

✓ Por um olhar epistemológico para a avaliação em políticas públicas: história, teoria e método

Rosana de Freitas Boullosa¹

¹ Universidade de Brasília (UnB) /
University of Brasília (UnB)
Departamento de Gestão de Políticas
Públicas/Department of Public Policy
Management

Resumo: Compreendendo a avaliação em políticas públicas como um campo estudos e práticas já denso e plural, este artigo busca problematizar seus desenvolvimentos históricos e sua atual configuração a partir de um olhar epistemológico. Este olhar buscou, de um lado, privilegiar a forte relação entre as principais expressões teóricas e metodológicas de tal campo; e, de outro, revelar e interpretar as disputas valorativas mais substanciais que modelam suas descontinuidades e contradições. Do ponto de vista do método, esta pesquisa foi desenvolvida como de tipo estado da arte, dando uma particular atenção à literatura que se debruçou sobre propostas classificatórias em avaliação de políticas públicas. Estruturado em seis seções, este artigo oferece ainda um caminho próprio para uma classificação, por epistemologias, da atual configuração do campo.

Palavras-chave: avaliação; políticas públicas, epistemologia.

EN For an epistemological look at public policy evaluation: history, theory and method

Abstract: Understanding the evaluation in public policies as a field of studies and practices already dense and plural, this article seeks to problematize its historical developments and its current configuration from an epistemological perspective. This look sought, on the one hand, to privilege the strong relationship between the main theoretical and methodological expressions of such field; and, on the other, to reveal and interpret the most substantial valuation disputes that shape its discontinuities and contradictions. From the point of view of the method, this research was developed as a state-of-the-art type, giving particular attention to the literature that focused on classificatory proposals in public policy evaluation. Structured in six sections, this article also offers its own path for a classification, by epistemologies, of the current configuration of the field.

Key-words: evaluation, public policy, epistemology

ES Para una mirada epistemológica a la evaluación de políticas públicas: historia, teoría y método

Resumen: Entendiendo la evaluación en las políticas públicas como un campo de estudios y prácticas ya denso y plural, este artículo busca problematizar sus desarrollos históricos y su configuración actual desde una perspectiva epistemológica. Esta mirada buscó, por un lado, privilegiar la fuerte relación entre las principales expresiones teóricas y metodológicas de dicho campo; y, por otro, revelar e interpretar las disputas de valoración más sustanciales que configuran sus discontinuidades y contradicciones. Desde el punto de vista del método, esta investigación se desarrolló como un tipo de estado del arte, prestando especial atención a la literatura que se centró en propuestas clasificatorias en la evaluación de políticas públicas. Estructurado en seis apartados, este artículo también ofrece su propio camino para una clasificación, por epistemologías, de la configuración actual del campo.

Palabras-clave: evaluación; políticas públicas, epistemología

FR Pour un regard épistémologique sur l'évaluation des politiques publiques: histoire, théorie et méthode

Résumé: Comprenant l'évaluation dans les politiques publiques comme un champ d'études et de pratiques déjà dense et pluriel, cet article cherche à problématiser ses évolutions historiques et sa configuration actuelle dans une perspective épistémologique. Ce regard cherchait, d'une part, à privilégier la relation forte entre les principales expressions théoriques et méthodologiques d'un tel champ; et, d'autre part, révéler et interpréter les litiges d'évaluation les plus importants qui façonnent ses discontinuités et contradictions. Du point de vue de la méthode, cette recherche a été développée comme un type de état de l'art, en accordant une attention particulière à la littérature axée sur les propositions classificatoires dans l'évaluation des politiques publiques. Structuré en six sections, cet article propose également son propre chemin pour une classification, par épistémologies, de la configuration actuelle du champ.

Mots-clés: évaluation; politiques publiques, épistémologie

1. Introdução: desvendando um pouco da polissemia em torno da avaliação

Há diferentes caminhos para compreender avaliação ou processos avaliativos em políticas públicas. Em comum, porém, podemos dizer que todos eles são voltados para a produção de um tipo muito específico de conhecimento: o conhecimento avaliativo. Discutir a sua natureza (a do conhecimento) pode ser, portanto, um bom ponto de partida para mergulhos mais profundos no tema da avaliação. Pode nos ajudar desde a problematizar a sua definição com campo multidisciplinar de estudos e práticas, suas possíveis configurações e representações – incluindo as principais disputas epistemológicas que encontramos por trás de suas polissemias, suas discordâncias de método, suas dinâmicas de construção de públicos – até chegar à investigação das suas lógicas formativas mais habituais. E estes e outros elementos nos ajudam, também, a compreender seu contexto social de produção, que, cada vez mais, vem ganhando terreno nos processos de governo e de governança, em diferentes escalas de atuação, de ação pública, de formação universitária (mas não só), além de outros contextos sociais que envolvem o grande universo dos processos decisórios públicos, sociais ou privados.

Assim, para esta primeira aproximação, de modo geral, quando a natureza do conhecimento avaliativo é vista como *a posteriori*, *situada* e *eminente empírica*, significa que a sua produção está associada mais diretamente a um quadro de respostas em torno de questões funcionais ou aplicadas à construção de uma resposta metodológica, para uma demanda muito específica de produção de conhecimento sobre um objeto bem definido. Pensar o conhecimento avaliativo deste modo significa concentrar os esforços de pesquisa na definição do seu próprio objeto; na precisão de seus objetivos; na explicitação de seu caminho metodológico e de suas técnicas de construção; na seleção dos materiais de pesquisa; na explicação de quais foram os públicos envolvidos em sua construção; e, claro, na apresentação de seus resultados, inclusive em termos de juízo avaliativo. Pode também significar definir as fronteiras do próprio trabalho avaliativo, ainda que com maior peso para os aspectos metodológicos do que para os epistemológicos – uma escolha que frequentemente se baseia nas ideias de objetivação da realidade e de uma desejada neutralidade das avaliações e de seus processos.

Já quando a natureza do conhecimento avaliativo, de modo geral, é vista como *a priori*, *universal* e *eminente normativa*, pode-se dizer, em grandes linhas, que a avaliação assume, de modo mais direto, uma preocupação com os significados do que é avaliação, com suas ancoragens epistemológicas, com sua compreensão de ciência, com a coerência normativa de quadros metodológicos, bem como com a compreensão dos seus propósitos e a explicitação dos seus interesses e preferências. Reflete, ainda, sua vinculação a campos diferentes de estudos, pois fazer avaliações vendo-se como parte do campo de estudos da Economia, da Sociologia, da Administração Pública, da Antropologia ou das Políticas Públicas, pode significar lidar com conjuntos de saberes, de possibilidades e de perspectivas bastante diferentes. Além disto, quando esta mesma natureza imputa uma dimensão reflexiva ao conhecimento avaliativo, é possível observar que a ideia de aprendizagem pode adquirir alguma centralidade na avaliação e em seus processos. Neste caso, estaríamos diante de experiências de avaliação que, ao serem problematizadas como parciais, também trazem para o debate as questões de valor e uma compreensão da realidade mais subjetiva do que a anterior.

Esta primeira aproximação com o universo da avaliação, por meio da noção de conhecimento avaliativo, embora não seja suficiente para compreender a história ou muito menos a pluralidade do atual universo de teorias e métodos em avaliação em políticas públicas, ajuda-nos a perceber que se trata de um universo com imensas disputas de pontos de vistas, de perspectivas e de narrativas. Ou, para sermos mais precisos: disputas de valores, sobretudo aquelas que se dão no plano da epistemologia, como orientadoras dos demais planos de pesquisa e ação. Para compreender melhor esta diferença, assumirei este universo como um campo de estudos e práticas, inspirando-me na ideia de campo científico de Bourdieu, que o compreende como um espaço estruturado por disputas simbólicas e por posições de poder, os quais criariam suas próprias estruturas de legitimação das verdades, gramáticas e práticas que consideram possíveis, por meio das noções de *habitus* e de agente (Bourdieu, 2004). Assim, se a avaliação fosse um campo, como este campo seria? Ou, ainda, como entender a polissemia em torno da avaliação? Como chegamos a uma construção tão enovelada e com tantos matizes metodológicos? Quais destas disputas estariam situadas no plano da epistemologia? Quais os efeitos disto sobre o próprio campo? E sobre os agentes,

como estes se comportariam no plano da deontologia? E, mais, o que podemos esperar em termos de movimentos mais gerais deste possível campo, inclusive em termos reflexivos, ou melhor, de aprendizagem por reflexividade?

Para tentar trazer alguma luz para estas e outras questões, esta pesquisa foi metodologicamente desenvolvida como um tipo de *estado da arte* ou o *estado da questão*, por buscar compreender, situacionalmente, o tema ou objeto de investigação no contexto atual da ciência (Nóbrega-Therrien & Therrien, 2004) ou de um campo científico, quando trazemos para Bourdieu (2004), como farei aqui. Alinho-me, ainda, à Messina (1998, p.1), para quem “um estado da arte é um mapa que nos permite continuar caminhando; um estado da arte é também uma possibilidade de perceber discursos que em um primeiro exame se apresentam como descontínuos ou contraditórios”. Neste sentido, como mapa, dou particular atenção à literatura que se debruçou sobre propostas classificatórias em avaliação de políticas públicas, para também propor uma classificação própria. Assim, para responder a estes desafios, este artigo está estruturado em seis seções, incluindo esta breve introdução. A segunda seção está voltada para a problematização da avaliação como campo de estudos e práticas; a terceira busca apresentar um breve panorama dos esforços classificatórios no campo; a quarta seção busca oferecer um caminho próprio para tal compreensão, a partir do plano epistemológico; a seguinte apresenta algumas breves palavras sobre meta-avaliação e monitoramento; finalizando com uma breve conclusão sobre uma possível agenda de trabalho para este campo em formação.

2. A avaliação em políticas públicas como campo

Poder ser visto como campo equivale a assumir que a avaliação em políticas públicas já possui uma massa crítica suficiente, para considerar que suas variações internas, suas disputas simbólicas e de posições de poder foram sendo construídas no tempo, até alcançar um estado atual de complexidade e densidade. Em outras palavras, significa assumir a dimensão histórica do que temos hoje, passando a considerar esta pluralidade como o resultado de um processo de desenvolvimento que começa a se aproximar de quase

um século de vida. Isso ocorre, sobretudo, se considerarmos, junto a outros autores, o *New Deal* como um momento que, de algum modo, marca o início da avaliação como (também) relacionada à esfera governativa. E, com tal assunção, emerge a demanda por estudos de reconstrução historicizada deste campo, uma demanda que, muito infelizmente, ainda não vem sendo suficientemente atendida. De fato, são cada vez mais raros os autores que se dedicam a tentar reconstruir a história da avaliação como campo (mesmo quando não o chamam assim), e, muitos deles, quando o fazem, acabam por associar tal reconstrução a um esforço de classificação tipológica para fins de apresentação de um novo percurso metodológico por eles criado e/ou defendido. Dito de outro modo, frequentemente, as classificações se apresentavam como um olhar historicamente perspectivado, para mostrar determinadas lacunas e vícios do “passado” e apontar para um novo “futuro” que estaria chegando.

Os estadunidenses Peter H. Rossi e James D. Wright (1984) estão entre os primeiros autores que buscaram interpretar histórica e conceitualmente a avaliação como campo de estudos. Para eles, as décadas que precedem os anos 1960 constituiriam um período de crescimento e reconhecimento da importância da avaliação, particularmente para o mundo das decisões públicas e governamentais, numa evidente aproximação com as teorias decisórias. Mas foi entre os anos 1960 e o começo dos anos 1970 (particularmente impulsionada pelos *Great Society Programs*) que a avaliação teria atingido o que eles chamam de “*golden age*” da avaliação. Uma idade marcada pela primazia do paradigma experimental controlado e randomizado (*the randomized, controlled experimental paradigm*), o qual teria dado resultados, segundo a leitura de tais autores, como muito próximos de zero. Esta supremacia teria chegado ao fim com a aproximação da era Reagan e com o desenvolvimento de avaliações de tipo compreensivo (*comprehensive evaluations*), defendidas por eles, que se caracterizavam não somente por construir juízos dicotômicos, mas também por explicar o porquê das dinâmicas de desenvolvimento dos seus objetos de avaliação.

Os também estadunidenses Egon Guba e Yvonna Lincoln (1989), centrando seus esforços na realidade dos Estados Unidos, propuseram que a ideia de gerações poderia explicar melhor o desenvolvimento dos estudos avaliativos do que, por exemplo, uma divisão por métodos. Assim, esta dupla de autores identificou quatro gerações de avaliações:

uma primeira que emergia no contexto do *New Deal* e se caracterizava por um viés eminentemente técnico, baseada na ideia de mensuração, pois chegava até mesmo a evitar a emissão de juízo avaliativo. A segunda geração foi batizada como descritiva e teria predominado até meados dos anos 1960, alargando o foco em relação à geração anterior, para os processos que envolviam o objeto avaliado. A geração seguinte passaria a incorporar fortemente a função de julgamento, relacionada, porém, à ideia de um tecnicismo e à descrição funcionalizada do problema avaliativo. Por fim, os autores defendem uma quarta geração, na qual eles mesmos se incluem, que começaria com a era Reagan, chamada de reativa e que representaria uma forte mudança na prática avaliativa, por responder a partir do paradigma construtivista aos dilemas e às falhas das gerações anteriores, tais como a supremacia da perspectiva gerencialista e métodos quantitativos, a dificuldade em considerar os contextos em seus quadros de investigação, bem como em lidar com pluralismo de valores, além da desresponsabilização moral e ética do avaliador.

Aqui, no entanto, considerarei – alinhada mais com Rosana F Boullosa (2017) e Rosana Boullosa e Janaina L. P. Peres (2020), ambas as autoras brasileiras – uma evolução histórica até os tempos mais recentes, caracterizada por cinco fases que se acumulam, tentando alargar o olhar para muito além da experiência dos Estados Unidos. A primeira fase, que em parte também coincide com a de Guba e Lincoln (1989), pois considera o *New Deal* como um importante *starting point*, se caracteriza pela percepção da época. Uma percepção de que o novo ciclo de desenvolvimento econômico desejado só seria possível se pautado por uma racionalidade de tipo linear-sequencial das decisões públicas, de natureza supostamente neutra e sem uma conclusão em formato de juízo ou de síntese avaliativa, já que a função de emitir juízo caberia ao gestor. Aos poucos, estas avaliações começaram a ficar mais extensas e descritivas, com relatórios infundáveis e evidentes dificuldades de construção de sínteses por parte de quem as lia. E, até então, este era um movimento que se restringia quase exclusivamente aos Estados Unidos e só se mantém hoje em grupos muito restritos de ação, ainda que tenha alimentado algumas epistemologias de matriz gerencialista.

Com o final da Segunda Guerra Mundial (1939-1945), este primeiro panorama mais asséptico da avaliação começou a mudar muito

rapidamente, chegando à segunda fase. Politicamente, a ideia de um mundo global mutuamente dependente reorganizava as relações entre os países. Projetos de desenvolvimento e de cooperação, como foi o caso do Plano Marshall (1948-1952), dedicado à reconstrução de parte da Europa, sob a liderança e o financiamento dos Estados Unidos, criou um contexto para que a ideia de avaliação, como subsídio à decisão racional, atravessasse suas primeiras fronteiras geográficas, começando pela Europa. Neste processo de expansão, porém, as avaliações começaram a extrapolar suas possibilidades iniciais de uso, passando a formular juízos para subsidiar, de modo mais direto, as decisões públicas e governativas.

O impulso para abandonar uma natureza meramente descritiva e para assumir uma natureza fortemente prescritiva chegou, inicialmente, com a aproximação aos *policy studies* de Harold Lasswell, de inspiração pragmatista, e, na década seguinte, com as aproximações com a filosofia analítica – período que viria a coincidir com o que Rossi e Wright (1984) chamaram de era de ouro da avaliação. Do ponto de vista epistemológico e metodológico, reforçam-se as raízes na tradição positivista, no paradigma *top-down* de ação dos Estados, na escala macro para as investigações de natureza empírica, com a franca primazia do método científico-experimental. Estas construções se mantêm com público cativo até os dias de hoje, alimentando diferentes abordagens avaliativas.

Esses “anos de ouro”, porém, acabaram sucumbindo à crítica trazida pelo movimento da contracultura na produção científica, responsável por evidenciar os efeitos perversos de um excesso de otimismo que tinha desencarrilhado para certa miopia relativa aos próprios objetos das avaliações, além da supervalorização das figuras do gestor e agora também do avaliador. Estávamos diante de um mundo em mutação, com um forte movimento crítico que colocava em discussão os pilares da construção racionalista e positivista da ciência (Fischer, 2016). No campo da avaliação, este mesmo positivismo, de tradição empiricista, começava a perder força rapidamente e, como consequência, passava-se a assistir à emergência de novas experiências de metodologias *bottom-up* (Albaek, 1998), particularmente aquelas que reforçam a perspectiva da participação social e do empoderamento social.

Esta terceira fase, que começa a ganhar densidade nos anos 1970 e se constrói a partir de uma reação à fase anterior, caracteriza-se pela busca de novas bases epistemológicas para a avaliação, para

além do positivismo que reinava, até então, quase exclusivamente sozinho. Foi nesta fase que as avaliações começaram a chegar, ainda que timidamente, à América Latina e à África, muitas vezes pelos braços da cooperação internacional, reagindo tanto na direção da aceitação passiva como, embora com menor ênfase, da reposta crítica que reforçava as ideias de empoderamento local. Estas novas buscas se ampliaram por toda a década de 1970 e entraram pela década seguinte, produzindo novas abordagens alinhadas, sobretudo, ao construtivismo e à hermenêutica, com períodos de maior ou menor fertilidade que também duram até os dias de hoje.

Na década de 1980, contudo, assistimos a uma retomada da tradição positivista, desta vez alinhada a uma nova compreensão de Estado, da sua estrutura e dos seus papéis. Junto com o movimento do *New Public Management* (Nova Gestão Pública), diante da escassez de recursos de um mundo que estava tentando sobreviver à forte crise do petróleo, assistiu-se a um movimento reformista dos Estados Nacionais, a partir da ideia de que estes deveriam aprender a melhorar as suas performances em termos de eficácia e eficiência, tirando lições da iniciativa privada (Osborne; Goebler, 1992). Estas ideias alimentaram muitas reformas administrativas, inicialmente em países como Estados Unidos, Inglaterra e Nova Zelândia, chegando, posteriormente, a alguns países da América Latina, como foi o caso do Brasil, do México, do Chile e da Argentina, já sob a cartilha do chamado Consenso de Washington¹ (1989), como também viria a acontecer com alguns países da África que estiveram sob os Programas de Ajuste Estrutural do Banco Mundial.

E foi neste contexto que a avaliação entrou de modo pesado na América Latina, como instrumento de gestão associado às reformas de seus Estados, associando-se às noções do novo gerencialismo, tais como as de orientação a resultados, flexibilidade e *accountability* (Bozzi, 2001; Farias, 2003). E em todas estas experiências, de modo mais ou menos intenso, a avaliação era vista como uma importante ferramenta de controle dos processos de reforma, em alguns casos inaugurando a entrada nos países. Neste novo contexto, a avaliação, além de retomar sua faceta mais positivista, pensada como campo de conhecimentos e práticas, passava a assumir, particularmente no contexto latino-americano, diferentes papéis relacionados à diminuição da estrutura administrativa do Estado, à melhoria de seus resultados, ao reforço da capacidade de competir e, claro, à responsabilização

dos tomadores de decisão. No Brasil, não foi diferente, e a avaliação também começou a ganhar espaço na agenda do Estado no contexto reformista de meados dos anos 1990 (Farias, 2003; Souza, 2006). De fato, muitos autores começam a concordar que, desde então, a avaliação vem progressivamente sendo consolidada na agenda governativa do país (Vaitsman *et. al.*, 2006; Gussi & Oliveira, 2016; Jannuzzi, 2018), estruturando, inclusive, suas próprias disputas (Jannuzzi, 2018).

Os anos 2000 começam, assim, com diferentes fases que se acumulam, conformando uma arena cada vez mais densa e plural, tanto no plano epistemológico, quanto no metodológico. Ainda que as visões mais alinhadas ao gerencialismo despolitizado dos anos 1980 ainda continuem a prevalecer, há também espaços para buscas alinhadas às críticas dos anos 1970, bem como às novas ancoragens críticas dos anos 2000, em diferentes partes do mundo. Estes movimentos podem ser vistos, pelo menos parcialmente, como resultados dos esforços realizados em diferentes países, para a formação de mais e mais associações nacionais de avaliadores, em lugares como Reino Unido (1994), Alemanha (1997), Itália (1997), França (1999), Japão (2000), Índia (2002), Senegal (2002), África do Sul (2002), Argentina (2005), Madagascar (2004), Filipinas (2007), Brasil (2008), Eslovênia (2008), Georgia (2008), Equador (2014), México (2014) e Rússia (2014), muitas delas seguindo os modelos pioneiros do Canadá (1981), dos Estados Unidos (1986) e da Austrália (1987).

Neste movimento de entrelaçamento global, assistimos também a grandes esforços de construção de redes regionais, a partir das redes nacionais, como são os casos da Afrea/África (1999), EES/Europa (1996), RELAC/América Latina e Caribe (2013) e APEA/Ásia-Pacífico (2012). Trata-se de organizações imensas, que precisam lidar com as diferentes tradições em avaliação de políticas públicas, quase nunca ancoradas nos mesmos campos disciplinares, com suas disputas internas, além das óbvias variedades de densidade das redes, contextos, línguas, dentre tantos outros elementos. Mesmo com tantas diferenças, assistimos, ultimamente, a uma crescente presença da avaliação nos contextos que envolvem as decisões públicas, seja dentro ou fora de governos. E, sem dúvida, uma das grandes conquistas deste movimento tem sido a campanha global Eval4Action, ativada pela Organização das Nações Unidas (ONU), para esta década (United Nations, 2020).

¹ Williamson, John: "What Washington Means by Policy Reform", in: Williamson, John (ed.): *Latin American Readjustment: How Much has Happened, Washington: Peterson Institute for International Economics, 1989.*

3. Percorrendo os mais expressivos esforços de classificação

Após compreendermos um pouco da história de construção da avaliação, precisamos falar sobre sua atual configuração. Ao longo da construção da avaliação em políticas públicas como campo, alguns autores se dedicaram a tentar entender e classificar a crescente diversidade que percebiam. Estes esforços de classificação são importantes porque, de um lado, ajudam os estudiosos e praticantes a realizarem um importante exercício de posicionalidade nas disputas que envolvem o campo; de outro, acabam estimulando alguns agentes do campo a imprimirem maior força para um dos lados que aparecem nas classificações, reforçando e estimulando disputas que são fundamentais para a expansão do conhecimento e para a formação do *habitus* de seus agentes, auxiliando na validação de suas estruturas de valores.

Dentre estes autores, vale começar pelo australiano Michael Scriven (1967), para quem as avaliações poderiam ser classificadas como formativa ou somativa, independentemente de outras eventuais características. As somativas são aquelas frequentemente posteriores ao fato avaliado, portanto, *ex-post*, que buscam responder questões avaliativas sobre continuidade, encerramento, ampliação ou adoção de seus próprios objetos avaliados (políticas, projetos, programas, ações), desde que estes já tenham apresentado algum resultado; enquanto que, por formativas, compreende-se aquele conjunto de avaliações que buscam produzir conhecimento avaliativo voltado para o desenvolvimento do próprio objeto em avaliação, o qual não precisa, portanto, ter começado a produzir resultados, e que são quase sempre dirigidas para a própria equipe envolvida com a gestão de tais objetos em avaliação. No primeiro caso, do ponto de vista do método, há maior inclinação ao uso de universos amostrais mais amplos do que no segundo caso.

O olhar classificatório de Scriven ganhou novas possibilidades nas mãos de diferentes autores. O estadunidense Marvin C. Alkin (1985) foi pioneiro, ao propor uma classificação de usos possíveis para as avaliações, baseada em uma lista de variáveis que considerava fatores humanos, fatores contextuais e fatores da própria avaliação. Logo em seguida, a também estadunidense Carol Weiss passou a interpretá-las a partir da noção de uso (instrumental, simbólico

e esclarecedor), propondo a ideia de que as avaliações deveriam construir uma Teoria do Programa (Weiss, 1998). Ainda segundo ela, o uso instrumental seria o mais difícil por causa de obstáculos de tipo cognitivo-intelectual (*intellectual-cognitive*) e estrutural-social (*social-structural*), sugerindo alguns caminhos para que os avaliadores aumentassem a potencialidade de uso de suas avaliações. Outros autores, com menor penetração na literatura especializada, também apresentaram olhares classificatórios, como foi o caso do alemão Hans-Ulric Derlien (2001), que propunha as funções de informação, realocação e legitimação; e, por fim, não é possível deixar de falar do estadunidense Michael Q. Patton, que não só utiliza a ideia de classificação a partir da noção de uso (Patton, 1998), como desenvolve sua própria abordagem em tal direção (Patton, 1987).

Também merece destaque a interessante distinção proposta por Marvin C. Alkin e Christina A. Christie, chamada de *evaluation theory tree*, na qual eles propõem um desenho metafórico de uma árvore que teria três ramos principais. O primeiro deles abrigaria todas as avaliações que explicam a si mesmas, a partir da noção de uso, reunindo autores como Daniel Stufflebeam, Michel Patton, Marvin Alkin e Jean King, com algumas pequenas variações e pequenas ramificações próprias. O segundo ramo reuniria os esforços que se definem mais como metodológicos, com nomes muito importantes tais como Carol Weiss, Shen, Thomas Cook, Peter Rossi e Donald Campbell. Por fim, um terceiro, que eles chamam de valorização (*valuing*), reunindo autores como Michael Scriven, Egon Guba e Lincoln, Robert Stake e Ernest House. Na raiz comum estariam as noções de *accountability* e controle, bem como de pesquisa social (*social inquiry*).

Esforços classificatórios também foram empreendidos por outros autores, como foi o caso do italiano Claudio Bezzi (2007), que classifica as avaliações quanto às suas finalidades. Para ele, estas podem ser simplesmente voltadas para informar os decisores sobre os programas que desenvolvem, quase sempre a partir de abordagens funcionalistas, em resposta a demandas específicas; podem oferecer um juízo profissional, com a apresentação de uma síntese na qual o avaliador expõe seus achados e procura dar respostas a perguntas mais concretas de seu demandante, ao invés de limitar-se a fornecer dados, como no primeiro caso; como, também, podem ser voltadas para responder a demandas mais simbólicas, quase sempre situadas no plano político dos objetos avaliados, tais como a prestação de contas de um governo,

a transparência das contas públicas, a comprovação das prioridades de um governo, etc.

Há classificações, também, que assumem somente um elemento das avaliações, como é o caso da classificação segundo a posição do avaliador em relação ao objeto avaliado, as quais poderiam ser interna, externa, mista, ou, para alguns autores, participativa. Para cada uma destas posições, discutem-se as vantagens e desvantagens relacionadas. Outras classificações assumem apenas um conjunto de caminhos metodológicos como linha de separação, como é o caso da classificação entre “avaliações gerencialistas” – ou seja, aquelas voltadas para a gestão – e “avaliações não gerencialistas”, que compreenderiam todo o universo plural não contido no primeiro caso. Há ainda classificações quanto ao *timing* da avaliação, de acordo com o ciclo de vida do objeto avaliado. Neste caso, as avaliações podem acontecer antes, durante ou depois do “fato concreto”. Para as anteriores ao fato (*ex-ante*), podemos falar em diagnósticos, avaliações de viabilidade, avaliação de problemas públicos, avaliações de cenário, avaliações de preferências ou de opinião, dentre outras; para as que acontecem no decorrer da existência do fato, ou seja, *in-itineri*, podemos chamar genericamente de avaliações de implementação, de avaliação de gestão, de avaliação instrumental, dentre outras; e para as que são posteriores ao fato, ou seja, *ex-post*, podemos chamar de avaliação de resultados, de impacto, de efeitos, dentre tantos outros nomes.

As avaliações podem, ainda, ser classificadas a partir de suas tradições de conhecimento: de ontologia realista ou ontologia naturalista. No primeiro caso, de matriz positivista, a missão do avaliador, preferencialmente externo, é explicar o objeto avaliado, pois, nesta ontologia, o papel da ciência é descobrir as leis (verdades) que governam o mundo (realidade). A verdade é sempre objetiva e existe concretamente, independentemente do observador, o que explica, para eles, que a avaliação deve ser isenta, neutra e metodologicamente baseada no controle das variáveis para chegar ao coração da verdade. Já na ontologia naturalista, de matriz construtivista, a missão do avaliador, preferencialmente interno, é compreender o objeto avaliado, pois, nesta ontologia, o papel da ciência é compreender o mundo, as suas partes, sempre de modo aproximado. Neste caso, a verdade é sempre subjetiva, provisória, situacional e dependente do observador, o que explica, para eles, que a avaliação deve ser parte do mundo do

observador e do observado, não neutra e metodologicamente baseada na abordagem hermenêutica e interacional.

4. Uma nova classificação por epistemologias

As classificações existentes são muito úteis, para explorarmos, a partir de diferentes ângulos, os caminhos avaliativos que percorremos ou os relatórios, artigos, cadernos e outras mídias com que nos deparamos. Vale o esforço, assim, de buscarmos enquadrar o que temos em mãos (ou diante dos olhos) nestas diferentes possibilidades descritas na seção anterior. No entanto, a maioria delas acaba por nos falar muito mais sobre as avaliações quando compreendemos o tal do conhecimento avaliativo como de natureza *a posteriori*, *situada e eminentemente empírica*, do que quando o fazemos a partir de uma compreensão de sua natureza como *a priori*, *universal e eminentemente normativa*. Para este segundo caso, parece que estamos diante de certa lacuna, o que nos leva não somente a conclamar a importância de novos estudos avaliativos, como também a nos arriscar neste terreno a ser mais bem ocupado. E o farei, novamente, inspirando-me nos estudos de Boullosa (2018, 2020), Boullosa (*et. al.*, no prelo) e Boullosa e Peres (2020), deslocando, porém, a centralidade da discussão que fazem do plano metodológico, para um cruzamento basicamente entre os planos ontológico e epistemológico.

Nesta tentativa de interpretação da atual configuração da avaliação como um campo cada vez mais denso e plural, o cruzamento proposto nos leva à proposição de classificação por *escolas de pensamento*, que disputam o campo em seus diferentes matizes, sistemas simbólicos, ancoragens valorativas, estruturas de poder, inclusive no que concerne à capacidade de refletir sobre si mesmo e sobre o campo, gramáticas de ação, comportamentos admissíveis para os seus agentes, quadros conceituais, definições metodológicas, inclusive no que concerne ao lócus possível para a investigação empírica, dentre muitos outros elementos. Pensar por escolas significa, ainda, lidar com as diferentes possibilidades de a avaliação relacionar-se com outros campos de estudo, para além dela. Aqui, porém, por força analítica, posicionarei o olhar a partir do campo de estudos em políticas públicas, para interpretar a avaliação como campo, a partir das seguintes cinco

escolas de pensamento: científico-experimental, gerencialista, antropológica, participativa e axiológica.

A escola de pensamento científico-experimental, de matriz positivista sequencial-linear, assume uma noção de realidade como a mais objetiva possível, defendendo a completa dissociação entre fatos e valores – o que significa dizer que os fatos existem como tal, independentemente dos valores de quem, por exemplo, os observa. O campo disciplinar mais próximo é, na maioria das vezes, visto como a Economia. A ideia de conhecimento que possuem é frequentemente definida, por eles mesmos, como de tipo científico, objetivo, empírico, mensurável, neutro e passível de reconstrução. Além disto, tal conhecimento, sobretudo quando articulado no formato de um juízo avaliativo, deve ser baseado em evidências experimentais e quantitativamente mensuráveis.

Filiar-se a esta escola de pensamento também significa considerar o objeto de avaliação como uma *variável independente* que *produz efeitos* sobre uma *dada realidade ou grupo experimental*, aspectos que podem e devem ser empiricamente investigados, do modo mais neutro possível, com a ajuda de grupos de controle (que não receberam a mesma *variável independente*). Esta investigação, de natureza comparativa, se dá por meio da decomposição dos efeitos produzidos em um conjunto articulável e possível de *variáveis dependentes*. Como um dos valores desta epistemologia é a crença na neutralidade da ciência e, portanto, na da avaliação, o avaliador, visto preferencialmente como externo, deve garantir tal neutralidade, apoiando-se em técnicas matemáticas de aleatoriedade, randomização, cálculos diferenciais e integrais, etc. Seus métodos de referência são o experimental, quase-experimental e suas variações de muitas ordens.

A segunda escola de pensamento, a gerencialista, apoia-se numa compreensão funcionalista da ciência, para a qual o conhecimento é visto, também, como objetivo (embora alguma subjetividade seja admissível) e preferencialmente empírico. Do ponto de vista da posicionalidade, é a escola que está mais próxima ao campo de estudos em Políticas Públicas, mas também em Administração Pública, sobretudo em suas versões mais tradicionais, também positivistas e ancoradas aos princípios, respectivamente, da análise racional de políticas públicas e do *new public management*, particularmente nas noções de eficácia, eficiência e efetividade, levadas para a gestão pública. Esta escola se caracteriza por uma forte crença em

uma racionalidade de tipo instrumental, para os sistemas de gestão, podendo até ser superficialmente interpretada como de base pragmatista, bem como em uma racionalidade limitada, para os gestores e envolvidos direta ou indiretamente com os processos de gestão e com os seus objetos – esta última devido ao jogo pela busca de interesses e preferências que interfeririam em suas racionalidades.

Filiar-se a esta escola de pensamento também significa assumir que as avaliações devem produzir conhecimento útil à ação do Estado e, preferencialmente, para contribuir instrumentalmente com um projeto de governo, mesmo quando isso não é explicitamente declarado. Ainda que seja possível pensar em avaliadores internos, aqui também há uma preferência de mercado pelos avaliadores externos, os quais acabam, em diferentes realidades nacionais (países), se dividindo em mercados orientados pelo tipo de avaliação e demandas em função do ciclo de vida dos objetos avaliados. Nesta escola, provavelmente a mais expressiva delas, o avaliador também se assume como um sujeito neutro, filosoficamente menos preocupado com a separação entre fato e valor, mas mais preocupado com a separação entre política e políticas públicas. Seus métodos de referência são a avaliação de eficácia interna e a avaliação custo-benefício.

A escola de pensamento participativa nos remete a uma tradição científica construtivista e defende um tipo de conhecimento que é prático, situado e subjetivo. O campo disciplinar mais próximo é, na maioria das vezes, visto como a Sociologia, ainda que se possam encontrar relações com o Planejamento Urbano e Regional. O tipo de racionalidade mais próxima é a comunicativa, não raro inspirada no filósofo alemão Jürgen Habermas, assumindo que o conhecimento avaliativo socialmente útil é aquele coletivamente construído e pactuado. Por isto, quando efetivamente participativas, as avaliações alinhadas a esta escola, admitem tanto as posições interna ou externa do avaliador, já que o mais importante mesmo é a ativação de um público diretamente relacionado com o objeto avaliado, público este que deverá assumir, metodologicamente, a função de uma “comunidade de avaliadores”.

Filiar-se a esta escola de pensamento também significa assumir que a avaliação carrega consigo uma função democratizante, muitas vezes relacionada à concepção de democracia deliberativa ou participativa. É importante ressaltar, também, a dimensão fortemente pedagógica que

muitas destas avaliações abraçam para si. Além disto, problematizar a aprendizagem em seus processos mais públicos exige um olhar atento para as relações de poder simbólico, contidas nos processos avaliativos. Estas relações podem ser interpretadas ou redirecionadas para objetivos de empoderamento individual e coletivo *dos* e *entre* sujeitos que integram as comunidades de avaliadores, possivelmente incluindo a aprendizagem como uma das categorias a serem observadas. Seus métodos de referência são o *practical participatory evaluation* (P-PE) e a avaliação centrada em participantes. Antes de passarmos para a próxima escola, temos o dever de ressaltar que nem toda avaliação com participação pode ser considerada de epistemologia participativa.

A quarta escola de pensamento, a antropológica, ainda que pequena e muito menos globalmente difundida que as demais, caracteriza-se por uma forte ancoragem no método etnográfico. Possui maior proximidade com a tradição culturalista na ciência, particularmente pelo fascínio que a Antropologia como campo lhe exerce. Assume sem pudores uma noção de conhecimento como uma matriz normativa e subjetiva de significados socialmente construídos, historicamente estruturados e não reproduzíveis. O tipo de racionalidade assumido é frequentemente subjetivista e as avaliações são voltadas para a compreensão da rede de significados que permeiam e estruturam o objeto em estudo, modeladas e compartilhadas por grupos e/ou territórios. A posicionalidade do avaliador não é definida em termos dicotômicos (interno x externo) e este pode chegar a descrever a si mesmo como um observador participante.

Filiar-se a esta escola significa, também, assumir certa proximidade com o método de estudo de caso, circunscrevendo-se, frequentemente, em uma escala micro de investigação. Significa ainda reconhecer que os significados, sobretudo aqueles compartilhados, conformam seus principais materiais de investigação. E, por uma coerência de método, os avaliadores filiados a esta escola, acabam, também, por problematizar suas próprias estruturas de significados. Seus produtos são quase sempre descrições densas e muito aprofundadas dos seus objetos avaliados. Justamente por isto, não raro, as experiências deste tipo sofrem rejeição por parte dos gestores das políticas públicas, por exemplo, que não conseguem perceber a utilidade prática de seus resultados, desperdiçando a riqueza de significados nelas contidas. Seus métodos de referência são a *black box evaluation* e a avaliação antropológica.

Por fim, a escola de pensamento axiológica parte da ideia da existência de múltiplas racionalidades nos processos públicos decisórios. Ancorada em uma tradição construtivista de ciência, com viés, porém, mais interpretativista, esta escola é a que mais se reconhece como próxima ao campo de estudos das Políticas Públicas. Este reconhecimento, contudo, leva-a a se posicionar fora da chamada análise racional e mais próxima dos chamados estudos críticos, para alguns pós-positivistas. Nela, o conhecimento possui uma natureza de tipo normativa, subjetiva, socialmente construída e, o mais importante, implicitamente ancorada em estruturas ativas de valores. A ideia de valor vista como “definidora” dos fatos, acaba por singularizar as avaliações que se alinham a esta escola de pensamento, por lhe atribuírem centralidade como material primordial de investigação. Se os valores conformam a matéria, os caminhos para lidar com ela são interpretativistas.

Filiar-se a esta escola de pensamento pode significar, também, compreender que os valores só podem ser extraídos a partir de argumentos, já que lhe subjazem (Fischer, Forester, 1993). As avaliações, então, deveriam se preocupar em construir as matrizes argumentativas em torno de seus objetos de avaliação – daí, não raro, tais avaliadores receberem o nome de argumentativistas. Além disto, a posicionalidade do avaliador não conforma uma questão importante de pesquisa, pois a discussão da posicionalidade é transferida do avaliador para a avaliação. Neste sentido, o processo de construção das avaliações passa a ser visto como parte de um contexto mais amplo (e, também, fortemente valorativo) de políticas públicas, por compreender que o argumento que está sendo produzido somente encontra realização quando confrontado nas arenas de políticas públicas, inclusive em termos de aprendizagem. Admite a multiplicidade de métodos e instrumentos que podem ser utilizados, desde que normativamente definidos. Seu método de referência é o *practical deliberation model* e suas derivações, a partir do reforço das dimensões política, argumentativa e/ou instrumental das abordagens axiológicas.

Compreender o campo da avaliação por escolas de pensamento pode, também, reforçar três importantes movimentos, quase cadenciados, da avaliação e dos avaliadores. Um primeiro que diz

respeito à própria compreensão do campo e das disputas que o estruturam, o que pode levar ao adensamento do mesmo, ao reforço de suas escolas, bem como da busca de diálogo entre elas. Em seguida, um movimento de busca por uma posicionalidade própria por parte dos avaliadores, mais coerente epistemologicamente e mais afinada deontologicamente com os valores e com suas redes de significados, passando de uma atividade de produção aplicada a uma atividade de produção mais implicada com seu próprio projeto de transformação da realidade (Boullosa, 2020). E, por fim, um movimento de coesão e conformidade metodológica das avaliações e de seus processos, evitando, assim, a mera reprodução de fórmulas e receitas mágicas que terminam por produzir pouco conhecimento avaliativo relevante, contribuindo para a imensa dificuldade de retroalimentação das avaliações nos processos de políticas públicas, ao ponto de a colombiana Karen M. Mokate (2009) ter escrito um artigo sobre como converter o “monstro” da avaliação em “aliado”.

Por último, espera-se que estes movimentos cadenciados consigam, por sua vez, promover o que parece ser uma das características não tão positivas do campo, que é a sua baixa reflexividade - compreendendo a reflexão na mesma linha de Donald Schön (1983), que a define como a capacidade de perceber, refletir e aprender no curso das práticas, de modo situado, mesmo quando se trata de uma prática profissional. Em parte, esta se deve, muito certamente, à natureza excessivamente aplicada do próprio campo, onde o fazer profissional se mostra urgente, disputado, pouco regulado externamente e com ainda insuficientes redes de produção e reprodução do conhecimento, como são os casos das redes profissionais, mas também das redes de formação. Mas, em parte, também se deve a uma baixa articulação entre a produção mais sistematizada de conhecimento normativo, classificatório e/ou histórico e certa abundância de produtos avaliativos construídos sobre seus próprios objetos de avaliação. Aumentar a reflexividade do campo significaria, assim, desta vez referindo-me a Chris Argyris & Donald Schön (1983), investir na ampliação de seus processos de aprendizagem.

5. Breves palavras sobre meta-avaliação e monitoramento

Antes de caminharmos para o final, são necessárias algumas palavras sobre meta-avaliação e monitoramento. A primeira é razoavelmente antiga, ainda que somente agora comece a ganhar seu merecido destaque. Há um consenso de que o termo meta-avaliação teria sido cunhado por Michael Scriven em 1969, não obstante o filipino Petro T. Orata já tivesse falado sobre “avaliação de avaliação” em 1940 (Cook & Gruder, 1978). De modo geral, por meta-avaliação compreendemos as avaliações que possuem como objeto uma outra avaliação, ou mesmo mais de uma, normalmente realizada por um/a outro/a avaliador/a (Shadish, 1998). Contudo, há diferentes caminhos para compreendermos as meta-avaliações, o que abre espaço para uma possível classificação.

Para começar, podemos distingui-las, com Scriven, entre formativas ou somativas. As primeiras estariam voltadas para melhorar a qualidade do desenho das avaliações. O melhor exemplo deste primeiro tipo viria do próprio Scriven, para quem a meta-avaliação deveria ser uma atividade que aconteceria no decorrer do próprio período de formulação da avaliação, de modo a imputar maior qualidade, rigor e profissionalismo à mesma - de fato, ele propõe que as avaliações, antes de serem implementadas, sejam submetidas a um *checklist* (Scriven 2005), que englobaria nove quesitos meta-avaliativos (validade, credibilidade, clareza, concisão, pontualidade, efetividade, custo-viabilidade, custo-efetividade e generalização). Considerar a avaliação como formativa significa, também, assumir como central o problema da aprendizagem, fortemente relacionado ao rigor metodológico, mas, também, à responsabilidade compartilhada pela construção de um campo profissional capaz de autovalidar-se continuamente.

Já as avaliações de tipo somativo estariam voltadas à mensuração ou à validação *ex-post* da qualidade de avaliações já realizadas. Estas aconteceriam ao final do processo avaliativo e poderiam se debruçar sobre toda a avaliação realizada, sobre seu processo, sobre seus resultados, sobre o método, sobre seus instrumentos, dentre outros elementos que a compõe. Além disto, uma meta-avaliação vista como somativa poderia assumir como objeto uma avaliação por inteiro, mais do que uma avaliação ou mesmo somente uma

parte de uma ou mais avaliações. As possibilidades, portanto, seriam muitas, mas todas elas estariam voltas à construção de um juízo meta-avaliativo sobre a qualidade ou o rigor de seus objetos. Esta classificação dicotômica entre somativas e formativas, exclui, porém, outro tipo de meta-avaliação que, apesar de pouco utilizada, começa a ganhar algum destaque. As meta-avaliações concorrenciais (Baslé, 2013), na prática, são avaliações realizadas paralelamente a outra avaliação (geralmente maior e mais importante). Realizadas de modo independente, contemporâneo e separado (muitas vezes até não revelado), estas avaliações concorrenciais são chamadas de meta-avaliações porque objetivam mensurar, quase simultaneamente, a qualidade das avaliações principais, bem como validar os seus resultados à medida que eles vão sendo apresentados.

Há, ainda, outra possível classificação, se pensarmos novamente em termos epistemológicos, mas desta vez em forte relação com o plano deontológico. Trata-se de compreender a meta-avaliação como voltada para responder a questões relativas aos padrões “profissionais” e técnicos das avaliações e aos princípios éticos da profissão, como explica Daniel Stufflebeam (2001), com sua *“Evaluated-centred metaevaluation”*, ao defendê-la como um espaço necessário para discutir a utilidade, viabilidade, propriedade e precisão das avaliações (e dos avaliadores), mas também a integridade, o respeito e a responsabilidade social das mesmas. Do outro lado, situam-se as defesas por uma meta-avaliação que responda a questões pertinentes à relação entre os avaliadores e entre suas avaliações e seus contratantes, como a *“Client-Centred Meta-Evaluation”* de Reneike e Welch (1986) - muitas vezes estas meta-avaliações podem acontecer para mediar relações de conflito entre as partes envolvidas no processo avaliativo.

As meta-avaliações começaram a ganhar um pouco mais de espaço após a formação das muitas associações de classe de avaliadores, como vimos nas seções anteriores, em diferentes países do mundo, culminando com a criação de associações nacionais. Alguns autores são mais otimistas ao defenderem que estaríamos diante de um quase-subcampo da avaliação. Baslé (2013), por exemplo, chega a falar que vivemos uma fase que pode ser considerada como de maturidade (*Maturity period*), após termos vivenciado as fases do desenvolvimento (*Development*), da propagação inicial (*Propagation*) e da difusão (*Enrolment*). Mas o fato é que estas ainda não formam uma massa crítica consistente. Se, por um lado, são muitos os

argumentos de defesa de meta-avaliações, quase sempre ancoradas no aumento da qualidade e do rigor das avaliações, bem como em possíveis interpretações como instrumentos de controle social, por outro lado, são também muitos os argumentos contrários às mesmas, tais como um forte anacronismo em relação às urgências das decisões públicas, à concorrência com as avaliações em contextos de escassez de recursos, bem como à discordância em torno da figura profissional do meta-avaliador.

Sobre o monitoramento, também são importantes algumas palavras, mesmo que breves. Avaliação e monitoramento são duas atividades muito próximas, mas com algumas características que as singularizam. É comum, infelizmente, encontrar autores que utilizam os termos avaliação e monitoramento como sinônimos, sobretudo quando se trata de trabalhos empíricos (Rodrigues, 2014; Hamberger, 2012; Antero, 2008; Wollman, 2007). Esta falsa sinonímia é especialmente prejudicial ao conceito de monitoramento, pois, quase sempre, tais autores estão de fato se referindo à avaliação, e não ao monitoramento, relegando este último a um apêndice ou a uma ferramenta do primeiro. Em comum, avaliação e monitoramento possuem a natureza de pesquisa aplicada, o propósito de apresentar respostas sobre o andamento de objetos (como programas, projetos, ações, etc.), a partir de perguntas específicas, que podem ou não incluir a necessidade de um juízo ou síntese final. De diferente, possuem a relação que estabelecem com a causalidade, pois o monitoramento depende sempre de uma hipótese de funcionamento do objeto que observa. Em outras palavras, o monitoramento lida com padrões de comportamento de dados, podendo ser associado a alertas em caso de variações, para além do que foi previamente considerado como aceitável. Ele é uma ação permanente e continuada de observação de um ou mais objetos que ocorrem na forma de fenômenos, de objetos sobre os quais incide um fenômeno ou de acontecimentos ou eventos que despertam interesse, para algum propósito ou utilidade dentro de um espaço de inserção humana. Em outras palavras, o monitoramento deve ser visto como um processo de acompanhamento contínuo das condições internas e externas de um objeto em seu contexto, à medida que ocorre no tempo e no espaço, abrangendo todas as suas características de interesse.

Essas proximidades e diferenças, entretanto, acabaram por abrir um novo caminho de relação entre tais atividades: o da

complementação (Boullosa, 2009). Ainda que seja muito raro encontrarmos esforços de avaliação e monitoramento sobre um mesmo objeto, as vantagens em reuni-los podem ser muito grandes. Em primeiro lugar, porque teríamos a oportunidade de construir perguntas muito específicas sobre um mesmo objeto. Em segundo, porque um conhecimento avaliativo concreto e complexo sobre o estado, de fato, de um objeto nos daria a possibilidade de construirmos um sistema mais preciso dos pontos ou componentes deste mesmo objeto, passíveis de monitoramento – sobretudo no que concerne à capacidade de tais componentes de traduzirem ou representarem o andamento geral do estado do objeto em um tempo específico. Por fim, porque uma boa avaliação nos forneceria uma espécie de “marco zero” para o monitoramento, ajudando-o a ampliar seus próprios propósitos de controle.

6. Para concluir, podemos falar em uma nova agenda de trabalho?

Para falarmos em uma possível agenda de trabalho, precisamos observar com atenção os atuais movimentos em curso no campo da avaliação em políticas públicas. E parece que eles são dois. O primeiro é mais institucionalizado e tem a ver com a confluência de muitos esforços oriundos das novas associações, agências e redes em avaliação, que encontramos em diferentes países do mundo, mas também com as redes de segundo nível que vão sendo montadas a partir destas redes nacionais. A partir deste movimento, que foi fundamental para a criação da campanha global pela avaliação ativada pelas Nações Unidas, a Eval4Action, assistimos aos esforços complementares de profissionalização das redes nacionais e regionais e de criação de canais permanentes de divulgação. São exemplos as revistas especializadas, a inclusão sistemática em agendas oficiais de governos, as empresas e organizações da sociedade civil, a promoção qualificada pelas agências de cooperação internacional e de desenvolvimento local, além de uma oferta mais sistematizada para a formação qualificada em avaliação.

O segundo movimento é mais sutil, menos articulado, mas talvez ainda mais poderoso. Ele encontra raízes na ideia formativa

de Scriven (1967), no modelo compreensivo de Rossi e Wright (1994), no paradigma construtivista de Guba e Lincoln (1989), na teoria do programa de Weiss (1998), na argumentação de Fischer (1995) e na trama valorativa de Boullosa (2020). Mas, vai além, buscando compreender a avaliação em sua dimensão política de transformação social por meio de uma noção mais plural e fortemente democrática de conhecimento avaliativo, ancorado em novas epistemologias, como vem acontecendo com os estudos sobre linguagem na Europa, os estudos subalternos na Ásia, os estudos decoloniais na América Latina e na África e os estudos feministas, distribuídos nas Américas como um todo. Estas novas incursões já nascem de uma busca por uma posicionalidade normativa e axiológica que, ao mesmo tempo em que enrique os processos avaliativos – trazendo duplas conceituais com profundidades normativas inovadoras, tais como poder e ética (Foucault, 2008), interseccionalidade e opressão (Crenshaw, 2017), corpo e política (Safatle, 2015), emoção e cidadania (Marcus, 2008) aprendizagem e transgressão (Hooks, 1994), experiência e políticas públicas (Peres, 2020), aprendizagem e multiatorialidade (Boullosa, 2019) – também apontam para um futuro engajado da avaliação em políticas públicas em processos de democratização das sociedades, por meio da democratização do conhecimento.

Referências

ALBAEK, Erik. “Knowledge, interests and the many meanings of evaluation: a developmental perspective”. *Scandinavian Journal of Social Welfare*, 7: 94-98, 1998.

ALKIN, Marvin C.. *A guide for evaluation decision makers*. Newbury Park, CA: Sage, 1985.

_____; CHRISTIE, Christina A. *Evaluation Theory Tree Revisited*, in Alkin, M.C. (Ed.), ‘Evaluation Roots: Tracing Theorists’ Views and Influences’ (pp. 381 – 392). Sage Publications, 2004.

ARGYRIS, Chris. & SCHON, Donald. *Organizational learning*. Reading, Mass.: Addison – Wesley, 1978.

BASLÉ, Maurice. Méta Évaluation Des Politiques Publiques Et Qualité Des Évaluations. In: Revue-Francaise d'Administration Publique. Nº 148, 2013, p. 1017-1027, 2013.

BEZZI, Claudio. Il Disegno della Ricerca Valutativa. Milão: Franco Angeli, 2007.

BOULLOSA, Rosana de Freitas. Introdução à Avaliação e Monitoramento de Projetos Sociais. In: Boullosa, R.F. & Araújo, E. T. (2009) Avaliação e Monitoramento de Projetos Sociais. 1. ed. Curitiba: IESDE. v. 1000. 262p, 2009.

_____. Gestão Social e Avaliação. In: João Martins de Oliveira Neto; Jeová Torres Silva Junior. (Org.). Gestão Social. 1ed. Fortaleza: Fundação Demócrito Rocha/UANE/BID/STDS, v. 1, p. 267-288, 2018.

_____. Mirando ao revés as políticas públicas: o desenvolvimento de uma abordagem crítica e reflexiva para o estudo das políticas públicas. Publicações da Escola da AGU, série especial, pp. 89-105, 2019.

_____. Para onde tem nos levado a pandemia? Entre tantos desamparos públicos, precisamos também falar sobre avaliação em políticas públicas. Revista NAU Social v.11, n.21, p. 427 – 442 Nov 2020 / Abr 2021.

_____; OLIVEIRA, Breyner; ARAÚJO, Edilson T.; GUSSI, Alcides. Por um antimanual de avaliação de políticas públicas. Aprovado pela Revista da Rede Brasileira de Monitoramento. In press.

_____; PERES, Janaína Lopes Peres. Análise e elaboração e gestão de programas e projetos de desenvolvimento socioterritorial. Salvador: UFBA, Escola de Administração, 2020.

_____; _____; BESSA, Luiz Fernando Macedo. Por dentro do campo: uma narração reflexiva dos estudos críticos em políticas públicas. Revista Organizações & Sociedade, Escola de Administração, Universidade Federal da Bahia, Salvador. In press.

BOURDIEU, Pierre. Le champ scientifique : Actes de la recherche en sciences sociales. Paris: Ed. du Seuil, 1976.

BOZZI, Sonia Ospina. “Evaluación de la gestión pública: conceptos y aplicaciones en el caso latinoamericano”. Revista do Serviço Público, 52 (1): 25-55, 2001.

COOK, Thomas D.; GRUDER, Charles L. Metaevaluation research. Evaluation Quarterly, 2(1), 5-51, 1978.

CRENSHAW, Kimberlé. On intersectionality: essential writings. Publisher : New York: The New Press, 2017.

FARIA, Carlo Aurélio Pimenta. A política da avaliação de políticas públicas. Revista Brasileira de Ciências Sociais, São Paulo, v. 20, n. 59, p. 97-109, out, 2005.

FISCHER, Frank. Evaluating Public Policy. Chicago: Nelson-Hall, 1995.

_____. Beyond empiricism: Policy analysis as deliberative practice. In M. HAJER & H. WAGENAAR (Eds.), Deliberative Policy Analysis: Understanding Governance in the Network Society (Theories of Institutional Design, pp. 209-227). Cambridge: Cambridge University Press, 2003.

_____; FORESTER, Jonh. The Argumentative Turn in Policy Analysis and Planning. London: Duke University Press, 1993.

FOUCAULT, Michael. O Nascimento da Biopolítica. 1. ed. São Paulo: Martins Fontes, 2008.

GUBA Egon. & LINCOLN Yvonna S. Paradigmatic controversies, contradictions, and emerging confluences. In: Denszin N, Lincoln YS editors. Handbook of qualitative research. 2nd ed. Thousand Oaks: Sage Publication; p. 163-88, 2000.

GUSSI, Alcides F.; OLIVEIRA, Breyner. R. Políticas Públicas e outra perspectiva de avaliação: uma abordagem antropológica. Revista Desenvolvimento em Debate, v. 4, n. 1, p. 83-101, 2016.

HOOKS, Bell. *Teaching to Transgress: Education as the Practice of Freedom*. Routledge, 1994.

JANNUZZI, Paulo Martino. Mitos do Desenho Quase-experimental na Avaliação de Programas. *Revista Nau Social*: Vol. 9, No 16, 2018.

MARCUS, George. E. *Le citoyen sentimental: Emotions et politique en démocratie*. Collection Sociétés en mouvement. Presses de Sciences Po, Paris, 2008.

MESSINA, Graciela Estudio sobre el estado da arte de la investigación acerca de la formación docente en los noventa. 1998. Trabalho apresentado à Reúñion de Consulta Técnica sobre investigación en Formación del Profesorado, Organización de Estados Ibero-Americanos para la Educación, la Ciencia y la Cultura, México, 1998.

MOKATE, Katia. Convirtiendo el “monstruo” en aliado: la evaluación como herramienta de la gerencia social. Instituto Interamericano para el Desarrollo Social – INDES, 2009.

NÓBREGA-TERRIEN, Sílvia Maria; TERRIEN, Jacques. O estado da questão: sua compreensão na construção de trabalhos científicos: reflexões teórico-metodológicas. *Estudos em Avaliação Educacional*, v. 15, n. 30, p. 5-16, jul./dez, 2004.

ORATA, Pedro T. “Evaluating evaluation.” *J. of Educational Research* 33: 641, 1940.

OSBORNE, David; GAEBLER, Ted. *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector*. New York: Addison – Wesley, 1992.

PATTON, Michael Quinn, *The evaluator’s responsibility for utilization*. *Evaluation Practice*, 9, 5-24, 1988.

_____. *Utilization-focused evaluation* (3rd ed.). Beverly Hills, CA: Sage, 1997.

PERES, Janaína Lopes Peres. *Reinterpretando o fluxo de políticas públicas a partir da experiência: do pragmatismo crítico ao Hip Hop da Ceilândia/DF* (Tese de Doutorado), Programa de Pós-Graduação em Desenvolvimento, Sociedade e Cooperação Internacional – PPGDSCI. Brasília: UnB, 2020.

REINEKE, Robert A., & WELCH, Wayne. Client centered meta-evaluation. *American Journal of Evaluation*, 7(3), 16-24, 1986.

RODRIGUES, Roberto Wagner Silva. Monitoramento (entry). In: Rosana de Freitas Boullosa. (Org.). *Dicionário para a Formação em Gestão Social*. 1ed. Salvador: Editora CIAGS, v. 1, p. 144-148, 2014.

ROSSI, Peter H.; WRIGHT, James. D. Evaluation research: an assessment. *Annual Review of Sociology*, v. 10, p. 331-352, 1984.

SAFATLE, Vladimir. *Circuito dos Afetos: Corpos políticos, Desamparo, Fim do Indivíduo*. São Paulo: Cosac Naify, 2015.

SCHÖN, Donald. *The Reflective Practitioner: How Professional Think in Action*. New York: Basic Books, 1983.

SCRIVEN, Michael. The methodology of evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39-83). Chicago: Rand-McNally, 1967.

_____. “An introduction to meta-evaluation”. In: *Educational Product Report 2*: 36-38, 1969.

_____. Key Evaluation Checklist. Oct 23, 2005. pp. 5-7.

SHADISH, William. “Evaluation Theory is Who We Are”. In: *American Journal of Evaluation*, Vol. 19, (1), p. 1-10. Shaver, P. & Staines, 1998.

SOUZA, Celina. *Políticas públicas: uma revisão da literatura*. *Sociologias*, Porto Alegre, n. 16, dez, 2006.

STUFFLEBEAM, Daniel. L. The methodology of metaevaluation as reflected in metaevaluations by the Western Michigan, 2000.

UNITED NATIONS. The Eval4Action campaign – A Decade of Action to deliver the SDGs by 2030. ONU, 2020. Access: <https://www.eval4action.org/>.

VAITSMAN, Jeni; RODRIGUES, Roberto W. S.; PAES-SOUSA, Rômulo. O sistema de avaliação e monitoramento das políticas e programas sociais: a experiência do Ministério do Desenvolvimento Social e Combate à Fome do Brasil. Brasília, DF: Unesco, 2006.

WEISS, Carol. Evaluation Research: Methods for Assessing Program Effectiveness. Englewood Cliffs: Prentice-Hall, 1972.

WEISS, Carol. Evaluation: Methods for Studying Programs and Policies. Pearson, 1994.

WOLLMAN, Helmut. Policy evaluation and evaluation research. In: FISCHER, F, M, G.; SIDNEY, M. (eds). Handbook of public policy analysis – theory, politics and methods, 2007.