos”, caracterizada pela assunção da ideia da avaliação como descrição, uma vez que se passou a estudar o desempenho dos alunos relacionado ao comportamento de seu grupo, quando os objetivos educacionais instituíram-se como novo foco da avaliação (STUFFLEBEAM; SHINKFIELD, 1987). Na literatura acadêmica (ARREDONDO; DIAGO, 2009; GUBA; LINCOLN, 2011; STUFFLEBEAM; SHINKFIELD, 1987), reconhece-se que o grande expoente da avaliação nesse período foi Ralph Tyler, na verdade o primeiro a cunhar a expressão “avaliação educacional” (ARREDONDO; DIAGO, 2009). Com a publicação de *Basic principles of curriculum and instruction*, em 1949, Tyler consolidou uma perspectiva nos campos educacional e da avaliação, sendo considerado, anos mais tarde, o “pai” da avaliação educacional.

Baseando-se em Arredondo e Diago (2009), podemos dizer que a fase que se caracteriza pela concepção de avaliação orientada por objetivos foi marcada pelo incentivo à gestão científica da escola e ao desenvolvimento curricular

e institucional por meio de objetivos previamente estabelecidos. Tyler definiu os objetivos educacionais em termos de conduta dos estudantes. Assim, para ele, deveria haver coerência entre a avaliação e os objetivos educacionais, resultado de cuidadoso planejamento curricular. Utilizando-se do método experimental, procurou verificar se os alunos eram capazes de alcançar os objetivos previstos no processo de aprendizagem. Para Tyler, ensinar significava mudar padrões de comportamento e, como consequência, o currículo passou a ser concebido com base na especialização das habilidades esperadas expressas em objetivos a serem atingidos.

Na problematização entre currículo tradicional e currículo progressista, Tyler coordenou amplo estudo longitudinal, denominado *The eight-year study* (1932-1940), para estudar a eficácia da escola tradicional em relação à escola progressista, pois as escolas tradicionais e as universidades recusavam-se a democratizar o acesso às classes populares. Nesse contexto, a contribuição de Tyler para o estudo vai apresentar uma nova concepção de prática avaliativa na qual a avaliação deve focar o aluno em suas habilidades, conhecimentos e atitudes, rompendo com a prática estrita de mensuração (WATRAS, 2006).

A proposta de Tyler enfatizava a relação professor-aluno quando esse afirmava a necessidade de que os docentes expusessem os objetivos de aprendizagem aos alunos por meio de diversas práticas curriculares. Nessa perspectiva, evidenciam-se os aspectos da solidariedade, pois o ensino e a aprendizagem não seriam atos isolados, mas uma atividade que envolve um esforço cooperativo, com o envolvimento de distintos agentes educacionais e sociais. Para esse autor, a educação é um processo que deve almejar a criação de padrões de comportamento relativos aos objetivos educacionais. Nesse contexto, o êxito de um programa educacional só poderia ser verificado por meio da avaliação, quando se concretizam os objetivos propostos. Na verdade, o modelo proposto por Tyler enfoca a avaliação como ferramenta para julgar o currículo. Nesse modelo, não era possível associar avaliação exclusivamente à mensuração. No entanto, Tyler não negava a relevância dos testes e das provas (GUBA; LINCOLN, 2011).

Segundo Stufflebeam e Shinkfield (1987), as reações à proposta de Tyler não tardaram a aparecer, sobretudo ao uso da psicometria em avaliação, fundamentada em distintas posições epistemológicas referentes à pesquisa e à avaliação. Mesmo assim, a importância desse teórico consolidou-se no campo da avaliação educacional norte-americano, com o desenvolvimento do National Assessment of Educational Progress (Naep)² (TYLER, 1966), nos anos de 1960. Os pressupostos teórico-metodológicos formulados por Tyler parecem dominar a cultura de avaliação nos Estados Unidos até os dias atuais, mesmo diante do surgimento de modelos alternativos à sua proposta. Segundo Vianna (2000), a influência de Tyler não se restringiu à área do currículo, pois compreendeu a avaliação do rendimento escolar e as taxonomias dos objetivos educacionais a partir dos anos 1970.

² Ver o site <<http://nces.ed.gov/nationsreportcard/>> para uma visão detalhada do Naep na atualidade.

O modelo de Tyler baseava-se na metodologia de análise quantitativa, com orientação geral do individualismo metodológico, na linha dos delineamentos experimentais (ARREDONDO; DIAGO, 2009). Essa proposta passou a ser questionada pelos teóricos sociais a partir dos anos 1960. A questão enquadrava-se no âmbito do conflito entre os objetivos das ciências sociais e das ciências físicas, com metodologias diferentes. Problematicava-se a possibilidade de estudar os fenômenos naturais e sociais a partir de uma unidade metodológica. Segundo Guba e Lincoln (2011), diversos teóricos do campo da avaliação, a partir dos anos 1960, passaram a criticar o modelo de Tyler na avaliação educacional, entre eles, Cronbach (1963), Scriven (1967) e Stake (1967). Cronbach (1963) critica a restrição da avaliação aos aspectos ritualísticos da mensuração. Essa deveria oferecer meios que possibilitassem o aprimoramento dos currículos, compreendendo o caráter multidimensional dos resultados da aprendizagem, exigindo da avaliação uma coleta diversificada de informações, a fim de descrever os currículos e programas educacionais. Scriven (1967) introduz a ideia de obrigar o avaliador a determinar o valor (mérito) do objetivo, diferenciando os papéis dos objetivos, para poder determinar a eficácia das práticas educacionais. Em sua obra mais relevante, *The countenance of educational evaluation*, Stake (1967) defende a tese de que a avaliação deveria ser exaustiva

e voltada à compreensão do todo, rompendo, assim, com seu caráter episódico e pontual.

Naquele contexto, a avaliação adquiriu uma nova característica, de tal forma que cabia aos objetivos originar os critérios e normas e à avaliação determinar as similaridades e diferenças entre o prometido e o cumprido. O papel essencial da avaliação seria determinar o grau de mudanças comportamentais por meio dos objetivos educacionais traçados em relação ao currículo e ao desenvolvimento das práticas pedagógicas.

Segundo Fernandes (2009), o corpo teórico e o prático da avaliação assumem outras características nessa fase. No campo pedagógico, a avaliação apropriou-se da ideia da eficiência. Esse aspecto já vinha adquirindo visibilidade desde os anos 1920, quando o currículo foi vinculado à ideia de eficiência social. Dessa maneira, a avaliação tornou-se um mecanismo para diagnosticar o rendimento e a eficácia da escola, dos processos pedagógicos e administrativos, envolvendo a taxa de sucesso escolar, o tempo na consecução das atividades e a taxa de aprovação nos níveis escolares, dentre outros.

A avaliação permanecia essencialmente técnica, porém não mais restrita ao âmbito do conceito de medida; passa a ser referida como um processo de aperfeiçoamento contínuo para a reformulação do currículo. Ensaia seus primeiros voos no sentido da *accountability*, para demonstrar a eficiência e produtividade dos sistemas e unidades educacionais, sobretudo no momento posterior a essa fase, quando, em meados dos anos 1960, teóricos como Cronbach e Bloom começaram a contestar a ideia de que o fracasso escolar era produto exclusivo da ação dos alunos individualmente, sem a participação do sistema educacional, da escola e dos docentes (VIANNA, 2000). Segundo Afonso (2005), no campo da educação, o emprego do conceito de *accountability* refere-se à prestação de contas e à responsabilização docente, entendido como necessário às políticas de avaliação nos sistemas educacionais.

Em síntese, a fase da avaliação orientada por objetivos caracterizou-se pela “descrição” do padrão de pontos fortes e fracos referentes aos objetivos estabelecidos, denominada

“geração dos objetivos” (STAKE, 1967). O avaliador cumpria a função de descritor, mantendo os aspectos técnicos anteriores a essa função. A mensuração não desaparece, mas se entende que poderia ser mais um dos instrumentos empregados no processo da avaliação.

A CONCEPÇÃO DE AVALIAÇÃO BASEADA EM JUÍZO DE VALOR

A terceira fase refere-se à concepção de avaliação como “juízo de valor”. Desenvolveu-se entre os fins dos anos 1950 e o início dos anos 1960. De maneira geral, nessa fase buscou-se empreender esforços para que as avaliações permitissem formular juízos de valor quanto aos seus respectivos objetos.

Denominada por Stufflebeam e Shinkfield (1987) de a “era do realismo”, e por Madaus e Stufflebeam (2000) de a “era do desenvolvimento”, essa fase passou a lograr a assunção do revigoramento das práticas avaliativas, tanto no campo da avaliação quanto no cenário mais amplo da educação nos Estados Unidos. A avaliação ampliou seus horizontes e adquiriu sofisticação do ponto de vista teórico. No cenário mais amplo, 1957 foi importante para a corrida espacial, pois se iniciava a era pós-Sputnik. O avanço soviético na corrida espacial suscitou sucessivas e massivas críticas ao sistema educacional norte-americano que impulsionaram o revigoramento das práticas avaliativas ancoradas em novas abordagens.

Nesse cenário, a abordagem das práticas avaliativas da segunda fase da avaliação – orientada por objetivos – mostrou-se inapropriada, pois parte dos avaliadores passou a questionar e rever a concepção de avaliação orientada por objetivos. Não queriam comprometer-se enquanto não tivessem uma visão mais nítida acerca do que de fato avaliavam, nem mesmo estabelecer objetivos provisórios para evitar a perda de sua criatividade. Na verdade, passaram a não adotar uma estratégia de avaliação enquanto o programa que avaliavam não tivesse tido desenvolvimento pleno (GUBA; LINCOLN, 2011).

Essa mudança no campo da avaliação e nas práticas avaliativas adquiriu notável visibilidade no governo do

presidente norte-americano John Kennedy e nos governos subsequentes, criando condições para a ação estatal no setor educacional por meio das políticas de avaliação. Logo, a avaliação passou a ser um conteúdo concreto e simbólico de decisões políticas e do artifício de construção e atuação das decisões estatais e “multicêntricas” (SECCHI, 2010), ou seja, tanto do monopólio de atores estatais quanto de organizações privadas, organizações não governamentais e organismos multilaterais, disseminando essas abordagens a partir dos Estados Unidos para o restante do mundo.

No contexto norte-americano, a avaliação passou a ser parte obrigatória da educação e posteriormente dos programas federais, iniciando políticas e iniciativas que seriam consolidadas na fase seguinte, caracterizada pela profissionalização do campo e consolidação de determinadas teorias da avaliação (HOUSE, 2000). Esses programas estão bem documentados no clássico artigo *Course improvement through evaluation*, de Cronbach, publicado em 1963.

Segundo House (2000), as escolas passaram a ser responsabilizadas pelos baixos rendimentos dos alunos, ao tempo em que apareceram diversas críticas em torno da eficácia e eficiência do financiamento público no campo educacional, exigindo-se transparência e prestação de contas dos educadores aos usuários do sistema educacional. Com vistas a superar as críticas ao sistema educacional, o governo norte-americano passou a empregar, nas escolas, o modelo produzido pelo Pentágono,³ a fim de elaborar “programas que possam ser enunciados, medidos e avaliados na perspectiva da relação custo/benefício” (House, 2000, p. 185). Essa medida, ao almejar vincular os requisitos operacionais às obrigações financeiras do sistema educacional, procurou desvelar a ineficácia das metodologias das práticas avaliativas amparada na abordagem da avaliação por objetivos.

Ao diagnosticar as novas e sofisticadas demandas postas às políticas educacionais e à ação estatal no campo educacional, tais como a organização dos estudantes, a regulação administrativa e a melhoria dos cursos, o campo da avaliação educacional passou a reconhecer não apenas os objetivos, mas, sobretudo, as funções dos diagnósticos das práticas

3 O modelo de avaliação do Pentágono era o *Planning, Programming and Budgeting System* (PPBS). Esse modelo busca integrar uma variedade de técnicas em um processo de planejamento e orçamento, para identificar, custear e atribuir uma complexidade de recursos, para estabelecer prioridades e estratégias em um programa importante e para prever custos, despesas e realizações no exercício financeiro imediato ou durante um período mais longo, vinculando os requisitos operacionais às obrigações financeiras.

avaliativas para as tomadas de decisão, relegando ao segundo plano o alcance dos objetivos, justificadas pela necessidade de introduzir, também, alterações no processo, e não apenas no produto educacional; o juízo de valor passou a distinguir o processo do produto. Como consequência, no âmbito escolar, as avaliações não apenas avaliariam os alunos, mas passariam a avaliar as escolas, os professores, os conteúdos, as metodologias e as estratégias de ensino e aprendizagem, superando a tese dos objetivos como elemento organizador da avaliação, com a introdução da noção das tomadas de decisão, a qual ampliaria os horizontes do campo da avaliação educacional.

Dessa maneira, conforme Guba e Lincoln (2011), a concepção de avaliação como juízo de valor rompe com o paradigma positivista – caracterizado pelo quantitativismo; orientado ao resultado; de realidade estável, dedutível, inferencial e hipotética; busca dos fatos ou causas dos fenômenos sociais –, passando-se a utilizar, gradativamente, também, os enfoques naturalistas – de observação naturalista e sem controle, orientada ao processo; descritivo e indutivo; métodos qualitativos. No caso, a avaliação passou a assumir e integrar contribuições da sociologia, da antropologia e da psicologia social, permitindo-lhe ultrapassar os limites da perspectiva *behaviorista* e lograr o desenvolvimento de conceitos importantes até hoje presentes no campo da avaliação educacional e, consequentemente, nas políticas educacionais de avaliação, tais como “avaliação somativa”, “avaliação formativa” e “avaliação global”, entre outros. Os dois primeiros conceitos foram desenvolvidos por Scriven (1967), e o último, por Stake (1967).

Scriven (1967) formulou argumentos fundamentais para a compreensão da lógica da avaliação educacional. Explicitou que a avaliação desempenha muitos papéis, mas deveria ter um único objetivo: determinar o mérito ou o valor do que é avaliado. Dessa forma, a diferenciação entre “papéis” e “objetivos” foi uma importante contribuição de seu trabalho para a teorização da avaliação. Tal esforço teórico procurou identificar as expectativas sociais estruturadas em relação às quais os indivíduos deveriam orientar-se dentro e fora da escola. Scriven (1967) também demonstrou que a avaliação formativa deve ocorrer ao longo do desenvolvimento

de um programa ou projeto educacional, para compreender e proporcionar informações úteis aos responsáveis pela implementação do objeto da avaliação. Assim, demonstrou a necessidade de a avaliação formativa ser contínua. No tocante à avaliação somativa, Scriven assumia que deveria ser conduzida ao final de um programa de avaliação, proporcionando, aos seus futuros beneficiários, elementos para julgar sua importância, seu valor e seu mérito. Considerava-a indispensável aos usuários em potencial de qualquer serviço educacional. Por isso, a avaliação somativa de um mesmo programa teria destinatários distintos, beneficiários potenciais, futuros discentes, professores, recrutadores, instituições de fomento, agências de supervisão estatal e gestores locais. Pode-se dizer, com segurança, que a avaliação formativa relaciona-se diretamente à decisão de desenvolver um programa ou à sua modificação/revisão, e a avaliação somativa associa-se à decisão de continuar, alterar ou encerrar um projeto, produto ou programa. Nessa perspectiva, com papéis distintos, os dois tipos de avaliação são relevantes para determinar o mérito ou valor em relação ao objeto educacional avaliado. É interessante notar que os avaliadores, de modo geral, tendem a trabalhar mais com o tipo de avaliação somativa, que se limita à constatação do sucesso e do fracasso do objeto avaliado, em detrimento do uso da avaliação formativa.

Porém, na visão de Scriven (1967), é preciso que os avaliadores acautelem-se nessa questão, pois sem a articulação entre esses dois tipos de avaliação não é possível superar proeminentes lacunas nas práticas avaliativas. Por isso que, na visão do autor, tais perspectivas são complementares e determinam o sucesso da avaliação. Essa problemática é uma das mais visíveis nas políticas de avaliação implementadas no Brasil, tanto na etapa da educação básica quanto na educação superior.

Há autores que defendem que a avaliação formativa e a avaliação somativa podem e devem ser combinadas com as avaliações internas e externas. Casanova (2007) classifica essas avaliações a partir da origem dos avaliadores no desenvolvimento da prática avaliativa. Nesse caso, apropriando-se

de Scriven (1967), compreende-se que, quando a avaliação interna assume a dimensão formativa, a aceitável insuficiência de objetividade do avaliador não tem as mesmas implicações da dimensão somativa. Essa seria mais bem conduzida por um avaliador externo, coordenado por agência externa, mesmo diante dos obstáculos de natureza financeira, logística e de *expertise*.

Já o trabalho de um mesmo avaliador interno, na dimensão somativa, na visão de Scriven (1967), padece de objetividade e credibilidade. Entretanto, é possível superar as dificuldades em torno da questão, desde que o avaliador não esteja relacionado com a instituição, programa ou objeto que está sendo avaliado. Scriven argumentava que as práticas de avaliação nas dimensões formativa e somativa, imersas nas dimensões interna e externa articuladas, são indispensáveis à realização de avaliações de qualidade nos sistemas educacionais.

Esse elemento dialoga com uma importante questão das políticas de avaliação educacional, qual seja, a dimensão formativa da avaliação externa, raramente considerada. A ausência de familiaridade dos avaliadores com o objeto avaliado e o seu contexto, bem como os pré-conceitos que carregam, fragiliza a compreensão dos aspectos observados na avaliação e circunscreve as avaliações externas ao âmbito da perspectiva somativa em detrimento dos aspectos formativos. No caso, há uma nítida inversão de papéis.

Nesse contexto, Scriven (1967) reconheceu que uma das maiores dificuldades da avaliação relaciona-se às inquietações de ordem metodológica. Trata-se, sem sombra de dúvida, de um aspecto relevante a ser considerado. As múltiplas abordagens metodológicas são frutos de distintos posicionamentos epistemológicos, reveladores de preferências paradigmáticas. Esse mesmo autor associou o julgamento de valor em avaliação à tomada de decisão, mas foi bastante criticado – principalmente por aqueles que advogam a democratização da avaliação (MACDONALD, 1974; HOUSE, 1994; SIMONS, 1999) – pelo seu posicionamento em defender que a avaliação deveria ficar restrita aos tomadores de decisão. É importante destacar que não existe, na realidade, um

conjunto de regras e procedimentos para a materialização das avaliações formativa e somativa.

Outro teórico que se destaca entre os adeptos da concepção da avaliação como “juízo de valor” é Cronbach. Segundo Vianna (2000), Cronbach adquiriu notoriedade com a publicação de seu artigo *Course improvement through evaluation*, em 1963. Também não pretendeu propor um modelo avaliativo naquele momento. Contudo, suas ideias foram essenciais para a área de avaliação educacional e decisivas na construção teórica de outros dois expoentes da avaliação, Stake e Scriven, já referidos. Cronbach discutiu aspectos importantes da avaliação, tais como a associação entre a avaliação e o processo de tomada de decisão, o desempenho do estudante como critério para avaliação de cursos, a disposição de técnicas de medida à disposição do avaliador educacional e os diferentes papéis da avaliação educacional, elementos essenciais nas políticas de avaliação na atualidade.

Cronbach foi capaz de visualizar e abordar questões do campo da avaliação educacional, muitas das quais persistem como objeto de estudo e pesquisa até os dias atuais: as decisões possíveis em consequência da avaliação, a avaliação no aprimoramento e na revisão de cursos, a comparação entre cursos, procedimentos de medidas em avaliação, o uso da observação em avaliação educacional, entre outras.

Cronbach (1963) foi um dos primeiros teóricos a defender que a avaliação tem função política, porque, para ele, as reações aos dados da avaliação têm motivação política. Para esse autor, a avaliação era requerida por três tipos de decisão, quais sejam: necessidade de verificação da eficiência do objeto educacional avaliado, diagnóstico das demandas dos alunos associando-as ao planejamento da instrução e julgamento da eficiência do sistema de ensino e dos docentes. Dessa forma, advogava a flexibilidade no planejamento das avaliações para atender à multiplicidade de interesses das suas várias audiências. Esse aspecto evidencia as relações de poder que influenciam diretamente as avaliações em suas distintas modalidades.

Esse cenário revela a frágil relação entre o avaliador, o Estado e os sujeitos do objeto educacional avaliado. Por essa

razão, Cronbach (1982) argumentava que a responsabilidade na e pela avaliação deve ser compartilhada. A avaliação, nessa perspectiva, constitui-se numa espécie de caminho para eliminar complexos mecanismos e proporcionar decisões sobre ações. Por isso, a avaliação deve elucidar questões, dirimir dúvidas e possibilitar ações que resultem da compreensão do objeto avaliado. Cronbach (1982) defendia uma posição mais equilibrada entre aqueles que se posicionavam a favor da abordagem científica e os que propunham uma abordagem holística ou naturalista (GUBA; LINCOLN, 2011). Esses autores classificam a abordagem da avaliação amparada no paradigma positivista como gerencial e científica, enquanto a avaliação baseada na negociação encontra identidade no paradigma naturalista e holístico. Entretanto, é preciso esclarecer que a abordagem naturalista e holística também é científica, porém baseada em outras bases ontológicas, epistemológicas e metodológicas (GUBA; LINCOLN, 2003, 2011).

Para Stake (1967), a avaliação concebida como juízo de valor assume outra dimensão, superando sua característica essencialmente descritiva do currículo, amparada na tradição psicométrica. Em *The countenance of educational evaluation*, Stake (1967) realizou importante avanço no tratamento da questão da avaliação educacional desde Tyler. Stake teve papel fundamental na teorização da avaliação, em especial na metodologia do estudo de caso, contribuindo de forma original para o entendimento da avaliação qualitativa, desencadeando, em contrapartida, a problematização em torno da avaliação institucional. Ressaltou que os atores envolvidos no processo educacional, independentemente de suas posições, acreditam na educação, porém valorizam a avaliação de diferentes maneiras. As ideias de Stake deram origem à avaliação responsiva e proporcionaram o aparecimento da avaliação naturalista, por visualizar dois atos indissociáveis na avaliação: a descrição contextualizada e o julgamento de valor.

Em relação à avaliação naturalista, Stake (1967) atribuiu-lhe um significado amplo, pois nela os sujeitos são observados na sua atividade habitual, em seu *habitat*, e as observações

são apresentadas numa linguagem não técnica, empregando o vocabulário dos sujeitos. Para ele, o avaliador naturalista deve relatar, na linguagem cotidiana, suas observações sobre os sujeitos no contexto natural de suas atividades. Sua posição gerou diversas críticas, entre elas a de que esse tipo de abordagem poderia comprometer o prestígio da teoria das medidas e as tecnologias de análise estatística (HOUSE, 2000). No entanto, Stake argumentava que o principal problema dessa abordagem não se restringia à subjetividade, mas aos custos de realização.

A avaliação responsiva, segundo Stake (1967), volta-se para as atividades de implementação do objeto da avaliação em detrimento de seus objetivos. Nesse caso, associa-se à necessidade de informações propostas por diferentes atores envolvidos, e distintas perspectivas de valor são apresentadas no relatório desse modelo avaliativo, pois a questão central é avaliar como as coisas funcionam. Dessa maneira, esse modelo de avaliação apresenta como característica o aspecto interpretativo, por enfatizar os significados das relações humanas a partir de distintos pontos de vista: o experiencial, por compreender que a realidade é uma obra humana, e o situacional, por estar direcionado aos objetos e às atividades em contextos únicos. Nessa perspectiva, o plano central na avaliação responsiva é a observação para coleta e registro dos dados.

Diversas críticas foram apontadas ao modelo responsivo de Stake, entre elas a de que a observação como método “*pa-dece*” direta e indiretamente dos problemas relacionados ao juízo de valor. Nessa avaliação, os aspectos subjetivos adquirem visibilidade. Stake (1967) argumentava que o valor de um objeto de avaliação não pode estar restrito à mensuração da eficiência dos processos e ao alcance dos resultados de suas metas, critérios e objetivos. O juízo de valor deveria fundamentar-se, também, nos aspectos da qualificação, tratando de avaliar os resultados independentemente das metas, critérios e objetivos preestabelecidos, voltando suas atenções para os impactos dos objetos educacionais nas pessoas. Por isso, Stake defendia a multiplicidade de instrumentos de avaliação, da observação sistemática ao uso de entrevista.

Esse debate propiciou, juntamente com as questões levantadas por Scriven (1967) e Cronbach (1963), a profusão de um conjunto de novos modelos de avaliação: modelos neotylerianos – modelo de fisionomia (STAKE, 1967) e modelo de avaliação da discrepância (PROVUS, 1972); modelos orientados à tomada de decisão – modelo de contexto, insumo, processo e produto (CIPP) (STUFFLEBEAM, 1971); modelos orientados aos resultados – modelo sem referência a objetivos (SCRIVEN, 1973); modelos de neomensuração sobre o pretexto da experimental social (CAMPBELL, 1977); e modelos críticos – modelo de crítica artística (EISNER, 1979).

Em síntese, o apelo para introduzir o juízo de valor no procedimento das avaliações foi o marco dessa fase. O avaliador assume o papel de julgador, tentando manter suas funções técnicas e descritivas do currículo das fases anteriores. Reconheceu-se que era necessário considerar também os próprios objetivos como algo problemático, os quais, tanto quanto o desempenho, deveriam ser submetidos à avaliação.

A “NEGOCIAÇÃO” COMO CONCEPÇÃO DE AVALIAÇÃO

A última fase da avaliação aqui tratada centra-se na concepção da negociação e iniciou-se sob a marca da ruptura epistemológica com os períodos precedentes a partir de uma teorização mais consistente, delineando um momento importante da avaliação, que coincide com sua profissionalização. A partir da década de 1970, a avaliação passa a ser, mais do que antes, uma área de inúmeras práticas, adensando sua relevância acadêmica e política como objeto de estudo, ao tempo em que o Estado passa a ampliar seu braço avaliativo no campo das políticas públicas de educação.

Os avanços teóricos permitiram o amadurecimento do conceito de meta-avaliação,⁴ o aumento da qualificação dos avaliadores e o deslocamento do foco das avaliações dos objetivos para a tomada de decisão, além de promover maior articulação entre partidários dos métodos quantitativos e qualitativos. O elemento mais proeminente desse período foi a superação do sentido restrito da descrição e diagnóstico da avaliação, embora se mantendo a tradição positivista (GUBA; LINCOLN, 2011).

⁴ Termo introduzido por Scriven (1969).

A manutenção dessa tradição caracteriza parte dos avanços e persistências no campo da avaliação educacional. A sofisticação de suas ferramentas, a intensificação de suas atividades e o fortalecimento de sua dimensão política consolidaram-se. Porém, a tradição dos exames de aprendizagem persiste até os dias atuais, embora ressignificada a partir de distintas concepções de avaliação, sob a ação de governos, agências estatais e organismos internacionais, suscitando uma série de debates sobre a eficácia e o alcance das avaliações estandardizadas (AFONSO, 2005; FERNANDES, 2009). A avaliação logrou ampliar seu âmbito de atuação e dimensão semântica devido às complexidades sociais. Dessa forma, a avaliação passa a ser reconhecida como atividade política e de sentido ético, incorporando a “negociação” como um de seus valores e procedimentos centrais, revelando a necessidade da promoção de valores democráticos (HOUSE, 2000).

Segundo Ângulo (1988), nessa fase, amparando-se no desenvolvimento dos estudos e ideias então produzidos pelas ciências sociais, orientados pela fenomenologia social, hermenêutica, interacionismo simbólico e etnometodologia, ocorreu uma mudança paradigmática que passou a questionar os pressupostos teórico-metodológicos da avaliação tecnológica e objetivista. Compreendeu-se que a avaliação não se limitaria ao aspecto técnico da mensuração ou investigação, que não é um processo científico de obtenção neutra de dados. Passou-se a reconhecer as dimensões humana, política, social e cultural que a avaliação necessariamente carrega em suas práticas.

Os modelos de avaliação da fase da negociação fundamentam-se no modelo responsivo desenvolvido por Stake (1967). Nas décadas seguintes, esse modelo permitiu à avaliação assumir distintas perspectivas paradigmáticas, a exemplo do paradigma construtivista e participativo, além de manter a tradição do paradigma racionalista, ancorado no positivismo. Nesse contexto, em ambos os casos, ressignificou-se o conceito de “participação” (MACDONALD, 1974) e se consideraram, como aspecto fundamental, as informações fornecidas pelos grupos de interesse nas avaliações.

Segundo Guba e Lincoln (2011), o campo da avaliação passa a considerar os atores envolvidos nos processos avaliativos,

seus valores e suas construções. Reconhecem a ideia de que todos os grupos de interesse colocados em jogo, por uma avaliação, devem apresentar suas demandas, independentemente do sistema de valores que adotam.

Para MacDonald (1995), nessa linha de argumentação, os modelos de avaliação podem ser adaptados de várias maneiras para conceder aos grupos de interesse exercício de poder ou privá-los dele. Dessa forma, busca-se, também, o envolvimento seletivo dos interessados no projeto e na implementação da avaliação. A seleção das perguntas operatórias da avaliação pode se apresentar num jogo aberto ou fechado. No jogo aberto, as tomadas de decisão são compartilhadas entre os sujeitos, avaliador e grupos de interesses. No jogo fechado, o processo restringe-se ao gestor e ao avaliador. Esse jogo revela e esconde determinadas relações de poder. Os interessados podem ser privados de poder no tocante à disseminação dos resultados das avaliações, uma vez que conhecimento é poder (MACDONALD, 1995).

Guba e Lincoln (2011) advogam a tese de que, na concepção de avaliação como negociação, o avaliador assume uma nova postura em relação às três primeiras fases da avaliação. Tomando como princípio da negociação o respeito aos diferentes conjuntos de valores, o avaliador passar a ter o papel de “orquestrador do processo de negociação”, bem diferente do papel de técnico que coleta informações. Os anos 1980 foram marcados por uma série de embates entre paradigmas, no entanto, ao final, logrou êxito a tese da “harmonização paradigmática” (ARREDONDO; DIAGO, 2009), baseada na tese da coexistência de abordagens, fundada em fatos empíricos. Por essa razão, foram incorporadas à avaliação novas noções derivadas de outros campos disciplinares, o que aprofundou seu sentido e tornou mais complexa sua conceituação. Nos anos 1990, fomentou-se, mais do que nunca, uma avaliação globalizada, integradora e formativa, aprofundando o sentido holístico (ARREDONDO; DIAGO, 2009).

Segundo Silva (2015), é possível reconhecer que o desenvolvimento da concepção de avaliação como negociação no campo da avaliação educacional é uma tendência que acompanha a ampliação da complexidade dos regimes de

regulação – regulação burocrático-profissional e regulação pós-burocrática (MAROY; DUPRIEZ, 2000; MAROY, 2004) – no contexto das políticas de avaliação, compreendendo o fluxo dos problemas, o fluxo de possíveis soluções, o fluxo das condições políticas favoráveis.

É possível compreender a regulação burocrático-profissional como uma aliança entre Estado e os profissionais da educação; já o modelo regulatório pós-burocrático está disposto em torno de dois eixos: o do “quase-mercado” e o do “Estado avaliador”. Esses eixos baseiam-se em modelos de governança mais abertos, que abarcam distintas dimensões, compartilhando traços que os opõem ao regime burocrático-profissional. No quase-mercado, o Estado não se retira do campo educacional. Continua com o desígnio de deliberar as finalidades e o currículo do sistema educacional. No caso do Estado avaliador (NEAVE, 1988), conjectura-se do mesmo modo para que as finalidades e os programas do sistema de ensino sejam deliberados de forma centralizada e para que as unidades educacionais adquiram certa autonomia de gestão pedagógica e financeira.

Esse movimento regulatório no campo da avaliação educacional acompanha, também, o que pode ser considerado “janelas de oportunidades” (MULLER; SUREL, 2002), dialogando com distintos modelos de racionalidade – absoluta, limitada, incremental e fluxos múltiplos (SECCHI, 2010) – no tocante aos processos de tomadas de decisão no sistema educacional, em especial nas políticas de avaliação.

Essas racionalidades podem ser caracterizadas da seguinte maneira: a racionalidade absoluta, configurada pelas condições cognitivas da certeza, baseada na modalidade do cálculo e no critério da decisão da otimização; a racionalidade limitada, condições cognitivas da incerteza, amparada na modalidade da escolha da comparação das alternativas com as expectativas e no critério da satisfação; a racionalidade incremental, nas condições cognitivas da parcialidade (interesses), na modalidade do ajuste mútuo de interesses e no critério de decisão do acordo; a racionalidade dos fluxos múltiplos, pelas condições cognitivas da ambiguidade, apoiada na modalidade de escolha do encontro de soluções e problemas e no critério de decisão casual.

A concepção de “negociação” abraça, preferencialmente, a racionalidade incremental, pois se ampara no desenvolvimento de práticas avaliativas com a lógica da parcialidade (interesses) e da decisão baseada no acordo. Todavia, esse aspecto relacionado aos variados modelos de tomada de decisão indica tensões no desenvolvimento das práticas avaliativas, especialmente no desenho que os modelos de avaliação assumem nos processos de implementação – arco temporal e cenário dos resultados concretos – nas políticas de avaliação.

Fernandes (2009) e Guba e Lincoln (2011) alertam para o fato de que a concepção de avaliação como negociação não está isenta de limitações, admitindo-se, no futuro, a revisão de seus pressupostos, conceitos e métodos ou mesmo seu esgotamento. Uma das maiores dificuldades desse período reside no fato de os atores institucionais tornarem a prática da avaliação formativa numa realidade palpável nas instituições educacionais.

EMBATES TEÓRICOS NO CAMPO DA AVALIAÇÃO

Na construção histórica do campo da avaliação, é evidente que o conceito de avaliação não é uniforme nem monolítico, e pode ser considerado o agrupamento de inúmeros fatores distintos entre si, como um esforço para formular um entendimento coerente, consistente e abrangente sobre as práticas de avaliação. Cabe destacar, também, a avaliação como uma ideia/prática muito discutida, criticada, desvirtuada e muito superficialmente compreendida; todavia, tornou-se discurso hegemônico ao nível dos fazedores de política educacional. Por isso, a tentativa de sistematizar as concepções de avaliação é tarefa complexa que, inevitavelmente, deixará em segundo plano, ou sem a ênfase necessária, aspectos importantes. Consideramos, mesmo ciente das limitações, que vale a pena tentar produzir um quadro-síntese das tendências do campo da avaliação educacional.

QUADRO 2 – Síntese das tendências do campo disciplinar da avaliação

CONCEPÇÃO DE AVALIAÇÃO	TENDÊNCIAS CIRCUNSCRITAS AO PERÍODO	TENDÊNCIAS PARA ALÉM DO PERÍODO
Mensuração	Tradição examinista Utilização de testes e exames Testes de memória Mensuração dos conhecimentos dos alunos (objeto)	<ol style="list-style-type: none"> 1. Tendência ao gerencialismo 2. Teoria social funcionalista 3. Paradigma racionalista de investigação 4. Abordagem predominantemente quantitativa 5. Perspectiva técnica da avaliação
Objetivos	Avaliação de programas, conteúdos, estratégias de ensino e os padrões organizacionais Avaliação para mudança do padrão de comportamento Alunos tornam-se objeto da avaliação Avaliação do currículo	
Juízo de valor	Juízo de valor sobre o mérito do objeto de avaliação Problematização dos objetivos dos programas educacionais Reconhecimento de distintos papéis da avaliação Avaliação como tomada de decisão Profusão de modelos avaliativos	
Negociação	Profusão de modelos avaliativos Avaliação como tomada de decisão Avaliação democrática Avaliação formativa Avaliação global	<ol style="list-style-type: none"> 1. Tendência holística 2. Teoria social pluralista 3. Paradigma dialógico-democrático 4. Abordagem qualitativa, modelo naturalista 5. Perspectiva política da avaliação
Hibridização	Combinação, harmonização ou conflitos das/entre concepções de avaliação Ênfase na avaliação externa dos sistemas de educação Debate, ascendência e articulação com políticas regulatórias Atuação estatal: reformas da educação	<ol style="list-style-type: none"> 1. Embates políticos recorrentes 2. Disputas entre as concepções gerenciais e democráticas de avaliação 3. Conflitos entre as formas de legitimidade técnica e política 4. Distintos modelos de tomada de decisão

Fonte: Elaboração dos autores.

Sintetizamos, no Quadro 2, o sentido e o significado das concepções de avaliação, construídos em distintos contextos históricos, apontando algumas das principais tendências da avaliação no âmbito da constituição do campo da avaliação educacional, inferindo determinadas características e sentidos da concepção de avaliação para o contexto atual, até mesmo com a proposição teórica de uma nova tendência, denominada concepção híbrida.

Parece evidente que houve consideráveis desdobramentos teórico-metodológicos quando consideramos as quatro fases estudadas. Na fase da mensuração, a coleta sistemática de dados a respeito dos indivíduos tornou-se possível com o desenvolvimento de instrumentos amparados na psicometria. Na fase da avaliação dirigida por objetivos, a avaliação

logrou avaliar programas, conteúdos, estratégias de ensino, padrões organizacionais e currículo. Na fase da avaliação como juízo de valor, passou-se a demandar dela o julgamento, tanto do mérito quanto do objeto avaliado, considerando a relevância do processo avaliativo. Entretanto, as três primeiras concepções de avaliação descritas apresentam como características: tendência a guiar-se pelo gerencialismo, incapacidade de acomodar o pluralismo de valores, orientação paradigmática racionalista com ênfase numa abordagem predominantemente quantitativista da avaliação, filiação às teorias sociais funcionalistas e *approach* tecnicista da avaliação.

De acordo com Simons (1999), nas práticas avaliativas e/ou políticas de avaliação caracterizadas pelo gerencialismo, um dos problemas frequentes refere-se à proeminência dos parâmetros e delimitações estabelecidos pelas autoridades e/ou gestores dos sistemas, os quais são efetivamente eximidos de responsabilidade, seus métodos de gestão não são avaliados, a eficiência e a eficácia da gestão não são questionadas e não há responsabilização sobre as tomadas de decisão realizadas.

Dessa maneira, a relação entre avaliador e gestor tende a ser retórica e formal. O gestor disputa o poder quando tem a capacidade de estabelecer as perguntas que a avaliação irá responder, a forma como os dados serão coletados e interpretados e a quem os resultados serão divulgados. É evidente que, em caso de desacordo, a decisão final será do próprio gestor. Essa perspectiva poderá ocorrer em distintos graus de alcance nas modalidades de avaliação desenvolvidas pelos avaliadores nas políticas de avaliação e dependerá, essencialmente, da perspectiva da cultura de avaliação adotada.

MacDonald (1995) compreende que essa tendência diminui a capacidade de agir dos sujeitos avaliados que também querem obter respostas para outras perguntas e outras interpretações. Porquanto, o gestor é levado ao mais alto pedestal de decisão e poder. Observa-se, também, que a relação característica entre gestor e avaliador tende a ser privadora de direitos. Várias estratégias têm sido adotadas pelos avaliadores para superar esse problema. Entretanto, no caso da educação superior no Brasil, essa questão é posta em outro contexto.

Pelas razões apresentadas, a relação entre agências (de avaliação) e avaliador apresenta, no geral, um frágil equilíbrio. Scriven (1983), problematizando essa questão, propôs, como solução, o envolvimento com um modelo de avaliação que fizesse perguntas supostamente interessantes para os sujeitos e que a avaliação prestasse contas a esse grupo. Essa abordagem avança por visualizar a relevância de outros grupos além dos gestores (HOUSE, 2000).

Nenhuma das três primeiras concepções de avaliação – mensuração, objetivos e juízo de valor – levou em consideração as diferenças valorativas, ou seja, a capacidade de acomodar o pluralismo de valores. Segundo Bobbio (2010), o pluralismo é caracterizado pelas rivalidades tradicionais entre indivíduos que se associam em grupos, tais como: partidos políticos, grupos étnicos, grupos de gênero, grupos de geração, entre outros, para satisfazer seus interesses, permitindo que inúmeros interesses se manifestem e se contraponham. No caso da avaliação, House (2000) o define como conflitos de interesses aderentes às práticas avaliativas. O reconhecimento do pluralismo no campo da avaliação adveio somente no final da década de 1960 e consolidou-se nas décadas posteriores. O apelo à prática plural na construção de juízos de valor na avaliação ocorreu, pela primeira vez, quase na mesma época em que houve certo reconhecimento do pluralismo como teoria das políticas sociais (HOUSE, 2000). Essa perspectiva vai de encontro ao que se defendeu por muito tempo: a confiança nos resultados de uma avaliação amparada na metodologia científica, que se acreditava ser isenta de valores.

Com o pluralismo, na visão de House (2000), amadurece-se a ideia de que o avaliador não tem garantias sobre o controle de como os resultados da avaliação serão empregados, isentando-o, a depender da abordagem – fenomenológica ou hermenêutica –, de certa responsabilidade no tocante a essa questão. Dessa forma, todo ato de avaliação se torna um ato político, e a pretensão de isenção de valores não é defensável. Compreende-se, então, que o pluralismo de valores na sociedade é uma questão crucial a ser levada em consideração na avaliação, assim como o é nas políticas públicas, uma vez que é ou deve ser constitutivo das sociedades democráticas e multiculturais.

Segundo Guba e Lincoln (2011), o comprometimento com o paradigma racionalista de investigação revela a perspectiva de que as premissas desse método caracterizam-se axiomáticamente como verdadeiras, dada como realidade objetiva que incumbe responsabilidade à ciência de descrever tal realidade e revelar suas leis. Também distancia o pesquisador/avaliador do ponto de vista da neutralidade do fenômeno investigado/avaliado, para não influenciar os resultados. É como se o pesquisador/avaliador pudesse prever e controlar à vontade e obter sua legitimidade por meio dos métodos das ciências positivas (ÂNGULO, 1988).

Em síntese, praticamente todas as concepções de avaliação dos três primeiros períodos usam o paradigma racionalista para orientar seu trabalho metodológico. Segundo House (1992, 2000), Simons (1999) e Guba e Lincoln (2011), várias críticas foram e são feitas ao uso do paradigma racionalista nas avaliações, em função dos resultados indezíveis, entre eles a descontextualização, a dependência em relação à mensuração, a coerção da verdade e a suposta isenção de valores.

Um dos mais evidentes problemas da descontextualização refere-se a avaliar determinada política ou programa como se não existisse em um contexto concreto, mas apenas sob as condições de controle que vigoram a partir da implementação da avaliação. Na visão dos autores citados, a descontextualização é um equívoco, pois essas generalizações não são possíveis. A descontextualização é um dos motivos pelos quais as avaliações são apontadas como despropositadas em nível local, ou seja, a falta do uso dos resultados, e nenhuma das três fases da avaliação – mensuração, objetivos e juízo de valor – lida com esse problema.

A dependência em relação à mensuração é caracterizada pelo rigor em que esse paradigma se apoia e à rigidez concernente à coleta de dados. Os dados quantificáveis também atenuam o problema associado ao prognóstico e ao controle associados ao uso de instrumentos matemáticos e estatísticos supostamente neutros e eficazes, na busca da legitimidade técnico-científica. Dessa forma, o que não pode ser mensurável, não é real.

Segundo House (2000), o problema da coerção da verdade no paradigma racionalista é algo inquietante para os avaliadores e persiste no campo da avaliação até os dias atuais. Na perspectiva tradicional, os dados não podem ser questionados e os métodos científicos são utilizados para reforçar e apoiar as tendências gerencialistas, circunscrevendo a avaliação ao âmbito do funcionalismo em detrimento da perspectiva do pluralismo de valores, aproximando avaliadores da gestão. Na perspectiva tradicional da avaliação, a ciência pretensamente neutra libera o avaliador de qualquer responsabilidade ética sobre qualquer resultado da avaliação ou do seu uso.

Observamos, na análise das políticas de avaliação contemporâneas, um embate cada vez mais acirrado entre concepções gerencialistas e democráticas na gestão dos sistemas educacionais, o qual remete ao panorama da crise fiscal do Estado e ao debate em relação ao compromisso financeiro-orçamentário estatal com o campo educacional. Esse debate, materializado nas políticas de avaliação, tem posto em evidência um momento de transição nas tendências das concepções de avaliação, com avanços e recuos importantes no que concerne ao alcance (dimensão técnica, política e cultural), objeto (aprendizagem, escola, professor, sistema educacional, políticas educacionais, entre outros) e lógica (modos e formas) da avaliação, o qual, neste trabalho, optamos por denominar tendência híbrida. Essa tendência caracteriza-se pela combinação de concepções, desvelando ora harmonização ora conflitos das/entre concepções de avaliação, com embates políticos recorrentes entre modelos gerenciais e democráticos de avaliação, com distintas formas de legitimidade técnica e política, materializados em variados enfoques de tomada de decisão.

Antes de avançarmos em nossa argumentação, é preciso esclarecer que o próprio desenvolvimento das concepções de avaliação no campo da avaliação educacional é resultado desse processo de tendência híbrida, superando qualquer reducionismo linear/sequencial entre o desenvolvimento das concepções de avaliação. Na verdade, disputas e tensões no campo da avaliação educacional não são um fenômeno de

sua atual fase, pois mesmo nos períodos anteriores, tais disputas já existiam, apesar de determinadas concepções serem dominantes. Hoje, o aparente hiato de dominação e o acirramento entre as concepções vêm contribuindo para o desenvolvimento e a propagação da visibilidade dessa tendência híbrida.

Essa tendência deve, nos próximos anos, apontar novas intenções no campo da avaliação educacional, desvelando o campo como um domínio aberto e fértil ao aparecimento de tendências e concepções para a avaliação. A destacada visibilidade das avaliações externas articulada ao desenvolvimento das políticas de regulação da educação possivelmente configura-se como indutora dessa tendência de desenvolvimento da concepção híbrida, pondo à prova a atuação estatal e o alcance das reformas da educação.

CONSIDERAÇÕES FINAIS

No decorrer deste artigo, realizou-se uma discussão a respeito das concepções de avaliação, com a finalidade de destacar e refletir que, no contexto presente, mesmo de forma transitória, a compreensão da avaliação educacional está associada ao hibridismo entre as concepções de avaliação como mensuração, objetivos, juízo de valor e negociação. Como caracterizado, a associação e os embates entre essas concepções revelam tensões, ao nível da delimitação conceitual, das práticas, dos modelos, desenvolvendo, no campo da avaliação educacional, possibilidades para as políticas de avaliação assumirem as concepções da mensuração, objetivos e juízo de valor, numa dimensão política que caracterize a avaliação como um processo de negociação.

Mesmo diante desse cenário adverso, é imprescindível avaliar o sistema educacional, seus resultados e processos, para tomar decisões que orientem, com legitimidade técnica e política, os fluxos de problemas, de soluções, observando as condições políticas favoráveis e a janela de oportunidades para as políticas de avaliação. Talvez o melhor fosse romper com a ortodoxia e particularidade de cada tendência de concepção de avaliação, indicando, como elemento na concepção híbrida, o princípio

da flexibilidade paradigmática nas práticas avaliativas, em face da complexidade do fenômeno educativo desvelado nas políticas de avaliação, ou seja, reconhecendo a necessidade de equilibrar a ênfase das políticas de avaliação entre a qualificação, objetivos, metas e processos.

REFERÊNCIAS

AFONSO, A. J. *Avaliação educacional: regulação e emancipação*. São Paulo: Cortez, 2005.

ÂNGULO, J. F. Evaluación de programas sociales de la eficacia a la democracia. *Revista de Educación*, Madri, v. 286, p. 193-207, mayo/ago. 1988.

ARREDONDO, C. S.; DIAGO, J. C. *Avaliação educacional e promoção escolar*. São Paulo: Editora da Unesp, 2009.

BOBBIO, N. *Dicionário de política*. 13. ed. Brasília, DF: Editora UnB, 2010. p. 656. v. 2.

CAMPBELL, D. T. Keeping the data honest in the experimenting society. In: MELTON, H. W.; WATSON, D. H. G. (Org.). *Interdisciplinary dimensions of accounting for social goals and social organizations: a conference of the Department of the Accountancy, University of Illinois, Urbana-Champaign*. Urbana-Champaign: University of Illinois, 1977. p. 37-76.

CASANOVA, M. A. *Manual de evaluación educativa*. Madri: La Muralla, 2007.

CRONBACH, L. J. Course improvement through evaluation. *Teachers College Record*, New York, v. 64, n. 8, p. 672-683, May 1963.

CRONBACH, L. J. *Designing evaluation of educational and social programs*. San Francisco: Jossey-Bass, 1982.

EISNER, E. *The education imagination*. New York: Macmillan, 1979.

FERNANDES, D. *Avaliar para aprender: fundamentos, práticas e políticas*. São Paulo: Editora da Unesp, 2009.

GUBA, E. G.; LINCOLN, Y. S. Controvérsias paradigmáticas, contradições e confluências emergentes. In: DENZIN, N. K.; LINCOLN, Y. S. *O planejamento da pesquisa qualitativa: teorias e abordagens*. Tradução de Sandra Regina Netz. Porto Alegre: Artmed, 2003. p. 169-192.

GUBA, E. G.; LINCOLN, Y. S. *Avaliação de quarta geração*. Tradução de Beth Honorato. Campinas: Unicamp, 2011.

HOUSE, E. R. Tendências em educação. *Revista de Educación*, Madri, n. 299, p. 43-56, sept./dic. 1992.

HOUSE, E. R. *Evaluación, ética y poder*. Madri: Morata, 1994.

HOUSE, E. R. *Evaluación, ética y poder*. Madri: Morata, 2000.

HYPOLITO, A. M. Gestão do trabalho docente e qualidade da educação. In: SIMPÓSIO BRASILEIRO DE POLÍTICA E ADMINISTRAÇÃO DA EDUCAÇÃO, 23., Porto Alegre, 2007. *Anais...* Porto Alegre: Associação Nacional de Política e Administração da Educação, 2007. Disponível em: <www.anpae.org.br/congressos_antigos/simposio2007/16.pdf>. Acesso em: set. 2016.

LEFRANÇOIS, G. R. *Teorias da aprendizagem*. São Paulo: Cengage Learning, 2015.

MACDONALD, B. Evaluation and the control of education. In: MACDONALD, B.; WALKER, R.; SAFARI I. *Innovation, evaluation, research and the problem of control*. Norwich: Centre for Applied Research in Education/University of East Anglia, 1974. p. 23-27.

MACDONALD, B. La evaluación como profesión de servicio público: perspectivas de futuro. In: SÁEZ, M. (Coord.). *Conceptualizando la evaluación en España*. Alcalá de Henares: Universidad Alcalá de Henares, 1995. p. 15-23.

MADAUS, G.; STUFFLEBEAM, D. Program evaluation: a historical overview. In: STUFFLEBEAM, D.; MADAUS, G.; KELLAGHAN, T. (Org.). *Evaluation models: viewpoints on educational and human services evaluation*. Dordrecht: Kluwer, 2000. p. 97-112.

MAROY, C. Sociologie des regulations de l'enseignement: une comparaison européenne. *Recherches Sociologiques*, Louvain, v. 35, n. 2, p. 1-4, mai/juillet 2004

MAROY, C.; DUPRIEZ, V. La regulation dans les systems scolaires. *Revue Française de Pédagogie*, Lyon, v. 130, n. 1, p. 73-87, janv./mars 2000. Disponível em: <http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-807_2000_num_130_1_1054>. Acesso em: dez. 2016..

MULLER, P.; SUREL, Y. *Análise das políticas públicas*. Pelotas: Educat, 2002.

NEAVE, G. On the cultivation of quality, efficiency and enterprise: an overview of recent trends in higher education in Western Europe, 1986-1988. *European Journal of Education*, Oxford, v. 23, n. 1/2, p. 7-23, Jan./Apr. 1988.

PROVUS, M. M. *The discrepancy evaluation model*. In: TAYLOR, P. A.; COWLEY, D. M. *Readings in curriculum evaluation*. Dubuque, Iowa: Wm. C. Brown Co. Publishing, 1972. p. 242-283.

SCRIVEN, M. The methodology of evaluation, In: TYLER, R. W.; GAGNE, R. M.; SCRIVEN, M. *Perspectives of curriculum evaluation*. Chicago: Rand McNally, 1967. p. 39-83. (American Educational Research Association Monograph Series on Curriculum Evaluation, n. 1).

SCRIVEN, M. An introduction to metaevaluation. *Educational Products Report*, n. 2, p. 36-38, 1969.

SCRIVEN, M. Goal-free evaluation. In: HOUSE, E. R. (Ed.). *School evaluation: the politics and process*. Berkeley, California: MacCutchan Publishing Corporation, 1973. p. 25-42.

SCRIVEN, M. Evaluation ideologies. In: MADAUS, F. G.; SCRIVEN M. S.; STUFFLEBEAM, D. L. (Org.). *Evaluation models*. Boston: Springer, 1983. p. 3-112.

SECCHI, L. *Políticas públicas: conceitos, esquemas de análise, casos práticos*. São Paulo: Cengage Learning, 2010.

SILVA, A. L. *Avaliação institucional no Sistema Nacional de Avaliação da Educação Superior (Sinaes)*. 2015. 372 f. Tese (Doutorado em Educação) – Centro de Educação, Universidade Federal de Pernambuco, Recife, 2015.

SIMONS, H. *Evaluación democrática de instituciones escolares*. Madrid: Morata, 1999.

STAKE, R. E. The countenance of educational evaluation. *Teachers College Record*, New York, v. 68, n. 7, p. 523-540, Jan./Apr. 1967.

STERNBERG, R. J. *Psicologia cognitiva*. Porto Alegre: Artes Médicas Sul, 2000.

STUFFLEBEAM, D. L. The relevance of the CIPP Evaluation Model for Educational Accountability. *Journal of Research and Development in Education*, v. 5, p. 19-25, 1971.

STUFFLEBEAM, D. L.; SHINKFIELD, A. J. *Evaluation sistemática*. Guia teórica e prática. Madrid: Paidós, 1987.

TAYLOR, F. W. *Princípios da administração científica*. 8. ed. São Paulo: Atlas, 1990.

TYLER, R. W. The objectives and plans for a national assessment of educational progress. *Journal of Educational Measurement*, Chicago, v. 3, n. 1, p. 3-15, Feb. 1966.

VIANNA, H. M. *Avaliação educacional: teoria, planejamento e modelos*. São Paulo: Ibrasa, 2000.

WATRAS, J. The eight-year study: from evaluative research to a demonstration project, 1930-1940. *Education Policy Analysis Archives*, Phoenix, v. 14, p. 21, Aug. 2006.

Recebido em: 21 OUTUBRO 2017

Aprovado para publicação em: 17 ABRIL 2018



This content is licensed under a Creative Commons attribution-type BY-NC.

