

<http://dx.doi.org/10.18222/ee.v30i73.5879>

“QUÃO ‘BOM’ É SUFICIENTE?”¹ DEFINIÇÃO DE CRITÉRIOS AVALIATIVOS DE VALOR E MÉRITO²

ADRIANA BAUER¹

RESUMO

Este texto procura discutir conceitos associados à noção de valor e à ação de valorar na avaliação de programas, tendo em vista a necessidade de fazer julgamentos acerca da qualidade de uma política, programa, projeto ou ação social que lhe é inerente. Busca-se contribuir para fomentar a discussão a respeito da prática de valoração nas avaliações de programas sociais. Para isso, é feita uma revisão das literaturas estadunidense e brasileira, a fim de identificar como as questões de valoração têm sido propostas teoricamente e explicitadas em pesquisas avaliativas. São examinados, de forma exploratória, relatórios de avaliação e artigos sobre avaliações de programas, com foco na área educacional. Conclui-se, com base na análise bibliográfica realizada, que a discussão sobre a atribuição de valor raramente tem sido feita na literatura nacional de forma explícita, ressaltando-se a necessidade de seu aprimoramento.

PALAVRAS-CHAVE AVALIAÇÃO DE PROGRAMAS • CRITÉRIOS DE AVALIAÇÃO • MÉRITO • VALOR.

1 A sugestiva pergunta é inspirada em questionamento feito por Cousins e Shulha (2008), em seu texto sobre as complexidades inerentes à escolha de critérios e padrões de valoração.

2 Dedico este texto a Chris Coryn, por ter me apresentado, durante suas aulas no Evaluation Center, as discussões sobre valoração e as teorias avaliativas, propiciando um intenso aprendizado acerca da avaliação de programas; a Nelson Antonio Simão Gimenes, colega da Fundação Carlos Chagas que, há anos, ouvindo as discussões que fiz sobre mérito e valor em um curso, sugeriu que as sistematizasse na forma de texto; e a Paulo de Martino Jannuzzi, cuja produção intelectual proficua no campo da avaliação, do monitoramento e dos indicadores sociais tem iluminado meus estudos e inspirado minhas formas de pensar acerca de importantes problemas relativos ao campo.

1 Fundação Carlos Chagas (FCC) e Universidade de São Paulo (USP), São Paulo-SP, Brasil; <http://orcid.org/0000-0002-5942-9181>; adbauer@fcc.org.br

“¿CUÁN ‘BUENO’ ES SUFICIENTE?” DEFINICIÓN DE CRITERIOS EVALUATIVOS DE VALOR Y MÉRITO

RESUMEN

Este texto busca discutir conceptos asociados a la noción de valor y a la acción de valorar en la evaluación de programas, teniendo en cuenta la necesidad de juzgar acerca de la calidad de una política, programa, proyecto o acción social que le es inherente. Se busca contribuir a fomentar la discusión acerca de la práctica de valoración en las evaluaciones de programas sociales. Para ello, se hace una revisión de la literatura estadounidense y brasileña, a fin de identificar cómo las cuestiones de valoración han sido propuestas teóricamente y explicitadas en investigaciones de evaluación. Se examinan, de forma exploratoria, informes de evaluación y artículos sobre evaluaciones de programas, con enfoque en el área educativa. Se concluye, en base al análisis bibliográfico realizado, que la discusión sobre la atribución de valor se realiza raras veces en la literatura nacional e internacional de forma explícita, lo que resalta la necesidad de su perfeccionamiento.

PALABRAS CLAVE EVALUACIÓN DE PROGRAMAS • CRITERIOS DE EVALUACIÓN • MÉRITO • VALOR.

“HOW ‘GOOD’ IS GOOD ENOUGH?” DEFINING EVALUATIVE WORTH AND MERIT CRITERIA

ABSTRACT

This text discusses concepts associated with the notion of value and the action of valuing in program evaluation, considering the need to make judgments about the quality of a policy, program, project or social action that is inherent to it. It seeks to contribute to foment the discussion about valuation in the evaluation of social programs. To that end, it reviews the US and Brazilian literature in order to identify how valuation questions have been theoretically proposed and made explicit in evaluative research. Reports and articles on program evaluation with a focus on education are examined. It concludes, based on this bibliographical analysis, that the discussion about the attribution of value has seldom been conducted in the national literature in an explicit way, which reinforces the need for improving it.

KEYWORDS PROGRAM EVALUATION • EVALUATION CRITERIA • MERIT • WORTH.

INTRODUÇÃO

Ruim é ruim e bom é bom, e é a tarefa da avaliação dizer qual é qual. (SCRIVEN, 1986, p. 19)

A polêmica citação que inicia este artigo, tomada de um texto de Michael Scriven (1986), refere-se a um dos aspectos constituintes da avaliação de políticas, programas, projetos e ações sociais: a atribuição de valor à entidade avaliada,³ a partir da qual são (ou deveriam ser) tomadas decisões.

O aspecto valorativo parece indissociável da avaliação de programas, como pode ser observado inclusive pela própria definição do que é esse tipo de avaliação, na perspectiva de diversos autores. Em geral, se diz que avaliação é:

Determinação do valor ou mérito de um objeto de avaliação (seja o que for que estiver sendo avaliado). Uma definição mais extensa diria que avaliação é identificação,

³ Essa entidade pode ser um objeto, um programa, um projeto, uma ação ou uma pessoa.

esclarecimento e aplicação de critérios defensáveis para determinar o valor (valor ou mérito), a qualidade, a utilidade, a eficácia ou a importância do objeto avaliado em relação a esses critérios. (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 35)

O processo de determinar o mérito, valor ou significância de algo, ou o produto desse processo. (SCRIVEN, 1991, p. 139)

[...] fazer um julgamento de valor a respeito de uma intervenção ou sobre qualquer um de seus componentes, com o objetivo de ajudar na tomada de decisões. Este julgamento pode ser resultado da aplicação de critérios e de normas (avaliação normativa) ou se elaborar a partir de um procedimento científico (pesquisa avaliativa). (CONTANDRIOPOULOS *et al.*, 1997, p. 31)

Se há certa concordância acerca da ideia de que mérito e valor são predicados básicos da avaliação de programas, cabe especificar o que significa cada um desses conceitos e que relações se tecem entre eles, bem como suas diferenças.

Primeiramente, nota-se que, segundo Guba e Lincoln (1980, p. 61), a distinção entre mérito (*merit*) e valor (*worth*) é arbitrária, visto que ambos são elementos constituintes da noção de valor (*value*).⁴ Já para Davidson (2005, p. 247), “mérito é a qualidade absoluta ou relativa de algo, seja globalmente ou em relação a um critério particular”, sem que a autora faça distinção entre mérito e valor.⁵

De todo modo, mérito tem sido compreendido como o *valor intrínseco* de algo e seu julgamento pode variar de acordo com os diferentes juízes que

4 Os autores explicam suas escolhas teóricas e analisam sua proposta em relação aos postulados de Scriven, outro teórico importante da avaliação de programas estadunidense, que também reconhece que a noção de valor abriga mais de uma dimensão. Não é objetivo deste texto aprofundar a discussão semântico-conceitual desses autores, mas o leitor pode encontrá-la em Guba e Lincoln (1980, p. 61).

5 Esse texto não trata de aspectos metodológicos para a determinação de mérito ou valor, ainda que se reconheça que os desenhos metodológicos utilizados nas avaliações dependem dos valores públicos que lhes são subjacentes (JANNUZZI, 2016), bem como dos valores e interesses pessoais dos envolvidos (VIANNA, 2005). Para os fins desse trabalho, basta comentar que, segundo Davidson (2005), o julgamento do mérito necessita: da coleta de dados relevantes que permitam ao avaliador julgar o que é avaliado em uma dada dimensão de qualidade; e da síntese das performances por ele obtidas quando se trata de uma dimensão multicriteriada.

fazem parte do processo avaliativo, visto que diferentes atores tendem a considerar diferentes indicadores de mérito (GUBA; LINCOLN, 1980).

Já o valor (*worth*) tem sido compreendido como o valor de uma entidade avaliada para um dado contexto, ou seja, o *valor extrínseco*. Segundo Mathison (2005), determinar esse valor requer uma compreensão completa do contexto particular em que o programa, projeto ou ação é implantado, bem como conhecimento das qualidades e atributos do que está sendo objeto de avaliação.

Em termos gerais:

- mérito (*merit*): qualidade intrínseca ao programa ou objeto avaliado;
- valor (*worth*): qualidade extrínseca do programa ou objeto avaliado, em relação ao contexto para o qual foi proposto, ao atendimento das necessidades reais do público-alvo, entre outros elementos exteriores à proposta que devem ser considerados em sua avaliação;
- significância (*significance*): pode ser entendida como a importância do programa/projeto/objeto avaliado, ou seja, a conclusão geral sobre o mérito e valor quando todas as considerações pertinentes foram sintetizadas.

Conclui-se, pelo exposto, que há variação nas bases de julgamento tanto por mérito quanto por valor e significância, ainda que a variabilidade seja maior nessas últimas, dado o caráter contextual do valor e da atribuição de significância. Nas palavras de Guba e Lincoln (1980, p. 64, tradução nossa):

[...] essa variação decorre do fato de que os julgamentos de mérito estão vinculados a características intrínsecas da própria entidade e, portanto, relativamente estáveis, enquanto os julgamentos de valor dependem da interação da entidade com algum contexto e, portanto, variam conforme os contextos variam. [...] Se, por outro lado, alguém estiver falando sobre as diferentes dimensões em que o mérito e o valor podem variar, é necessário considerar o tempo, o grau de consenso e os fatores de limite. As avaliações de mérito são feitas em termos de critérios que são relativamente invariáveis ao longo do tempo, enquanto avaliações de valor são feitas em termos de critérios que podem se alterar rapidamente com mudanças sociais, econômicas ou outras condições de curto prazo. Além disso, é prová-

vel que haja um grau relativamente alto de consenso sobre os critérios de mérito, enquanto os critérios que valem a pena dependem, entre outras coisas, de conjuntos de valores que podem ser muito diferentes entre e dentro de grupos sociais e organizacionais.⁶

Deve-se notar, no entanto, que nem todas as três dimensões ocorrem de forma concomitante. Um programa pode ter mérito, mas não ter valor ou significância em um determinado contexto. Se não atende às demandas dos seus *stakeholders*,⁷ diz-se que ele não tem valor. Além disso, todo programa (ou política) que não tenha mérito não possui valor, pois, se não faz bem o que se propôs a fazer, não pode ser empregado para suprir as necessidades de seus *stakeholders*. Um programa mal formulado, nesse sentido, não tem mérito e nem valor.

Caberia, pois, ao avaliador, determinar com clareza o que pretende avaliar (objeto) e as dimensões de qualidade relacionadas a esse objeto, a fim de definir possíveis critérios de mérito e de valor do programa. Há, no entanto, autores que divergem dessa concepção. Segundo Schwandt (2005, p. 443-444), muitos avaliadores discordam que a atribuição de mérito, valor ou significância é inerente à avaliação de programas sociais, advogando que seu papel é, antes, descrever e explicar a realidade, sem emitir julgamento.

Neste texto, defende-se que o debate acerca do estabelecimento dos critérios de mérito, valor e significância de um programa é fundamental para a garantia da transparência pública e da ética nos processos avaliativos, ainda que os diversos autores que se dedicaram ao desenvolvimento de teorias sobre avaliação de programas tenham posturas diferentes sobre a questão.

Shadish, Cook e Leviton (1991) preocuparam-se em analisar, no pensamento de teóricos de diferentes épocas da história recente da avaliação de programas, como a questão da valoração é compreendida e trabalhada. Na obra intitulada *Foundations of program evaluation: theories of practice*, eles analisam

6 Do original: "that variation arises from the fact that judgments of merit are tied to intrinsic characteristics of the entity itself, and are therefore relatively stable, while judgments of worth depend upon the interaction of the entity with some context and thus vary as contexts vary. [...] If, on the other hand, one is talking about the different dimensions along which merit and worth may vary, one needs to consider time, degree of consensus, and boundary factors. Assessments of merit are made in terms of criteria which are relatively invariant over time, while worth assessments are made in terms of criteria that may alter rapidly with changing social, economic, or other short-term conditions. Moreover, there is likely to be a relatively high degree of consensus about merit criteria, while worth criteria depend, among other things, on value sets likely to be very different both between and within social and organizational groupings".

7 Diz-se dos interessados na avaliação, podendo ser o público-alvo do programa ou ação avaliada, os propositores das políticas, seus executores e a sociedade em geral.

estudos de sete autores⁸ que consideram representantes de três diferentes estágios das teorias avaliativas: o estágio 1, representado por Campbell e Scriven, cuja principal preocupação teria sido iluminar a solução dos problemas sociais; o estágio 2, com pensadores como Weiss, Wholey e Stake, em que haveria o comprometimento com a geração de alternativas à avaliação de programas, enfatizando o uso dos resultados e o pragmatismo; e o estágio 3, de Cronbach e Rossi, centrado em teorias que tentam integrar “o passado”, considerando as principais preocupações exploradas nos estágios anteriores.

Em suas análises das obras e teóricos selecionados, Shadish, Cook e Leviton (1991) consideraram cinco diferentes dimensões: construção do conhecimento (como construir conhecimento válido acerca do objeto avaliado); valoração (como atribuir valor aos resultados das avaliações); programa social (que seria compreender as posições dos teóricos acerca de como os programas mudam); uso dos resultados (suas indicações acerca de como usar os resultados no processo da política); e, finalmente, suas posições sobre como organizar a prática avaliativa.

Neste artigo interessa explorar as considerações dos autores sobre a dimensão de valoração que, segundo eles, discute essencialmente o papel ocupado pelos valores e processos de valorar em cada perspectiva teórica, tendo como discussão central os caminhos e as formas pelas quais valores podem ser associados à descrição dos programas. Os autores defendem que os valores são onipresentes na avaliação de programas e, em geral, os teóricos que se preocupam com a questão acreditam ser importante responder se o objeto ou entidade avaliada é bom/adequado, deixando evidentes a noção de “bom” considerada e as justificativas para essa conclusão. Enquanto componente de uma teoria de avaliação, a dimensão valorativa preocupa-se em discutir “como os avaliadores podem tornar os problemas de valor explícitos, lidar com eles abertamente e produzir uma análise sensível às implicações de valores dos programas” (SHADISH; COOK; LEVITON, 1991, p. 47). Os autores se baseiam em Beauchamp para afirmar que as teorias de valoração podem ser metateorias, teorias prescritivas ou teorias descritivas, que definem como segue:

Metateoria: estudo da natureza e justificativa para valorar;
Teoria prescritiva: teoria que defende a primazia de interesses particulares; e

⁸ São eles: Michael Scriven, Donald Campbell, Carol Weiss, Joseph Wholey, Robert Stake, Lee Cronbach e Peter Rossi.

Teoria descritiva: teoria que descreve os valores, sem defender que um deles é melhor.⁹ (SHADISH; COOK; LEVITON, 1991, p. 48, tradução nossa)

Para exemplificar suas colocações, os autores defendem que Michael Scriven é o único teórico que possui uma metateoria de valoração, enquanto Ernest House teria proposto uma teoria prescritiva, advogando a necessidade de uma teoria de justiça estar subjacente às análises dos resultados das avaliações. Argumentam, ainda, que Joseph Wholey propõe uma teoria descritiva, pois defende que cabe ao avaliador apenas tentar compreender e descrever os valores dos propositores de cada política/programa ou ação social. Os autores salientam que teóricos que assumem uma teoria prescritiva têm uma perspectiva crítica e uma “autoridade intelectual” que não combinam com as teorias descritivas, visto que se defrontam com a necessidade de refletir sobre o que é “bom” para a condição humana, para o bem público de forma mais geral, e não apenas no escopo do programa (SHADISH; COOK; LEVITON, 1991, p. 49).

Para os autores de *Foundations of program evaluation*, teóricos do primeiro estágio de avaliação, como Scriven, tendem a definir os critérios avaliativos com base em valores de justiça social, sendo possível assumir que teriam, como horizonte, teorias prescritivas de valoração. Já autores da segunda geração da avaliação, como Carol Weiss e Robert Stake, reconheceriam a natureza plural dos valores, admitindo que os usuários intencionais dos resultados da avaliação contribuam na construção dos critérios de avaliação, em uma perspectiva descritiva da valoração; finalmente, autores representativos da terceira geração de avaliação, como Cronbach, Rossi e Wholey, defenderiam que os critérios valorativos fossem estabelecidos pelos beneficiários dos programas, também em uma perspectiva descritiva.

Apesar de construírem uma teoria da valoração consistente e defenderem que a preocupação sobre como fazer os julgamentos acerca dos programas avaliados deveria fazer parte da teorização sobre avaliação, Shadish, Cook e Leviton reconhecem que este aspecto não tem recebido a atenção necessária, nem em termos teóricos, nem práticos, pelo menos não explícita ou sistematicamente nos trabalhos que foram selecionados. Em suas palavras:

⁹ Do original: “Metatheory: the study of the nature of and justification for valuing; prescriptive theory: theory that advocates the primacy of particular values; descriptive theory: theory that describes values without advocating one as best”.

[...] os avaliadores reconhecem que os valores merecem mais atenção (House, 1980). Quase todos os teóricos neste livro concordam que a avaliação é sobre determinar [a noção de] valor, mérito ou valor,¹⁰ não apenas sobre a descrição de programas. Mas poucos teóricos fazem mais.¹¹ (SHADISH; COOK; LEVITON, 1991, p. 49, tradução nossa)

Essa realidade, descortinada pelos estudiosos, aponta para uma necessidade de discutir os significados e propósitos da própria avaliação de programas, campo que se expandiu nos últimos anos, mas que precisa estar continuamente em discussão, visando ao seu aprimoramento teórico e prático.

AVALIAÇÃO DE PROGRAMAS X PESQUISA AVALIATIVA X ANÁLISE DE POLÍTICA

Conforme se tentou explicitar no início deste artigo, o processo de avaliação de um programa ou política determina, ou deveria determinar, critérios para julgar o valor ou mérito do objeto a ser avaliado a partir de dados coletados por meio de investigação empírica realizada pela aplicação de técnicas das Ciências Sociais.

No entanto, nem toda apreciação acerca de uma política ou programa traz, em seu bojo, um juízo valorativo. Como visto anteriormente, há discordâncias, entre os estudiosos, se toda avaliação deveria levar a um ajuizamento. Mas há, também, estudos que não se dedicam à discussão sobre o mérito ou valor do seu objeto por assumir outra perspectiva teórica de análise. Nesse sentido, antes de prosseguir com a discussão acerca das teorias de valoração, cabe fazer a distinção entre avaliação de programas e análise de políticas, ferramentas de estudo sobre os programas públicos que se diferenciam em vários aspectos, inclusive sobre a questão do julgamento de valor.

Neste trabalho, assume-se que a determinação do mérito ou valor é uma tarefa essencial que diferencia a avaliação de outros tipos de pesquisa ou de

10 Nota da autora: No inglês, há duas palavras que, na tradução para o português, assumem o sentido de valor: *value* and *worth*. No entanto, ainda que a tradução de ambas seja "valor", esses termos têm diferentes significados. *Value* assume um significado mais amplo e refere-se comumente à importância ou significância de algo em um determinado momento e contexto, enquanto *Worth* apresenta um significado mais "pragmático", contábil, muitas vezes se referindo ao custo, à relação custo-benefício, etc. Por exemplo: um determinado objeto pode ter um valor baixo (*worth*), mas um alto valor (*value*) para uma determinada pessoa.

11 Do original: "Yet evaluators acknowledge that values deserve more attention (House, 1980). Nearly all the theorists in this book agree that evaluation is about determining value, merit, or worth, not just about describing programs".

análise a que são submetidas as políticas, independentemente dos propósitos da avaliação, somativos ou formativos, ou de abordagens mais tradicionais ou colaborativas (DAVIDSON, 2005; COUSINS; SHULHA, 2008).

Como já explorado em outro trabalho (BAUER, 2011), o elemento distintivo entre avaliação e pesquisa é justamente a realização de um juízo de valor sobre um dado objeto ou programa, inerente à primeira. Ainda que a pesquisa (*research*) possa produzir evidências sobre valor ou mérito de um dado objeto de estudo, entende-se que essa produção é indireta, secundária, visto que o foco da investigação é a geração de novos conhecimentos para o campo. Já a avaliação de programas (*program evaluation*) se concentra na busca de evidências que possibilitam a realização de um julgamento. Essa mesma preocupação é central, também, nas pesquisas avaliativas.¹²

Os diversos autores parecem concordar que o produto da avaliação e da pesquisa avaliativa é *necessariamente* um juízo de valor a ser utilizado para *tomadas de decisão*.

Essa questão se apresenta de forma diferente quando o analista recorre ao suporte da análise de políticas.

Cabe abrir parênteses para distinguir avaliação (*evaluation*) e análise de política (*policy analysis*). Enquanto a análise de políticas interessar-se-ia principalmente em identificar tendências e procurar princípios generalizáveis, a avaliação de programas tem um caráter de subsídio à tomada de decisões acerca da qualidade do programa, sendo-lhe inerente o ajuizamento de valor. Para Figueiredo e Figueiredo (1986, p. 107), a análise de políticas teria como preocupação o estudo sobre como as decisões são tomadas, sobre a identificação de “que fatores influenciam o processo de tomada de decisões e as características desse processo”. Muller e Surel (2002) apontam que, se inicialmente a análise de políticas teve al-

¹² Cabe abrir parênteses para tratar da distinção existente na literatura entre avaliação e pesquisa avaliativa, discussão datada da década de 1970 e que, hoje, parece superada. A moderna avaliação de programas se disseminou na década de 1960, nos Estados Unidos, tendo como foco os programas educacionais. Tal expansão e investimento nesse tipo de avaliação não foi necessariamente acompanhada de rigor teórico-metodológico. Suchman (1971), preocupado com a qualidade das avaliações que eram produzidas e consciente da influência dos seus resultados sobre a política social americana, propôs a distinção entre avaliação e pesquisa avaliativa: “primeiramente, distinguimos avaliação como o processo social geral de fazer julgamentos, independentemente da base para tal, e pesquisa avaliativa como sendo a aplicação do método científico para coleta de dados sobre o grau em que uma atividade específica alcança um efeito desejado” (SUCHMAN, 1971, p. 45, tradução nossa). Do original: “First, we distinguish between evaluation as the general process of making judgments of worth regardless of the basis for such judgments, and evaluative research as referring to the use of the scientific method for collecting data concerning the degree to which some specified activity achieves some desired effect”. Essa distinção foi incorporada por outros autores. Worthen, Sanders e Fitzpatrick (2004, p. 36), por exemplo, definem pesquisa avaliativa como “qualquer avaliação que empregue uma metodologia rigorosa do campo das ciências sociais”. Na literatura brasileira, encontra-se o respeito aos “cânones científicos mínimos” também nos escritos de Draibe (2001, p. 18). Atualmente, a distinção entre avaliação e pesquisa avaliativa tem sido menos enfatizada na literatura, havendo uma tendência a incorporar a ideia do rigor metodológico ao conceito de avaliação.

guma pretensão de fazer afirmações valorativas sobre as políticas e programas, esse campo de estudos foi redirecionado para análises mais complexas, focadas em entender a ação pública estatal (ou a ausência dessa ação):

Originalmente formulada nos Estados Unidos, como um conjunto de dispositivos de pesquisa, que tinha por ambição fornecer as receitas do “bom” governo, a *policy analysis* distanciou-se progressivamente de sua orientação operacional para tornar-se uma disciplina por inteiro da Ciência Política, progressivamente autonomizada nas estruturas de ensino e de pesquisa. Alimentando pouco a pouco outro olhar sobre o Estado, a análise das políticas públicas contribuiu desde então para mostrar que o Estado não é (não é mais?) esta forma social absoluta na história.

[...]

O trabalho do analista deve, portanto, levar em conta, ao mesmo tempo, as intenções dos tomadores de decisão, mesmo se estas são confusas, e os processos de construção do sentido na prática ao longo da fase de desenvolvimento da ação pública. No entanto, em todos os casos o pesquisador deverá cuidar para não tomar o lugar dos atores da política na determinação do sentido da política. (MULLER; SUREL, 2002, p. 8, 20-21)

No Brasil, uma definição possível de análise de políticas públicas pode ser encontrada em trabalho de Marta Arretche (2009, p. 30):

Qualquer política pública pode ser formulada e implementada de diversos modos [...] A análise de políticas públicas busca reconstruir estas diversas características, de forma a apreendê-las em um todo coerente e compreensível. Ou, melhor dizendo, para dar sentido e entendimento ao caráter errático da ação pública.

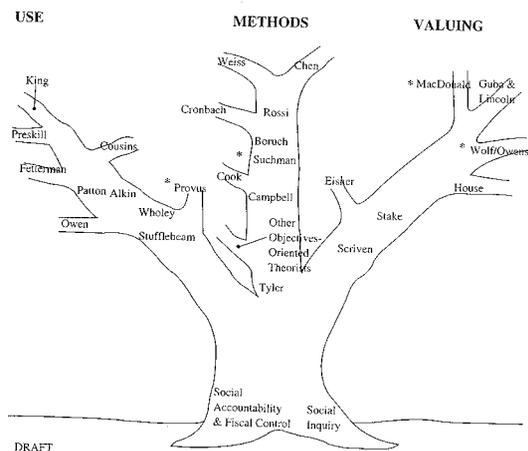
As colocações da autora, que permitem diferenciar o escopo da análise de políticas daquele da avaliação de políticas públicas, autorizam que se reafirmem considerações já realizadas alhures:

Quer seja sob a égide da avaliação ou da pesquisa avaliativa, o campo se desenvolveu vertiginosamente, e a busca pelo aprimoramento e rigor metodológico ainda não chegou ao fim. Várias são as abordagens teóricas e os movimentos em avaliação, e muitos são os autores que buscam categorizar as tendências, correntes ou gerações em avaliação. (BAUER, 2011, p. 10)

Para ilustrar essa diversidade, cabe mencionar o trabalho de Madaus, Scriven e Stufflebeam (1983), em que se identifica a evolução dos movimentos em avaliação em cinco períodos principais: o período pré-tyleriano (antes de 1930); a era tyleriana (1930-1945); a era da inocência (1946-1957); a idade do realismo (1958-1973); e a idade do profissionalismo, a partir de 1973. Em relação às abordagens avaliativas, Stufflebeam e Shinkfield (2007, p. 145) classificam e analisam 26 avaliativas, inclusive denominando algumas delas de pseudoavaliações, se elas “falham em produzir e reportar avaliações válidas de mérito e valor” para as audiências que estariam interessadas nesses resultados.

Outra sistematização das diferentes abordagens teóricas em avaliação de programas foi proposta por Alkin e Christie (2004). Os autores apresentam graficamente uma árvore genealógica da avaliação (Figura 1), ressaltando que a consolidação e a universalização do campo da avaliação apoiaram-se numa dupla base de *accountability*, no sentido de responsabilização e de investigação social, ambas situadas no tronco da árvore. A partir desse tronco comum, os autores organizam as contribuições dos teóricos em três ramificações que indicam ênfases ou preocupações de diferentes naturezas: os teóricos que se ocuparam com a discussão dos métodos de avaliação; os que focalizaram os usos dos resultados das informações obtidas por meio da avaliação; e os que discutem a questão do valor e do mérito na avaliação. O pensamento de alguns desses autores, em relação à questão valorativa, é explorado na próxima seção.

FIGURA 1 - Árvore teórica da avaliação revisitada



Fonte: Alkin e Christie (2004, p. 389).

CONTRIBUIÇÕES TEÓRICAS INTERNACIONAIS À QUESTÃO DA VALORAÇÃO

Michael Scriven é, sem dúvida, o primeiro teórico a expressar sua preocupação com a questão da valoração e da realização de julgamentos de valor, assumindo um papel importante na teorização sobre esses aspectos. Para ele, todo processo avaliativo deveria seguir uma lógica, definida em quatro etapas: selecionar critérios de mérito; definir padrões de desempenho; medir o desempenho; e sintetizar os resultados em uma declaração de valor. Na primeira etapa, defende o autor, deveriam ser consideradas as necessidades da sociedade.

A avaliação não é apenas a aplicação de métodos de ciências sociais para resolver problemas sociais; em vez disso, os avaliadores usam os métodos das ciências sociais para examinar o mérito, o valor e a importância de um programa ou política para fins de descrição de valores associados a diferentes grupos de partes interessadas, além de conclusões avaliativas sobre boas e más soluções para problemas sociais. (SCRIVEN, 2003, p. 21, tradução nossa)

Scriven critica fortemente a tendência que percebe de os avaliadores, principalmente em algumas abordagens, não assumirem a perspectiva

valorativa de seu trabalho. Seu incômodo é compreensível quando se observa que, para o autor, a avaliação é instrumento de garantia de direitos sociais, cabendo ao avaliador protegê-los. O avaliador é, pois, um “protetor público” na sua área (SCRIVEN, 1971, p. 53). Consequentemente, quem realiza a avaliação, ou pesquisa avaliativa, não pode se eximir da tarefa de julgar se um programa funciona ou não, se é bom ou ruim, se atinge ou não seus objetivos.

Além disso, Scriven sugere que o avaliador deve buscar uma síntese, um julgamento final, mesmo quando a avaliação de um programa, uma política ou ação apresente múltiplos resultados. Segundo Alkin e Christie (2004), ele é o único autor que sustenta e advoga a necessidade dessa síntese, em termos de julgamento de qualidade da entidade avaliada.

Mas, ainda que o avaliador assuma um papel importantíssimo na teorização de Scriven à medida que deve determinar critérios de mérito e valor e padrões de qualidade na análise dos resultados (SCRIVEN, 1976), o autor é cuidadoso ao ressaltar que a decisão sobre quais critérios e padrões escolher não é única e exclusivamente do avaliador. Cabe a ele discutir amplamente os valores com especialistas, apoiado em fatos empíricos, para determinar os valores que são fundamentais. Tal cuidado, para o autor, propiciaria validade às afirmações e considerações valorativas (SHADISH; COOK; LEVITON, 1991; ALKIN; CHRISTIE, 2004). Note-se que Scriven considera mais a opinião de especialistas sobre os valores que devem ser considerados na avaliação do que os envolvidos na proposição do programa.¹³

Já Elliot Eisner assume um posicionamento bastante diferente do de Scriven, ainda que ambos atribuam importância central à questão da valoração em suas teorizações.

Para Eisner (1998, p. 80 *apud* ALKIN; CHRISTIE, 2004, p. 34-36), a avaliação “se preocupa em fazer julgamentos de valor sobre a qualidade de algum objeto, situação ou processo”. O papel do avaliador, para o autor, é muito parecido com a do crítico de arte. Ele seleciona, a partir de seu conhecimento (*connoisseurship*), os aspectos centrais que determinam a qualidade do que é avaliado. A crítica, que juntamente com conhecimento constitui os aspectos centrais da teorização do autor, se daria pela tentativa de resposta a três questões essenciais: o que está acontecendo, qual o sentido do que está acontecendo e qual o valor do que está acontecendo. Sua proposta, pois, baseia-se em

¹³ Nos limites desse texto, o foco não é discutir aspectos positivos e negativos de cada abordagem. Para um aprofundamento no assunto, pode-se consultar Shadish, Cook e Leviton (1991) e Alkin e Christie (2004).

abordagens qualitativas, tendo a observação um papel crítico na emissão do juízo de valor (ALKIN; CHRISTIE, 2004; MATHISON, 2005).

Ernest House (2004) também dedica especial atenção ao componente de valor em suas considerações sobre avaliação. Para ele, os avaliadores devem se concentrar em tirar, a partir de fatos e valores, as conclusões avaliativas:

Os avaliadores devem concentrar-se em tirar conclusões, um processo que implica empregar métodos apropriados como um componente, mas também descobrir os critérios e padrões de comparação corretos e abordá-los na coleta de dados, análise e inferência de conclusões. (HOUSE, 2004, p. 219)

O autor defende que os princípios da teoria da justiça de Rawls¹⁴ deveriam ser levados em conta no processo avaliativo, embasando o estabelecimento de julgamentos que priorizem econômica e politicamente os interesses dos desfavorecidos. Partindo do princípio de que não há, nem deve haver, neutralidade de valor na avaliação, House (2004) aceita que o avaliador assuma um posicionamento favorável à diminuição das injustiças sociais. Isso não significa que os interessados na avaliação devam ser desconsiderados. Para ele, tanto a natureza da entidade sendo avaliada quanto as intuições dos envolvidos com o que está sendo avaliado devem ser respeitadas na definição dos critérios avaliativos e dos dados a serem coletados.

Com abordagem distinta dos autores anteriormente mencionados, Robert Stake (2004) acredita que não há um único, verdadeiro ou melhor valor a partir do qual um programa possa ser julgado, visto que os diversos programas podem ter diferentes valores, para pessoas diversas. Segundo Shadish, Cook e Leviton (1991), para Stake os valores dos interessados na avaliação devem ser considerados pelo avaliador ao fazer os julgamentos sobre os resultados, bem como estar neles refletidos. Assim, sua teorização admite uma diversidade de perspectivas entre os envolvidos no programa. Em suas palavras:

¹⁴ John Rawls, em sua obra *Uma teoria da justiça*, publicada originalmente em 1971, defende a teoria da justiça como equidade e estabelece dois princípios de justiça para as instituições, diferindo-os dos princípios aplicados aos indivíduos e suas ações, que considera fundamentais: o princípio da liberdade e o da igualdade de oportunidades. Não é objetivo deste trabalho o aprofundamento na discussão acerca da teoria rawlsiana, mas apenas contextualizar a discussão de House.

[...] direcionada para a descoberta de mérito e lacunas no programa, a avaliação responsiva reconhece múltiplas fontes de valoração, bem como múltiplos fundamentos. É respeitosa com múltiplos e até às vezes contraditórios padrões, mantidos por indivíduos e grupos diferentes, com uma relutância em pressionar por consenso. (STAKE, 2004, p. 210)

Wolf e Owens, na década de 1980, desenvolveram uma abordagem avaliativa que claramente seguia uma orientação valorativa. A preocupação de ambos era a de minimizar a possibilidade de *bias* (desvios) involuntários, realizados pelos avaliadores, quer na definição dos objetivos a serem avaliados, quer na escolha da base para a valoração dos resultados. Como em um julgamento, são chamados dois ou três grupos de avaliadores, que assumem posições contrapostas em relação ao programa (método de contraposição). A atenção que os autores deram à forma como o julgamento seria feito possibilitou sua classificação, por Alkin e Christie, nesse braço da avaliação. Considerando que um julgamento tem como objetivo final o interesse público, presume-se, a partir do exposto, que as discussões dos grupos oponentes acerca dos resultados obtidos – o levantamento de aspectos favoráveis e contrários ao programa – possibilitam o ajuste das bases valorativas sem, contudo, que o avaliador decida sobre quais valores deveriam ser considerados no âmbito do programa.

Já para Barry MacDonald (1995), os diversos interessados no programa possuem visões distintas acerca dos seus objetivos e das dimensões e níveis de qualidade do mesmo. Cabe ao avaliador, em sua perspectiva, ser um “negociador”. Sua atuação principal, no que se refere à valoração, seria “coletar e comunicar as definições alternativas, as perspectivas e julgamentos das pessoas de dentro e do entorno do programa”¹⁵ (MACDONALD *et al.*, 1979, p. 127-128 *apud* ALKIN; CHRISTIE, 2004, p. 39). Ainda que defenda que os especialistas devam conduzir avaliações democráticas, o autor reconhece que os contextos para avaliação são diferenciados e o avaliador pode se defrontar com a necessidade de conduzir propostas avaliativas burocráticas e autocráticas. Mesmo nesses contextos, o papel do avaliador seria o de representar o valor do outro e aceitá-lo. Isso não significa que o autor não reconheça que o especialista tem

15 Do original: “Rather he will collect and communicate the alternative definitions, perspectives and judgement held by people in and around the program”.

seus valores próprios; ao contrário, ele considera que o avaliador tem um papel político e que seus valores estão sempre presentes no processo avaliativo, desde a escolha do método (MACDONALD, 1995).

O Quadro 1 apresenta uma classificação das propostas avaliativas dos autores supracitados em relação ao tipo de teoria de valor que, acredita-se, está subjacente ao seu pensamento.

QUADRO 1 - Proposta de classificação dos autores de acordo com as teorias de valoração apresentadas por Shadish, Cook e Leviton (1991)

AUTORES	PROPOSTA AVALIATIVA ¹⁶	TIPO DE TEORIA DE VALOR
Scriven	<i>Goal-free evaluation</i>	Metateoria e teoria prescritiva
Guba e Lincoln	<i>Constructivist evaluation</i>	Descritiva
Stake	<i>Responsive evaluation</i>	Descritiva
House	<i>Deliberative democratic evaluation</i>	Prescritiva
Eisner	<i>Connoisseurship</i>	Descritiva
Wolf e Owens	<i>Adversary evaluation</i>	Descritiva
MacDonald	-	Descritiva

Fonte: Alkin e Christie (2004); Shadish, Cook e Leviton (1991); Mathison (2005).

Além dos autores destacados, podem ser encontradas iniciativas, tanto de associações de profissionais voltadas à avaliação quanto de organismos multilaterais, que indicam aspectos, valores e critérios que são desejáveis de ser considerados na prática avaliativa. Um exemplo do exposto pode ser encontrado no documento *Manual de planificación, seguimiento y evaluación de los resultados de desarrollo*, do Programa das Nações Unidas para o Desenvolvimento (PNUD, 2009). Entre os diversos elementos discutidos, observa-se uma clara posição acerca de critérios usualmente incluídos nas avaliações do PNUD: pertinência, eficácia, eficiência, sustentabilidade e impacto dos esforços de desenvolvimento, cuja importância é justificada no documento. Explicação sintética acerca desses critérios pode ser encontrada em Jannuzzi (2018, p. 16):

De forma simplificada, Relevância de um programa ou projeto está associado ao grau de pertinência do mesmo às demandas públicas prioritárias, isto é, sua aderência à agenda de prioridades políticas de uma

¹⁶ As abordagens avaliativas propostas por cada um dos autores foram mantidas com sua denominação original para evitar que a tradução dos termos modificasse os sentidos atribuídos pelos teóricos às suas propostas.

determinada sociedade. Eficácia de um programa ou projeto é um atributo relacionado ao grau de cumprimento dos objetivos do mesmo ou de atendimento da demanda motivadora do programa. Eficiência, por outro lado, está associado à qualidade e preocupação de como os recursos são utilizados na produção dos resultados do programa. A Sustentabilidade refere-se à capacidade do programa ou projeto em gerar mudanças permanentes na realidade em que atuaram. Por fim, o Impacto diz respeito aos efeitos de médio e longo prazo sobre os beneficiários e sociedade – positivos ou não – direta ou indiretamente atribuíveis ao programa ou projeto. Em geral, quando considerados os impactos positivos, emprega-se alternativamente o termo Efetividade.

Além de sugerir com clareza uma base valorativa pragmática para os programas, o documento parece consoante com a perspectiva prescritiva de valoração, ao defender que outros princípios, como o respeito aos direitos humanos e à equidade, traduzida no documento pela preocupação com a questão de gênero, devem ser considerados nas avaliações do PNUD, como ilustra o trecho a seguir:

En línea con los esfuerzos de desarrollo del PNUD, las evaluaciones de la organización están guiadas por los principios de igualdad de género, un enfoque basado en los derechos y el desarrollo humano. Por tanto, y según corresponda, las evaluaciones del PNUD valorarán el grado en el que las iniciativas de la organización han tratado los temas de inclusión social y de género, de igualdad y de empoderamiento; han contribuido a fortalecer la aplicación de estos principios a los diferentes esfuerzos de desarrollo en un país dado; y han incorporado en el diseño de la iniciativa el compromiso del PNUD con los enfoques basados en los derechos y la problemática de género. (PNUD, 2009, p. 169-170, grifos nossos)

Esses são posicionamentos de alguns dos teóricos do campo e instituições que se debruçaram sobre a temática. Percebe-se, a partir do exposto, que a questão da valoração ganha atenção em diferentes momentos da história da avaliação e, na grande parte das vezes, emerge da prática avaliativa dos autores, que os confronta com dilemas éticos e com a necessidade de se posicionar perante o desafio de estabelecer as bases para o julgamento dos programas. A seção seguinte busca explorar como tais questões são trabalhadas por autores e avaliadores nacionais. Para isso, fez-se uma revisão bibliográfica não exaustiva, em teses, dissertações, artigos e relatórios avaliativos, que tratam de programas e políticas educacionais, apresentada na próxima seção.

A DISCUSSÃO DE VALORAÇÃO EM AVALIAÇÕES DE PROGRAMAS NO TERRITÓRIO NACIONAL

No Brasil, raros são os trabalhos que se debruçam sobre a questão valorativa inerente às avaliações de programas, políticas e projetos educacionais. Em geral, os pesquisadores e avaliadores não assumem ou propõem claramente bases a partir das quais serão realizados os julgamentos sobre as propostas avaliadas, exceção feita ao pensamento de Paulo Jannuzzi (2016) e de Heraldo Vianna (2005).

Ao discutir o caráter coletivo do ato de avaliar programas, Vianna (2005) reconhece a diversidade de interesses, visões e valores que lhe são inerentes e defende, com base em House e Howe (2003), que esses aspectos sejam debatidos democraticamente para que se chegue a julgamentos sensatos e adequados. Para o autor, devem ser considerados valores efetivamente importantes para o programa, ainda que a busca desses padrões valorativos, a partir dos quais serão feitos juízos acerca do programa e de sua qualidade, seja tarefa complexa, que nem sempre resulta no estabelecimento de consensos. Referindo-se especificamente à avaliação de programas escolares, Vianna (2005, p. 51) ressalta que “devem ser levados em conta valores, crenças e múltiplas exigências sociais em relação aos vários tipos de aprendizagem e às diferentes situações da escola”. Infere-se, portanto, que sua referência para o estabelecimento das bases de julgamento de valor e mérito do programa são os diversos envolvidos em sua consecução.

Paulo Jannuzzi (2016) parte do princípio de que, às avaliações de programas propostas no território nacional, subjazem três concepções políticas e ideológicas em disputa: eficiência econômica, eficácia procedural e efetividade social, que influenciam escolhas teórico-metodológicas e os valores

públicos com os quais dialogam aqueles que se dedicam à tarefa de avaliar políticas públicas. Tais valores, que o autor identifica como “qualidade do gasto público”, “conformidade processual” e “aprimoramento em busca do maior impacto social”, atrelam-se a diferentes concepções de justiça e oferecem diferentes bases para a proposição de critérios de valor e mérito do que está sendo avaliado. O autor assume a defesa de que a efetividade social e, portanto, o aprimoramento em busca do maior impacto social devem ser os elementos balizadores dos avaliadores de programas, já que estão intimamente relacionados a princípios de justiça, equidade e bem-estar social que são basilares na Constituição Federal de 1988. Jannuzzi explora, com clareza dificilmente encontrada na literatura brasileira sobre avaliação de programas, como valores públicos republicanos influenciam (ou deveriam influenciar) os critérios de mérito e valor utilizados no julgamento dos programas e projetos, que coadunam com as teorias prescritivas de valor, exploradas anteriormente.

O interesse pela compreensão de como os autores e pesquisadores em avaliação brasileiros têm lidado com problemas relacionados à questão da valoração é antigo e remonta à época em que a autora deste artigo estava produzindo a tese de doutoramento (BAUER, 2011). Naquele momento, o levantamento bibliográfico realizado mostrou que o Brasil estava apenas engatinhando no que se refere às avaliações de programas e de seus impactos, que era o interesse principal da pesquisa. Assim, pelos caminhos teórico-metodológicos escolhidos, não se realizou uma discussão sobre como a questão da valoração estava posta na literatura nacional.

Nos limites do presente texto, busca-se preencher essa lacuna. Ainda que a experiência e as decisões da autora sobre a valoração sejam retomadas e relatadas, outros trabalhos são considerados. Para a sua identificação e seleção, escolheu-se fazer uma busca ampla, utilizando a ferramenta Google Acadêmico. A opção justifica-se pela intenção de identificar, em um curto período, trabalhos de diferente natureza, que normalmente são apontados por essa ferramenta.

A seleção foi realizada por meio do uso de descritores que combinavam diferentes expressões: avaliação de programas educacionais AND mérito; avaliação de programas educacionais AND valor; e avaliação de programas educacionais AND julgamento. Foram obtidas diversas referências que dialogavam com outras áreas (avaliação de programas de saúde, de informática, etc.), ou com outras dimensões da avaliação (avaliação da aprendizagem, institucional, do ensino superior, etc.). Dada a enorme quantidade de

referências, selecionaram-se somente aquelas que se relacionavam diretamente, pela leitura dos títulos, à área educacional e à avaliação de programas. O segundo passo foi consolidar as referências em um arquivo único, eliminando as repetições, visto que muitas vezes foram obtidos os mesmos trabalhos, ainda que tivessem sido utilizados descritores diferentes. Essa primeira fase de seleção resultou em um banco de 85 referências, com artigos, relatórios avaliativos, teses e dissertações.

No segundo momento, cada uma das referências foi localizada na internet. Quando não foi possível acessar o trabalho eletronicamente, ele foi descartado. A leitura dos resumos possibilitou uma nova seleção, restando uma base com 32 estudos. Nestes, procurou-se identificar se os autores faziam uma discussão própria sobre a questão da valoração (mais explícita ou implicitamente), se ela aparecia reproduzida pela revisão da literatura realizada, sem representar necessariamente o pensamento do autor, ou, ainda, se estava ausente.

Também se ensaiou uma classificação dos trabalhos em prescritivos ou descritivos. Foram considerados autores que procuraram parâmetros de valoração fora do programa ou ação avaliada, quer na base da literatura de referência, quer por meio da indicação de padrões avaliativos acordados no âmbito de associações de especialistas, como o *Joint Committee*, por exemplo. Os que basearam a análise dos resultados na opinião dos interessados ou nas indicações presentes na documentação do próprio programa foram entendidos como assumindo uma teoria descritiva. Finalmente, quando não ficou clara a base de julgamento do autor, ou o trabalho não comportava julgamentos valorativos, indicou-se “ausente” ou “não se aplica”.

QUADRO 2 – Análise de trabalhos de autores brasileiros com base nas teorias de valoração apresentadas por Shadish, Cook e Leviton (1991)

OBRAS	DISCUSSÃO SOBRE VALORAÇÃO	TIPO DE TEORIA DE VALOR
ABRAMOWICZ, Mere. Avaliação, tomada de decisões e políticas: subsídios para um repensar. <i>Estudos em Avaliação Educacional</i> , São Paulo, n. 10, p. 81-101, jul./dez. 1994.	Explícito	Prescritiva
ANDRIOLA, Wagner B. Utilização do modelo CIPP na avaliação de programas sociais: o caso do Projeto Educando para a Liberdade da Secad/MEC. <i>REICE: Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación</i> , Madrid, v. 8, n. 4, p. 65-82, 2010.	Ausente	Não se aplica
BAUER, Adriana. <i>Avaliação de impacto de formação docente e serviço: o programa Letra e Vida</i> . Tese (Doutorado) – Universidade de São Paulo, São Paulo, 2011.	Explícito	Prescritiva
BITTENCOURT, Jaqueline M. V. <i>Uma avaliação da efetividade do Programa de Alimentação Escolar no Município de Guaíba</i> . Dissertação (Mestrado) – UFRGS, Porto Alegre, 2007.	Implícita	Descritiva
BRANDÃO, Maria de F. R. Um modelo de avaliação de programa de inclusão digital e social. Tese (Doutorado) – UnB, Brasília, 2009.	Explícita	Descritiva
CAMPOS, Maria de F. H. et al. Avaliação de políticas e programas governamentais: experiências no mestrado profissional. <i>Avaliação: Revista Avaliação de Políticas Públicas</i> , v. 1, n. 1, p. 49-58, jan./jun. 2008.	Revisão de literatura	Não se aplica
CARNEIRO, Liliâne B. <i>Características e avaliação de programas brasileiros de atendimento educacional ao superdotado</i> . 2015. 178 f. Tese (doutorado) – UnB, Brasília, 2015.	Implícita	Descritiva
CORREA, Vera L. de A. <i>Avaliação de programas educacionais: a experiência das escolas cooperativas em Maringá</i> . Tese (Doutorado) – FGV, Rio de Janeiro, 1993.	Revisão de literatura	Descritiva
DAVOK, Delsi F. et al. <i>Modelo de meta-avaliação de processos de avaliação da qualidade de cursos de graduação</i> . Tese (Doutorado) – UFSC, Florianópolis, 2006.	Explícita	Prescritiva
DAVOK, Delsi F. Quality in education. <i>Avaliação: Revista da Avaliação da Educação Superior</i> , Campinas, v. 12, n. 3, p. 505-513, set. 2007.	Explícita	Não se aplica
DUTRA, Marina L. da S. <i>Avaliação de treinamento: em busca de um modelo efetivo</i> . Tese (Doutorado) – FGV, Rio de Janeiro, 1979.	Implícita	Prescritiva
FERREIRA, Helder; CASSIOLATO, Martha; GONZALEZ, Roberto. <i>Uma experiência de desenvolvimento metodológico para avaliação de programas: o modelo lógico do programa segundo tempo</i> . Brasília: Ipea, 2009.	Implícita	Descritiva
FIRME, Thereza P. Os avanços da avaliação no século XXI. <i>Revista Educação Geográfica em Foco</i> , v. 1, n. 1, p. 1-4, jan./jul. 2017.	Explícita	Descritiva
GIMENES, Nelson A. S. Estudo metavaliativo do processo de auto-avaliação em uma Instituição de Educação Superior no Brasil. <i>Estudos em Avaliação Educacional</i> , São Paulo, v. 18, n. 37, p. 217-243, maio/ago. 2007.	Implícita	Prescritiva
GOLDBERG, Maria A. A.; BARRETTO, Elba S. S.; MENEZES, Sônia M. C. Avaliação educacional e educação de adultos. <i>Cadernos de Pesquisa</i> , São Paulo, n. 8, p. 7-110, 1973.	Explícita	Prescritiva
MEDEIROS NETO, Benedito. <i>Avaliação dos impactos dos processos de inclusão digital e informacional nos usuários de programas e projetos no Brasil</i> . Tese (Doutorado) – UnB, Brasília, 2012.	Ausente	Descritiva
MIORANZA, Claudio. <i>Desenvolvimento e aplicação de modelo multidimensional para a avaliação da qualidade educacional no programa de Pós-Graduação Stricto Sensu do IPEN</i> . Tese (Doutorado) – Ipen/ USP, São Paulo, 2009.	Explícita	Prescritiva

(Continua)

(Continuação)

OBRAS	DISCUSSÃO SOBRE VALORAÇÃO	TIPO DE TEORIA DE VALOR
ORLANDO FILHO, Ovidio; SÁ, Virgínio, I. M. Avaliação externa da gestão escolar do Programa Nova Escola do Estado do Rio de Janeiro: um estudo reflexivo sobre o seu primeiro ciclo de realização (2000-2003), passados quinze anos de sua implementação. <i>Ensaio: Avaliação e Políticas Públicas em Educação</i> , v. 24, n. 91, p. 275-307, abr./jun. 2016.	Explícita	Descritiva
PINTO, Edvan W. F. <i>Programa Nacional de Integração da Educação Profissional com a Educação Básica na Modalidade de Educação de Jovens e Adultos (PROEJA): uma avaliação de impactos nas condições de trabalho e renda dos egressos no município de Açailândia-MA</i> . Dissertação (Mestrado) – UFMA, São Luís, 2016.	Ausente	Ausente
RIBEIRO, Claudia. <i>Programa Alfabetizar com Sucesso – Programa de acompanhamento dos anos iniciais da rede pública de Pernambuco: a avaliação do município de Condado</i> . Dissertação (Mestrado) – UFJF, Juiz de Fora, 2015.	Ausente	Ausente
ROCHA, Ana R.; CAMPOS, Gilda H. B. Avaliação da qualidade de software educacional. <i>Em Aberto</i> , Brasília, v. 12, n. 57, p. 32-44, 1993.	Ausente	Prescritiva
RÓDRIGUES, Rosângela S. <i>Modelo de avaliação para cursos no ensino a distância: estrutura, aplicação e avaliação</i> . Dissertação (Mestrado) – UFSC, Florianópolis, 1998.	Revisão de literatura	Prescritiva
ROSEMBERG, Fúlvia. Políticas de educação infantil e avaliação. <i>Cadernos de Pesquisa</i> , São Paulo, v. 43, n. 148, p. 44-75, jan./abr. 2013.	Explícita	Prescritiva
SERPA, Selma M. H. C. <i>Para que avaliar? Identificando a tipologia, os propósitos e a utilização das avaliações de programas governamentais no Brasil</i> . Dissertação (Mestrado) – UnB, Brasília, 2010.	Revisão de literatura	Não se aplica
SILVA, Danilma de M. <i>Desvelando o PRONATEC: uma avaliação política do programa</i> . Dissertação (Mestrado) – Universidade Federal do Rio Grande do Norte, Natal, 2015.	Revisão de literatura	Ausente
SOMERA, Elizabeth A. S. Reflexões sobre vertentes da avaliação educacional. <i>Acesso do Avesso</i> , Araçatuba, SP, v. 6, n. 6, p. 56-68, ago. 2008.	Revisão de literatura	Não se aplica
SOUSA, Clarilza P. Descrição de uma trajetória na/da avaliação educacional. <i>Ideias</i> , São Paulo, v. 30, p. 161-174, 1998.	Revisão de literatura	Não se aplica
VIANNA, Heraldo M. A prática da avaliação educacional: algumas colocações metodológicas. <i>Cadernos de Pesquisa</i> , São Paulo, n. 69, p. 40-47, maio 1989.	Revisão de literatura	Não se aplica
VIANNA, Heraldo M. Avaliação de Programas Educacionais: duas questões. <i>Estudos em Avaliação Educacional</i> , São Paulo, v. 16, n. 32, p. 43-56, jul./dez. 2005.	Explícita	Descritiva
VIANNA, Heraldo M. Avaliação educacional: problemas gerais e formação do avaliador. <i>Estudos em Avaliação Educacional</i> , São Paulo, v. 25, n. 60 (especial), p. 74-84, dez. 2014.	Revisão bibliográfica	Não se aplica
VIANNA, Heraldo M. Fundamentos de um programa de avaliação educacional. <i>Estudos em Avaliação Educacional</i> , São Paulo, n. 28, p. 23-38, jul./dez. 2003.	Revisão bibliográfica	Não se aplica
VIANNA, Heraldo M. Novos estudos em avaliação educacional. <i>Estudos em Avaliação Educacional</i> , São Paulo, n. 19, p. 77-169, jan./jun. 1999.	Revisão bibliográfica	Não se aplica

Fonte: Google Acadêmico e SciELO (*Scientific Electronic Library Online*). Elaboração da autora.

A análise do quadro revela alguns aspectos de interesse. Diversos são os trabalhos que se assumem como avaliativos e que não realizam explicitamente

uma discussão teórica sobre a questão do valor. A maioria deles limita-se a fazer uma revisão de literatura, apontando conceitos e problemáticas relacionados ao tema da valoração, mas sem posicionamento explícito dos seus autores sobre a questão.

Cerca de um terço dos trabalhos apresenta, implícita ou explicitamente, uma teoria descritiva da avaliação, representando uma realidade bastante parecida com a que os teóricos estadunidenses apontam em suas análises. Em geral são avaliações que assumem abordagens qualitativas de avaliação, com ênfase em estudos de caso e pesquisas de opinião. Deve-se compreender esse resultado em relação à tradição teórica da avaliação brasileira que, principalmente nos anos 1980, tem ressaltado a dimensão política da avaliação e a existência de diversos valores entre os interessados nos processos avaliativos. De todo modo, mesmo essa produção teórica, que influenciou as propostas avaliativas no final do século XX, não trata diretamente (ou com maior profundidade) das questões valorativas. Não foi possível identificar, nesse sentido, uma metateoria brasileira em torno da questão da valoração. O autor que dedica mais atenção a esse aspecto parece ser Heraldo Vianna, ainda que suas considerações estejam dispersas em diversos dos artigos que escreveu ao longo de sua carreira.

Finalmente, e a título de ilustração das decisões que devem ser tomadas pelo interessado em realizar uma avaliação de programas, explicitam-se, com base em experiência prévia da autora (BAUER, 2011), as decisões que foram tomadas na oportunidade de realização de uma avaliação de impacto de um programa para formação de professores alfabetizadores e que, certamente, influenciaram as análises e os resultados obtidos. Não é objetivo deste artigo aprofundar-se no modelo metodológico proposto,¹⁷ mas apenas colocar em discussão as opções que foram realizadas na época do estudo.

Para a análise dos impactos do referido programa, foi necessário identificar as variáveis que seriam consideradas resultados desejados do programa. Na época de sua realização, alegava-se que a alfabetização deveria ser um processo para entender o mundo e interagir com ele, tornando os alunos capazes de acessar e fazer uso de diferentes habilidades linguísticas para resolver problemas reais. Essa concepção trazia necessidades de um processo mais amplo e holístico de alfabetização e, portanto, os critérios de mérito para avaliar o Programa foram definidos levando em conta a capacidade de os

¹⁷ Para mais informações sobre este tema, consultar Bauer (2010, 2011).

alunos demonstrarem tais habilidades de linguagem, ainda que tal objetivo não fosse explicitado. Além disso, tendo em vista que o programa focalizava na aprendizagem dos professores, pensou-se ser necessário avaliar o que eles aprenderam depois de participar do programa, considerando as novas teorias de alfabetização e metodologia de ensino que destacam o desenho do programa. Na produção dos instrumentos de pesquisa e de observação do ambiente das salas de aula, buscou-se contemplar práticas esperadas de professores participantes do curso, segundo não apenas os materiais do programa, mas também os indicadores de qualidade na educação, com ênfase em alfabetização, produzidos pela Ação Educativa, em parceria com o MEC e o Fundo das Nações Unidas para a Infância (Unicef). Essas decisões não se apoiaram em pressupostos metodológicos, mas referiram-se à definição dos objetivos do programa a serem avaliados e aos critérios a partir dos quais seu alcance seria julgado. Proposições de Daniel Stufflebeam (2002, p. 1, tradução nossa) foram consideradas:

Como os sistemas de avaliação são dependentes do contexto, eles devem considerar as necessidades, desejos e expectativas dos constituintes, além de outras variáveis, como valores sociais pertinentes, clientela e costumes; leis e estatutos relevantes; dinâmica econômica; forças políticas; interesses de mídia; critérios substantivos pertinentes; missão organizacional, objetivos e prioridades; governança organizacional, gestão, protocolos e rotinas operacionais; e a história da organização e desafios atuais.¹⁸

Nesse sentido, pode-se dizer que os critérios utilizados na análise, entendidos como “estândares nos quais os julgamentos são baseados” (STUFFLEBEAM, 2001, p. 1), foram formulados a partir de valores e objetivos intrínsecos ao programa, mas também incorporaram “valores sociais pertinentes” ou “critérios substantivos pertinentes”. Estes últimos foram incorporados, por exemplo,

¹⁸ Do original: “Since evaluation systems are context dependent, they must take into account constituents’ needs, wants, and expectations plus other variables such as pertinent societal values, customs, and mores; relevant laws and statutes; economic dynamics; political forces; media interests; pertinent substantive criteria; organizational mission, goals, and priorities; organizational governance, management, protocols, and operating routines; and the organization’s history and current challenges”. (Disponível em: <http://www.wmich.edu/evalctr/checklists/institutionalizingeval.pdf>. Acesso em: 01 maio 2011)

com a decisão de usar os indicadores de qualidade na educação como indício da efetividade do programa e com a investigação de sua influência na melhora dos resultados educacionais dos alunos e dos níveis de alfabetização. Assim, ampliou-se o olhar sobre os critérios a serem avaliados, incorporando objetivos de interesse social (todos os alunos alfabetizados), independentemente da clareza com que esse objetivo era perseguido pelo programa. O estabelecimento dos critérios e indicadores de análise dos dados se deu com base no entendimento da pesquisadora sobre os pressupostos teórico-metodológicos do programa. Essas decisões foram discutidas na tese. Se a base para o julgamento dos resultados obtidos tivesse sido os valores dos interessados – quer propositores, quer professores que participaram do programa –, as conclusões do trabalho teriam sido diferentes.

CONSIDERAÇÕES FINAIS

Todo programa, projeto ou ação faz parte de uma política pública mais ampla, constituindo a tradução para a realidade das estratégias formuladas para o alcance dos objetivos e metas a que se quer chegar. Nesse sentido, a política pública traduz-se como uma ação intencional do Estado junto à sociedade visando, espera-se, ao bem comum.

A ação política apresenta, em seu bojo, as concepções dos representantes do Estado sobre princípios que norteiam os aspectos fundamentais da área sobre a qual a política está sendo formulada. E o fazer político pressupõe o estabelecimento de metas para a área em questão. Ao determinar princípios e metas, a política assume capacidade projetiva, direcionando por um período limitado os caminhos trilhados pela sociedade. Tais percursos refletem a concepção de homem, de mundo, de sociedade, de Estado, de justiça social, de direitos humanos (dentre tantos outros) dos governantes que instituem a política.

Nem sempre, entretanto, tais concepções são explícitas nos documentos que organizam a política, sendo que esta, muitas vezes, tem de ser inferida por meio dos atos traçados para sua execução, expressos no processo de implementação dos programas e projetos que a concretizam, como nas entrelinhas das leis, regulamentos e os documentos elaborados para sua execução e outros elementos indicativos da política (DRAIBE, 2001). Ou, ainda, pelo comportamento e concepções daqueles que são responsáveis pela formulação e concretização da política. Portanto, não é possível afirmar que um direcionamento político é neutro ou livre da subjetividade humana. Enquanto efetivação de uma determinada política em ações (em plano teórico, ou seja, de concepção),

os programas educacionais carregam pressupostos e intencionalidades que precisam ser analisados criticamente. Jannuzzi (2018) inclusive avança nessa análise, postulando a necessidade de os valores públicos republicanos estarem na base tanto da elaboração de programas como de sua avaliação.

Entender esses aspectos é fundamental, especialmente quando o objetivo do pesquisador é contribuir para a avaliação da ação pública, das concepções que a fundamentam e de sua aplicabilidade. O avaliador pode apreciar os objetivos, ações e estratégias delineados na proposta, deliberando acerca de sua adequação diante dos desafios e necessidades da sociedade, ou pode assumir a perspectiva dos vários interessados nessa política. Sua escolha não requer reflexão e investigação cuidadosa, pois determinará, a miúdo, a base para análise da qualidade do programa, tarefa à qual nenhum avaliador de políticas públicas deveria se eximir.

É a discussão sobre essas opções e sua explicitação no fazer avaliativo, quer no âmbito da avaliação de programas, quer no da pesquisa avaliativa, que possibilita o desenvolvimento de boa teoria sobre avaliação de programas. Assim, caberia ao avaliador (ou o pesquisador em avaliação) atentar para os aspectos valorativos, buscando discutir seus pressupostos e especificar seu próprio sistema de valores e crenças acerca da prática avaliativa, contribuindo para a ampliação do debate acerca desse importante aspecto, inextricável à prática avaliativa.

REFERÊNCIAS

ALKIN, Marvin; CHRISTIE, Cristina A. An evaluation theory tree. In: ALKIN, Marvin C. (ed.). *Evaluation roots: tracing theorists' views and influences*. Thousand Oaks: Sage Publications, 2004. p. 12-65.

ARRETCHE, Marta. Tendências no estudo sobre avaliação. In: RICO, E. M. (org.). *Avaliação de políticas sociais: uma questão em debate*. São Paulo: Cortez/IEE, 2009. p. 29-40.

BAUER, Adriana. Avaliação de impacto no Brasil: é possível mensurar impactos de programas de formação docente? *Estudos em Avaliação Educacional*, São Paulo, v. 21, n. 46, p. 229-252, maio/ago. 2010.

BAUER, Adriana. *Avaliação de impacto de formação docente e serviço: o programa Letra e Vida*. 2011. 232 f. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade de São Paulo, São Paulo, 2011. doi:10.11606/T.48.2011.tde-15072011-105940. Acesso em: 12 out. 2018.

CONTANDRIOPOULOS, André-Pierre; CHAMPAGNE, François; DENIS, Jean-Louis; PINEAULT, Raynald. A avaliação na área da saúde: conceitos e métodos. In: HARTZ, Zulmira M. de Araújo (org.). *Avaliação em saúde: dos modelos conceituais à prática na análise da implantação de programas*. Rio de Janeiro: Fiocruz, 1997. p. 29-47.

COUSINS, Bradley; SHULHA, Lyn M. Complexities in setting program standards in collaborative evaluation. In: SMITH, Nick L.; BRANDON, Paulo R. *Fundamental issues in evaluation*. New York: The Guilford Press, 2008. p. 139-158.

DAVIDSON, Jane. Merit. In: MATHISON, Sandra. *Encyclopedia of evaluation*. Thousand Oaks, CA: Sage Publications, 2005. p. 245.

DRAIBE, Sônia M. Avaliação de implementação: esboço de uma metodologia de trabalho em políticas públicas. In: BARREIRA, Maria Cecília R. Nobre; CARVALHO, Maria do Carmo B. (org.). *Tendências e perspectivas na avaliação de políticas e programas sociais*. São Paulo: IEE/PUC-SP, 2001. p. 13-42.

FIGUEIREDO, Marcus; FIGUEIREDO, Argelina. Avaliação política e avaliação de políticas: um quadro de referência teórica. *Análise e Conjunturas*, Belo Horizonte, v. 1, n. 3, p. 107-127, set./dez. 1986.

GUBA, Egon G.; LINCOLN, Yvonna S. The distinction between merit and worth in evaluation. *Educational Evaluation and Policy Analysis*, v. 2, n. 4, p. 61-71, 1980.

HOUSE, Ernest R.; HOWE, Kenneth R. Deliberative democratic evaluation. In: KELLAGHAN, T.; STUFFLEBEAM, D. L. (ed.). *International handbook of educational evaluation*. Boston: Kluwer Academic, 2003. v. 2, p. 79.

HOUSE, Ernest R. Intellectual history in evaluation. In: ALKIN, Marvin C. (ed.). *Evaluation roots: tracing theorists' views and influences*. Thousand Oaks: Sage Publications, 2004. p. 218-224.

JANNUZZI, Paulo M. Eficiência econômica, eficácia procedural ou efetividade social: três valores em disputa na avaliação de políticas e programas sociais. *Desenvolvimento em Debate*, Rio de Janeiro, v. 4, n. 1, p. 117-142, 2016.

JANNUZZI, Paulo M. A implementação como centro da avaliação de políticas públicas: reflexões à luz da experiência da SAGI/MDS 2004-maio 2016. In: SEMINÁRIO BRASILEIRO SOBRE IMPLEMENTAÇÃO DE POLÍTICAS PÚBLICAS, 1., Brasília, 20-21 de março de 2018. *Anais [...]*. Brasília: Enap, 2018.

MACDONALD, Barry. La evaluación como servicio público: perspectivas de futuro. In: SÁEZ BREZMES, María José (ed.). *Conceptualizando la evaluación en España*. Madrid: Universidad de Alcalá de Henares, 1995. (Colección Pautas, 3). Disponível em: <https://ueaeprints.uea.ac.uk/31854/>. Acesso em: 10 jul. 2018.

MACDONALD, Barry; JENKINS, David; KEMMIS, Stephen; VERMA, Gajendra. Racism and education evaluation. In: VERMA, Gajendra; BUGLEY, Christopher (ed.). *Racism and educational evaluation education and identity*. London: Macmillan, 1979.

MADAUS, George; SCRIVEN, Michael; STUFFLEBEAM, Daniel L. Program evaluation: an historical overview. In: MADAUS, George; SCRIVEN, Michael; STUFFLEBEAM, Daniel L. (ed.). *Evaluation models: viewpoints on Educational and Human Services Evaluation*. Boston: Kluwer-Nijhoff, 1983. p. 3-22.

MATHISON, Sandra. *Encyclopedia of evaluation*. Thousand Oaks, CA: Sage Publications, 2005.

MULLER, Pierre; SUREL, Yves. *A análise das políticas públicas*. Tradução: Agemir Bavaresco, Alceu R. Ferraro. Pelotas: Educat, 2002.

PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO. Garantía de calidad del diseño y la metodología de la evaluación. *Manual de planificación, seguimiento y evaluación de los resultados de desarrollo*. New York: PNUD, 2009. p. 161-178.

RAWLS, John. *Uma teoria da justiça*. São Paulo: Martins Fontes, 1997.

SCHWANDT, Thomas. Values. In: MATHISON, Sandra. *Encyclopedia of evaluation*. Thousand Oaks, CA: Sage Publications, 2005. p. 443-444.

SCRIVEN, Michael. Evaluating educational programs. In: CARO, Francis G. (ed.). *Readings in evaluation research*. New York: Russell Sage Foundation, 1971. p. 49-53.

SCRIVEN, Michael. Payoffs from evaluation, In: ABT, Clark C. *The evaluation of social programs*. Beverly Hill: Sage Publications, 1976. p. 217-224.

SCRIVEN, Michael. New frontiers of evaluation. *Evaluation Practice*. Science Direct, v. 7, n. 1, p. 7-44, Feb. 1986.

SCRIVEN, Michael. *Evaluation thesaurus*. Newbury Park, CA: Sage Publications, 1991.

SCRIVEN, Michael. Hard-won lessons in program evaluation. *New directions for program evaluation*. San Francisco, CA: AEA/Sage Publications, 1993.

SCRIVEN, Michael. Evaluation in the new millennium: the transdisciplinary vision. In: DONALDSON, Stewart I.; SCRIVEN, Michael (ed.). *Evaluating social programs and problems*. Mahwah NY: Erlbaum, 2003. p. 19-41.

SHADISH, William; COOK, Thomas; LEVITON, Laura. *Foundations of program evaluation: theories of practice*. Newbury Park, CA: Sage Publications, 1991.

STAKE, Robert. Stake and responsive evaluation. In: ALKIN, Marvin C. (ed.). *Evaluation roots: tracing theorists' views and influences*. Thousand Oaks: Sage Publications, 2004. p. 203-217.

STUFFLEBEAM, Daniel L. Evaluation values and criteria checklist. *Evaluation Checklists Project*. Kalamazoo, MI: The Evaluation Center, March 2001.

STUFFLEBEAM, Daniel L. Institutionalizing evaluation checklist. *Evaluation Checklists Project*. Kalamazoo, MI: The Evaluation Center, 2002. Disponível em: http://www.wmich.edu/evalctr/archive_checklists/institutionalizingeval.pdf. Acesso em: 23 fev. 2011.

STUFFLEBEAM, Daniel L.; SHINKFIELD, Antony J. *Evaluation theory, models, and applications*. San Francisco, CA: Jossey-Bass, 2007.

SUCHMAN, Edward A. Evaluating educational programs. In: CARO, Francis (ed.). *Readings in evaluation research*. New York: Russell Sage Foundation, 1971. p. 43-48.

VIANNA, Heraldo M. Avaliação de programas educacionais: duas questões. *Estudos em Avaliação Educacional*, São Paulo, v. 16, n. 32, p. 43-56, jul./dez. 2005.

WORTHEN, Blane; SANDERS, James; FITZPATRICK, Jody. *Avaliação de programas: concepções e práticas*. São Paulo: Edusp/Gente, 2004.

Recebido em: 15 JULHO 2018

Aprovado para publicação em: 22 OUTUBRO 2018



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.