

Ensaio teórico sobre as avaliações de políticas públicas

LILIAN RIBEIRO DE OLIVEIRA¹
CLAUDIA SOUZA PASSADOR¹

¹ UNIVERSIDADE DE SÃO PAULO (USP) / FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE, RIBEIRÃO PRETO – SP, BRASIL

Resumo

A partir da identificação da atual e crescente demanda por serviços públicos de maior qualidade, eficácia e eficiência, a avaliação de políticas públicas se torna cada vez mais indispensável. Sublinha-se neste artigo a importância dessa etapa do ciclo político, referente ao exercício da avaliação como ferramenta crucial para o aperfeiçoamento e o desenvolvimento da gestão de políticas públicas, principalmente as de caráter multidimensional. A pesquisa se caracteriza como um ensaio teórico, com o intuito de construir uma análise ampla e integrada sobre as avaliações de políticas públicas com base na literatura pertinente, além de compreender e propor reflexões quanto ao modelo vigente no que concerne, principalmente, às ações executadas pelos organismos multilaterais. Como conclusão, os dados demonstram que existe no Brasil, ainda que no meio acadêmico, o fortalecimento do estudo da avaliação para buscar as melhores práticas e teorias, contribuindo sobremaneira com a criação de uma “cultura” avaliativa capaz de abarcar a complexidade e as especificidades de cada política ou programa em questão.

Palavras-chave: Marco teórico. Avaliação. Avaliação de políticas públicas.

Theoretical essay on public policy evaluations

Abstract

From the identification of the current and growing demand for higher quality, efficient and effective public services, the evaluation of public policies becomes increasingly indispensable. This article highlights the importance of this stage of the political cycle regarding the exercise of evaluation as a crucial tool for the improvement and development of the management of public policies, especially those of a multidimensional nature. This research is characterized as a theoretical essay with the purpose of constructing a broad and integrated analysis on public policy assessments based on the published literature, in addition to understanding and proposing reflections on the model currently adopted mainly in relation to actions carried out by multilateral agencies.

Keywords: Theoretical Framework. Evaluation. Public Policies Evaluation.

Ensayo teórico sobre las evaluaciones de políticas públicas

Resumen

A partir de la identificación de la demanda actual y creciente de servicios públicos de mayor calidad, eficiencia y eficacia, la evaluación de las políticas públicas se vuelve cada vez más imprescindible. En este artículo se resalta la importancia de esta etapa del ciclo político en relación con el ejercicio de evaluación como una herramienta fundamental para la mejora y desarrollo de la gestión de las políticas públicas, en especial las de carácter multidimensional. Por lo tanto, la investigación se caracteriza por ser un ensayo teórico, con el fin de construir un análisis completo e integrado sobre la evaluación de las políticas públicas basadas en la literatura pertinente, además de entender y proponer reflexiones sobre el modelo vigente, principalmente, en relación con las medidas adoptadas por las organizaciones multilaterales.

Palabras clave: Marco teórico. Evaluación. Evaluación de Políticas Públicas.

Artigo submetido em 20 de junho de 2017 e aceito para publicação em 13 de novembro de 2018.

DOI: <http://dx.doi.org/10.1590/1679-395169657>

INTRODUÇÃO

Ainda que seja amplamente discutida no Brasil e no mundo, a expressão “política pública” assume diversos significados. De acordo com Secchi (2014), países de língua latina apresentam dificuldades para distinguir dois termos da língua inglesa: *politics* e *policy*. Para Bobbio, Matteucci e Pasquino (1998), o primeiro se refere a um conjunto de atividades humanas atinentes à *polis* ou ao Estado; e, conforme Secchi (2014), o segundo denota algo mais concreto, que tem relação com ações e decisões. Assim, a “política pública” se alinha ao significado de *policy*, ou seja, diz respeito ao processo de construção, atuação e decisão política. Pode-se dizer que a política pública está intimamente ligada a ações estatais, porém, existem outras correntes de pensamento que indicam um cenário multicêntrico (SECCHI, 2014). Na abordagem multicêntrica se constrói uma rede de políticas públicas (*policy networks*) conformada por atores governamentais, organizações privadas e não governamentais e organismos multilaterais (FREY, 2000; SECCHI, 2014). Sob o ponto de vista conceitual, além de envolver diversos atores na formulação ou implementação, as políticas públicas são multidisciplinares e repercutem na economia e nas sociedades (SOUZA, 2006; HOWLETT, RAMESH e PERL, 2013). Tendo em vista as definições supracitadas, infere-se que o estudo da política pública é um fenômeno complexo com multiplicidade de atores internos ou externos ao Estado (HOWLETT, RAMESH e PERL, 2013). Desse modo, a busca por processos analíticos é um caminho para diminuir a complexidade do objeto e simplificar as análises. Para tanto, estabeleceu-se a divisão em etapas, sob a égide do ciclo político (*policy cycle*) e em relação às diversas fases do processo político dispostas de forma sequencial e interdependente (FREY, 2000; HOWLETT, RAMESH e PERL, 2013; SECCHI, 2014). Dentre os formatos observados na literatura, optou-se por descrever as resoluções aplicadas de problemas e os estágios correspondentes ao ciclo político, com base em Howlett, Ramesh e Perl (2013) e Secchi (2014) – o Quadro 1 sintetiza essas conexões.

Quadro 1

Resolução aplicada de problemas e sua relação com o ciclo político

Resolução aplicada de problema	Estágios no ciclo político segundo Howlett, Ramesh e Perl (2013)	Estágios no ciclo político segundo Secchi (2014)
Reconhecimento do problema	Montagem da agenda	Identificação do problema
Propostas de solução	Formulação da política	Formação da agenda/formação de alternativas
Escolha da solução	Tomada de decisão política	Tomada de decisão
Efetivação da solução	Implementação da política	Implementação da política pública
Monitoramento dos resultados	Avaliação da política	Avaliação da política pública
Término ou extinção ¹	–	Extinção da política pública

¹Estudos sobre a extinção de políticas públicas começaram a ter destaque nos países desenvolvidos na década de 1970. Eles realizavam diversas ações sociais vinculadas ao Estado de bem-estar, postas em xeque devido à resolução de um problema específico: a ineficiência da ação ou a perda de importância e saída da agenda política (SECCHI, 2014).

Fonte: Elaborado pelas autoras.

Por se tratar de um campo holístico com vários desdobramentos, este estudo se propõe a analisar especificamente o processo histórico de estruturação da etapa do ciclo político denominada *avaliação*. Dada sua importância conceitual, metodológica e de desenvolvimento, a avaliação de políticas, programas ou processos públicos se insere na agenda de pesquisa contemporânea como ferramenta essencial para o aperfeiçoamento das políticas públicas em todas as esferas de governo e no cenário global. A Organização das Nações Unidas (ONU), o Banco Mundial, o Banco Interamericano de Desenvolvimento (BID), a Organização para a Cooperação e o Desenvolvimento Econômico (OCDE), a Comissão Econômica para a América Latina e Caribe (Cepal)

e o Centro Latino-Americano de Administração para o Desenvolvimento (Clad) surgem como protagonistas na geração de metodologias e projetos avaliativos (RAMOS e SCHABBACH, 2012).

O uso efetivo de processos avaliativos pode contribuir com a transparência dos atos públicos, bem como apresentar aos cidadãos formas de controle e acompanhamento das ações governamentais, garantindo, assim, a legitimidade das políticas ou programas desenvolvidos (WEISS, 1999; ALA-HARJA e HELGASON, 2000; MOKATE, 2002; VEDUNG, 2009; RAMOS e SCHABBACH, 2012). Contudo, vale ressaltar que, por ter um caráter político, as avaliações podem contribuir com a manutenção ou não de determinada política, por vezes atrelada às proposições e estratégias das partes envolvidas (RAMOS e SCHABBACH, 2012; ARRETCHÉ, 2001; CRUZ, 2015).

Carvalho (1999) pontua que a avaliação apresenta alto grau de complexidade e especificidade, fatores que elevam a dificuldade na incorporação da atividade avaliativa no dia a dia dos gestores, já que existem múltiplos atores, objetivos, causas, efeitos e desfechos. Nesse sentido, para contribuir com o debate sobre avaliação de políticas públicas, este artigo destina os próximos tópicos ao resgate histórico do estudo da avaliação, além do detalhamento dos modelos, tipos e metodologias existentes na literatura e as tendências internacionais recentes inseridas nessa discussão. Para a construção deste ensaio teórico, optou-se por busca avançada nas seguintes bases de dados: SCOPUS, Web of Science, Google Scholar, SciELO, Banco de Teses e Dissertações da USP, BVS e Portal de Periódicos da Capes. As principais palavras-chave da investigação foram: “políticas públicas”; “avaliação”; e “avaliação de políticas públicas”. Não obstante, tal busca também resultou em ampliação da literatura por meio dos artigos analisados.

REVISÃO TEÓRICA

A literatura considera como marco de proliferação de estudos avaliativos o final da década de 1950. De acordo com Vedung (2010), o processo histórico dos estudos sobre avaliação no contexto global pode ser dividido em quatro ondas de difusão. Cada onda é indicada por um período de tempo e moldada de acordo com questões sociais, políticas e econômicas. A construção dessa metáfora remete à constante renovação do cenário decorrente do movimento repetitivo das ondas e do constante depósito de sedimentos para a criação de uma base (VEDUNG, 2010). Utilizando esse enquadramento teórico, opta-se por descrever a evolução dos estudos avaliativos usando a concepção das ondas de difusão de Vedung (2010).

A primeira onda, ou fase de difusão dos estudos de avaliação, remonta ao final da década de 1950, com consolidação em meados de 1960 (DERLIEN, 2001; TREVISAN e BELLEN, 2008; VEDUNG, 2010; PICCIOTTO, 2015), e é caracterizada como “a onda orientada para a ciência” (VEDUNG, 2010). Os estudos pertencentes a esse recorte de tempo estão atrelados ao surgimento do *Welfare State*, principalmente nos países com aspirações reformistas naquele período, destacando-se: Estados Unidos da América (EUA), França, Alemanha, Suécia e Canadá (DERLIEN, 2001; VEDUNG, 2010; PICCIOTTO, 2015).

Os principais objetivos dos estudos dessa fase eram acompanhar e analisar os resultados das políticas implementadas; Derlien (2001) atribui a esse conjunto de estudos a função de informação. Trata-se de avaliações com o intuito de averiguar os efeitos e resultados decorrentes da política ou programa, bem como identificar possíveis consequências positivas ou negativas. Esses estudos também tinham por objetivo verificar pontos a melhorar e apontar falhas recorrentes (DERLIEN, 2001). “O foco era a melhoria dos programas, e os gerentes tinham interesse em usar a avaliação como mecanismo de *feedback*” (TREVISAN e BELLEN, 2008, p. 537). Para Vedung (2010), foi nesse período que as avaliações emergiram de um contexto de racionalidade radical; a estrutura do pensamento avaliativo foi pautada pela noção de “cientificação” dos processos de política pública e da administração pública. A avaliação traria ao governo ares de racionalidade com arcabouço científico e baseado em fatos (ROSSI e WRIGHT, 1984; WEISS, 1999; VEDUNG, 2010). Em seu artigo sobre reforma no setor público, Thoenig (2000) reforça a racionalidade descrita por Vedung (2010) e confirma a função de informação proposta por Derlien (2001), afirmando que, se o uso da avaliação é orientado à ação, ela deve concentrar-se em fornecer conhecimento utilizável. Para o autor, é muito mais provável que a avaliação seja aceita se seu aspecto de informação for ressaltado (THOENIG, 2000).

O segundo período de tempo, ou segunda onda descrita por Vedung (2010), denominada “onda orientada ao diálogo”, configura-se em meados da década de 1970, em um contexto no qual pesquisadores argumentavam que as avaliações deveriam ser mais pluralistas, impulsionados por ideais progressistas (VEDUNG, 2010; PICCIOTTO, 2015).

Discute-se a inserção de todos os atores envolvidos na intervenção a analisar, desde pesquisadores e políticos até operadores da política ou programa em questão (VEDUNG, 2010). Rossi e Wright (1984) destacam o crescimento do uso de pesquisas qualitativas nesse período, modelo convergente ao pensamento inclusivo da época, justificando seu uso baseado nos custos menores dessas pesquisas em nível local e com maior utilidade e flexibilidade para os gestores. Tais modelos avaliativos suportariam melhor a abrangência participativa e trariam maior contato com a realidade (ROSSI e WRIGHT, 1984; VEDUNG, 2010). Sintetizando as características dessa onda, Vedung (2010, p. 270, tradução nossa) explica que, “em contraste com a onda induzida pela ciência, a onda orientada ao diálogo repousou sobre racionalidade comunicativa. Em vez de produzir verdades, a avaliação dialógica faria gerar amplos acordos, consensos, aceitabilidade política e legitimidade democrática”.

Por outro lado, a terceira onda, denominada por Vedung (2010) “onda neoliberal”, introduzida no início da década de 1980, modificou mais uma vez o padrão dos estudos de avaliação. Vedung (2010) e Picciotto (2015) descrevem esse período como uma inundação neoliberal no campo de estudos avaliativos, com foco em redução do Estado e promoção do livre mercado e de parcerias público-privadas. Nos EUA, as novas concepções avaliativas acompanharam o desmonte dos programas sociais promovido por Ronald Reagan (ROSSI e WRIGHT, 1984). O modelo de gestão emergente era o *New Public Management* (NPM), com maior focalização nos resultados em relação aos processos. Vedung (2010, p. 270, tradução nossa) complementa argumentando que “descentralização, desregulamentação, privatização, a sociedade civil e, em particular, orientação ao cliente se tornaram novos *slogans*. Anteriormente considerado a solução para os problemas, o setor público se tornou o problema a ser resolvido”. Derlien (2001) atribui a essa onda a função de mapeamento e os estudos se concentram em diagnosticar as políticas ou programas que podem ser cortados a partir de resultados negativos apresentados e de levantamentos das consequências da privatização de certas funções ou atividades do Estado; busca-se a eficiência financeira na condução de políticas e programas públicos (DERLIEN, 2001).

Em contribuição, Ala-Harja e Helgason (2000) indicam que, nessa fase, a avaliação foi entendida como ferramenta capaz de justificar políticas e realocar recursos financeiros. No âmbito governamental, os avaliadores se concentram no Ministério da Fazenda e em órgãos de auditoria. Em conclusão, Vedung (2010) sintetiza que as atividades avaliativas da terceira onda se voltaram principalmente a prestação de contas, medição de desempenho e consumo, busca pela qualidade dos serviços prestados e avaliações comparativas.

Em continuidade, a quarta e última onda descrita por Vedung (2010) é denominada “onda da evidência: o retorno da experimentação”. Essa onda de estudos avaliativos tomou corpo nas décadas de 1990 e 2000, principalmente em países do Atlântico Norte e nos países nórdicos. Trata-se de estudos baseados em evidências e, nesse sentido, Picciotto (2015) enfatiza que os envolvidos em estudos de avaliação estão “surfando” essa quarta onda. Os estudos focalizados na evidência científica almejam separar o que funciona do que não funciona (VEDUNG, 2010; PICCIOTTO, 2015).

Os autores destacam que estudos com abordagens experimentais, com uso de randomização (grupos controle e tratamento) ou quase experimentais, têm por objetivo principal a verificação dos impactos causados pela política ou programa público (VEDUNG, 2010; PICCIOTTO, 2015). Complementarmente, Silveira, Vieira, Capobiango et al. (2013) retomam o conceito de avaliação que busca entender e conhecer as consequências de uma ação, bem como antecipar seus possíveis resultados. Nesse entremeio, organismos internacionais, acadêmicos e governos começaram a realizar revisões sistemáticas da literatura, com o intuito de fortalecer as bases da pesquisa, garantindo maior racionalidade ao processo. Vedung (2010) classifica esse movimento de busca incessante por provas científicas como um retorno ao modelo baseado na ciência, porém, com nova roupagem. Na visão de Derlien (2001, p. 107, tradução nossa), a avaliação nesse momento tem a função de legitimação: “acredita-se que a evidência científica justifica decisões políticas, seja para melhorar, reduzir ou eliminar os programas”.

Diante do panorama internacional apresentado, mostra-se necessário um recorte geográfico para a discussão sobre o campo de estudos de avaliação no Brasil. Diferente dos países de grande tradição na avaliação de políticas e programas públicos, o Brasil entrou tardiamente nessa fase do ciclo político. De acordo com Mokate (2002, p. 90, tradução nossa):

Apesar do que foi falado e conforme publicado sobre a importância dos processos de avaliação, ainda é incomum encontrar programas ou políticas sociais na América Latina que têm um rigoroso processo de avaliação sistemática incorporado em processos diários de gestão e tomada de decisão.

A preocupação em avaliar políticas ou programas públicos é algo recentemente incorporado à agenda política, dado que a prioridade dos gestores era canalizada para o processo de formulação de políticas, com pouca ênfase às outras etapas do ciclo

(COSTA e CASTANHAR, 2003; ALVES e PASSADOR, 2011; CRUMPTON, MEDEIROS, FERREIRA et al., 2016). Essa incorporação se reporta a meados da década de 1980, após a experiência de reformas estruturais no aparato estatal propostas em várias partes do mundo, com o intuito de reduzir os custos do governo e torná-lo mais eficiente (PAULA, 2005). Tem-se, então, uma utilização maior e menos fragmentada da avaliação no Brasil apenas no decorrer da terceira onda descrita por Vedung (2010) ou durante a função de legitimação proposta por Derlien (2001).

Por outro lado, a adoção de práticas avaliativas em países como o Brasil visam a fortalecer os arranjos institucionais e proporcionar aos gestores a análise das políticas adotadas (CRUMPTON, MEDEIROS, FERREIRA et al., 2016). Em seu estudo, Crumpton, Medeiros, Ferreira et al. (2016) compararam os estudos avaliativos realizados no Brasil e nos EUA, com o intuito de apresentar o grau de estabelecimento desse tipo de pesquisa nos países; como resultado, tem-se que, diferente dos EUA, a área de pesquisa em avaliação não está totalmente estabelecida no Brasil, ainda que haja um esforço por parte dos pesquisadores para alterar esse quadro.

Acredita-se que, no caso brasileiro, as bases de dados públicas disponíveis apresentam problemas de confiabilidade (algumas apresentam dados declarados), dados indisponíveis e descontinuidade na produção de indicadores; tais fatores podem contribuir com a baixa adesão a estudos avaliativos quantitativos. Assim, as avaliações no país estariam entre a segunda onda proposta por Vedung (2010), com maior quantidade de avaliações qualitativas de caráter local e análise de políticas específicas, e a terceira onda, voltada principalmente a prestação de contas, medição de desempenho e busca da qualidade dos serviços prestados.

De qualquer modo, é oportuno frisar que existe, no meio acadêmico, o fortalecimento do estudo da avaliação para buscar as melhores práticas e teorias, contribuindo de modo substancial com a criação de uma “cultura” avaliativa capaz de abarcar a complexidade e as especificidades de cada política ou programa em questão (CARVALHO, 1999; PNUD, 2016).

A trajetória de evolução dos processos avaliativos no Brasil e no mundo facilita a compreensão de que não há um único método ou padrão definido na atualidade. Percebe-se a multiplicidade de abordagens e metodologias abrangendo um amplo leque de opções. Partindo do pressuposto de que as avaliações não são iguais, Cohen e Franco (2008) e Ramos e Schabbach (2012) sugerem a descrição das diferentes formas de avaliação segundo uma série de critérios, tais como: uso, tempo, objetivos, quem realiza, a quem se destina, qual é o desfecho esperado, dentre outros. Costa e Castanhar (2003) denominam a variedade de conceitos e metodologias dos estudos de avaliação um *emaranhado conceitual*. Os autores indicam que o conhecimento de diversas alternativas teóricas contribui sobremaneira para a escolha do método mais adequado para cada tipo de política ou programa a ser avaliado.

Segundo Faria (2005), pode-se apontar ao menos quatro tipos de uso para a avaliação: instrumental; conceitual; persuasivo; ou de esclarecimento. O uso instrumental remete à importância da qualidade e à adequada divulgação dos resultados alcançados, para tornar a avaliação algo palpável e factível de empregabilidade. Weiss (1998) acrescenta que o uso instrumental da avaliação pode orientar os decisores, pois possibilita uma retomada à história do programa analisado, fornece *feedback* aos praticantes da avaliação e também confere destaque aos objetivos estabelecidos para o programa.

O segundo uso, denominado *conceitual*, é aplicado frequentemente aos que não têm contato direto com a formulação do programa (técnicos locais); trata-se, pois, de fornecer informações a respeito da política ou programa, apresentar sua natureza e operação, bem como introduzi-lo aos impactos e resultados conquistados (FARIA, 2005). Pode-se afirmar que o principal objetivo do uso conceitual é a aprendizagem proporcionada aos atores; a esse respeito, Mokate (2002, p. 127, tradução nossa) reforça que, “em resumo, a avaliação permite enriquecer os processos de gestão com uma aprendizagem dinâmica. Longe de ser um monstro diabólico, então, a avaliação pode ter uma nova imagem, a de um aliado da guarda”. Complementarmente, Ramos e Schabbach (2012) indicam que a aprendizagem pode ser potencializada quando os técnicos participam de modo ativo no processo de avaliação.

Em continuidade, o terceiro uso, denominado *persuasivo*, descrito por Faria (2005, p. 103), “dá-se quando ela é utilizada para mobilizar o apoio para a posição que os tomadores de decisão já têm sobre as mudanças necessárias na política ou programa”. A avaliação, nesse momento, tem a função de legitimar as mudanças desejadas e conquistar novos apoiadores para esse movimento. Por fim, o uso de esclarecimento é proposto como decorrente do acúmulo de conhecimento oriundo de várias avaliações já realizadas (FARIA, 2005) e torna-se capaz de pautar a agenda governamental. Mokate (2002) afirma que a maturação de um novo paradigma de gestão, voltado a práticas avaliativas, pode ser um aliado na incorporação da avaliação como processo natural e gradual para a consecução dos objetivos propostos.

Para embasar a diferenciação entre os vários momentos da avaliação, Cohen e Franco (2008) a dividem em: avaliação *ex ante* e avaliação *ex post*.

As avaliações *ex ante* são as realizadas antes da tomada de decisão e pode-se dizer que têm por finalidade auxiliar na tomada de decisão quanto à implementação de políticas ou programas sociais (COTTA, 2001; COHEN e FRANCO, 2008). Ademais, Ramos e Schabbach (2012) afirmam que, nesse tipo de avaliação, gera-se um diagnóstico sobre a situação observada, contribuindo com a adequação dos recursos disponíveis aos objetivos propostos (verificação de viabilidade), também podendo servir como ferramenta de direção para a manutenção e/ou formulação das políticas ou programas.

Já a classificação *ex post* indica que as avaliações são realizadas durante a execução de um projeto ou ao seu final. Quando a avaliação se refere ao momento intermediário (avaliação de processos), as conclusões podem basear decisões de continuidade ou descontinuidade de determinada política ou programa (COTTA, 2001; COHEN e FRANCO, 2008; TREVISAN e BELLEN, 2008; RAMOS e SCHABBACH, 2012). Já em relação às avaliações terminais (que medem impacto e resultado, por exemplo), as avaliações *ex post* fornecem informações concernentes à experiência alcançada e podem subsidiar futuras decisões sobre a mesma política ou programa ou, ainda, contribuir com a construção de um modelo avaliativo para outro objetivo ou experiências semelhantes (COTTA, 2001; COHEN e FRANCO, 2008; TREVISAN e BELLEN, 2008; RAMOS e SCHABBACH, 2012). Com base na literatura, pode-se inferir que as avaliações *ex post* são mais difundidas tanto no meio acadêmico quanto em aplicações práticas; assim, apresentam metodologias mais sedimentadas e sofisticadas.

Do mesmo modo, Scriven (1991) define dois termos para os momentos de avaliação: *avaliação formativa* e *avaliação somativa*.

A avaliação formativa se assemelha à função da avaliação *ex post* intermediária, sendo conduzida durante o processo de implementação ou desenvolvimento do programa, e tem como desígnio fornecer uma leitura analítica aos tomadores de decisão.

Já a avaliação somativa é dirigida após a conclusão do programa (*ex post* terminal), proporcionando aos gestores e usuários um julgamento do valor geral do programa (SCRIVEN, 1991; TREVISAN e BELLEN, 2008). Sob o mesmo ponto de vista, Ramos e Schabbach argumentam que a classificação das avaliações feitas por Scriven (1991) também apresenta conexões com a natureza de uma avaliação, podendo ser de formação de um programa (avaliações formativas) ou para a análise e produção de informações (avaliação somativa). Weiss (1998) argumenta que a simplicidade proposta na definição de Scriven (1991) não diminui o alcance dos propósitos avaliativos, mas produz informações que contribuem com a modificação ou manutenção do programa.

No tocante à classificação da avaliação sob a ótica de quem a realiza ou dela participa, pode-se afirmar que há consenso na literatura quanto à distinção em quatro tipos: avaliação externa; avaliação interna; avaliação mista; e avaliação participativa. Primeiro, Cohen e Franco (2008) alertam que avaliações realizadas por pessoas alheias à organização tendem a enfatizar o método aplicado em detrimento do conhecimento específico à área em que a política ou programa se desenvolve. Cotta (1998) contribui com a discussão informando que avaliações externas agregam isenção à prática avaliativa, porém, o avaliador externo tem maiores dificuldades de acesso aos dados. Em complementação, pode-se citar como vantagem a maior objetividade dos avaliadores externos, por não participarem dos processos internos, além da possibilidade de comparação com programas semelhantes (COHEN e FRANCO, 2008; RAMOS e SCHABBACH, 2012). Na visão de Thoenig (2000), em seu artigo sobre reforma no setor público, as avaliações externas podem desempenhar um papel crucial na tomada de decisão; o autor cita o caso da Grécia, que realizou avaliações externas para a tomada de decisão em relação aos fundos estruturais da União Europeia.

Em referência às avaliações internas, as vantagens apresentadas pela literatura versam sobre: a diminuição dos conflitos gerados pela inserção de pessoas ou organizações externas; a possibilidade de revisão e aprendizagem quanto aos processos executados em determinada política ou programa; o maior conhecimento teórico e técnico das especificidades avaliadas. Todavia, as avaliações internas se arriscam a perder objetividade e conclusões parciais, já que são desenvolvidas pelos envolvidos na formulação e implementação do programa analisado (WEISS, 1998; COTTA, 1998; COHEN e FRANCO, 2008; VEDUNG, 2009; RAMOS e SCHABBACH, 2012).

Em seguida, tem-se a avaliação mista, que combina os dois modelos descritos. Basicamente, o objetivo da avaliação mista é somar esforços e garantir a sinergia entre os atores envolvidos na ação. Combinam-se, então, as potencialidades de cada um dos modelos e diminuem os impactos ou vieses apresentados individualmente (COHEN e FRANCO, 2008; RAMOS e SCHABBACH, 2012).

Por fim, apresenta-se a avaliação participativa, que tem por principal objetivo diminuir a distância existente entre os organizadores e produtores da avaliação e os beneficiários das políticas ou programas avaliados. Permite-se, nesse modelo, a participação dos usuários em todas as fases do ciclo político, favorecendo a opinião e atendendo às demandas específicas desses atores (COHEN e FRANCO, 2008; RAMOS e SCHABBACH, 2012). Para Patton (1999), a avaliação participativa deve considerar as prioridades e os efeitos e estabelecer as definições comuns tanto para avaliadores quanto para beneficiários.

O autor descreve alguns passos para tornar a avaliação participativa viável, como: realização de sessões interativas para discussão dos aspectos da avaliação; discussão e modo de utilização das prioridades, finalidades e definições; e proposta de modelo avaliativo ideal que supra todas as condições expostas nas etapas anteriores (PATTON, 1999). Em complemento, Spink (2001) expõe a importância da avaliação participativa, já destacada, capaz de fortalecer os laços entre os envolvidos, colaborando com a manutenção da prática avaliativa democrática. No entanto, deve-se levar em conta a dimensão da política ou programa avaliado; a esse respeito, Ramos e Schabbach (2012) sugerem que as avaliações participativas funcionam de modo adequado apenas em pequenos projetos ou programas locais. Por sua vez, Carvalho (1999, p. 89) argumenta que:

De fato, tratando de unidades mais delimitadas (uma instituição ou um programa), a avaliação participativa torna-se procedimento rico, dado que sua realização é partilhada com os agentes e beneficiários envolvidos (no programa ou instituição), permitindo para além da avaliação uma apropriação reflexiva e socializada entre os diversos sujeitos da ação em movimento.

Salienta-se, portanto, que a avaliação participativa promove a incorporação dos sujeitos na tomada de decisão e favorece a aprendizagem social e a cidadania (CARVALHO, 1999). Ceneviva e Farah (2012) destacam que os burocratas não devem trabalhar de modo isolado, pelo contrário, a avaliação é uma ferramenta de transparência dos atos públicos e os cidadãos podem e devem informar-se e fiscalizar as ações promovidas pelos atores políticos.

No tocante à abordagem, as avaliações podem ser classificadas em qualitativas, quantitativas e de métodos mistos. Sinteticamente, Weiss (1998) classifica pesquisas qualitativas como aquelas que se apresentam em palavras e as quantitativas, em números. O autor complementa afirmando que avaliações quantitativas coletam dados e transformam-nos em informações numéricas, fortemente apoiadas por estatística, métodos matemáticos e econométricos. As abordagens quantitativas podem utilizar ferramentas sofisticadas na análise de dados, que identificam relações entre uma ou mais variáveis, e podem englobar um universo amostral de grande porte, além de fornecer informações sobre a extensão e distribuição de um fenômeno (WEISS, 1998). Por outro lado, as avaliações qualitativas tendem ao uso de ferramentas com menor estruturação, tais como entrevistas, técnicas observacionais e análise documental (WEISS, 1998). Garcia (2001, p. 29) sustenta que:

A avaliação tem que ser trabalhada com visão ampla, orientada por um julgamento de valor, algo eminentemente qualitativo, focalizada sobre processos complexos, em que os elementos em interação nem sempre produzem manifestações mensuráveis, podendo, inclusive, alguns desses elementos, não apresentar atributos quantificáveis.

Rossi e Wright (1984) afirmam que, para pesquisas sociais, como é o caso do estudo das políticas públicas, a abordagem qualitativa apresenta vantagens, pois tem a capacidade de promover aproximação com a realidade e é flexível para contemplar a complexidade que envolve tais políticas. Mokate (2002) faz alusão às grandes discussões existentes nos setores sociais que colocam as abordagens qualitativas e quantitativas em diferentes lados de uma mesma arena. Trata-se de uma competição exagerada sobre qual método é o mais eficiente, eficaz ou efetivo. Atribuem-se a essas diferenças os motivos de, em parte, as avaliações não serem institucionalizadas como rotina da prática administrativa.

Em relação aos resultados, grande parte da literatura nacional e internacional discute as vantagens e desvantagens de cada abordagem, porém, há crescente aceitação de que a integração entre elas é melhor saída para os embates teóricos (SILVEIRA, VIEIRA, CAPOBIANGO et al., 2013). As diferenças entre as abordagens geram oportunidades de complementação, agregando as similaridades e aumentando a sinergia entre as partes (MOKATE, 2002). É nessa seara que se insere a abordagem de métodos mistos, que nada mais é do que a combinação entre avaliações qualitativas e quantitativas. Mokate (2002, p. 114, tradução nossa) argumenta que, “em particular, a avaliação pode ser aplicada a métodos quantitativos e qualitativos também para gerar tipos diferentes informação que, em conjunto, correspondem às diferentes questões propostas pelo processo de avaliação”.

Scriven (1996) estabelece as competências de um avaliador e assinala a utilização de métodos mistos, argumentando que o avaliador deve apresentar domínios qualitativos básicos, como utilização de questionário e métodos observacionais. Também

é necessário o controle da polarização entre as abordagens, para combinar tais testes práticos, procedimentos de medição, avaliação de julgamento e narrativa e técnicas de estudo de caso. Por conseguinte, Bloch, Soresen, Graversen et al. (2014) afirmam que o uso de métodos mistos colabora com o emprego das melhores fontes de dados disponíveis, com o intuito de fornecer respostas abrangentes e robustas. Contudo, os autores apresentam como limitações para o uso de métodos mistos o tempo, os recursos e o compromisso de edificar os aspectos pragmáticos da avaliação (BLOCH, SORENSEN, GRAVERSEN et al., 2014). Costa, Pegado, Ávila et al. (2013) acrescentam que a pesquisa quantitativa leva a uma generalização de resultados conformada por informações padronizadas. Em contrapartida, a pesquisa qualitativa provavelmente fornecerá dados relativos a circunstâncias e ambientes sociais, creditando destaque às dimensões culturais e contextuais.

Considerando a metodologia aplicada à avaliação, Costa e Castanhar (2003) agrupam as avaliações em: avaliação de processo; avaliação de metas; e avaliação de impacto. Quanto à avaliação de processo, segundo Costa e Castanhar (2003, p. 980):

Seu objetivo é detectar possíveis defeitos na elaboração dos procedimentos, identificar barreiras e obstáculos à sua implementação e gerar dados importantes para sua reprogramação, através do registro de eventos e de atividades. Assim, o uso adequado das informações produzidas durante o desenvolvimento do programa permite introduzir mudanças no seu conteúdo durante a própria execução.

Com propósitos semelhantes aos da avaliação formativa descrita por Scriven (1991), a avaliação de processos averigua de modo sistematizado a evolução das políticas ou programas, bem como acompanha seus processos internos. Seus usos permitem contemplar o conteúdo do programa, a relação entre o que foi planejado e o que vem sendo executado, o alcance dos usuários-alvo e o sucesso ou fracasso na entrega correta dos benefícios (COSTA e CASTANHAR, 2003; RAMOS e SCHABBACH, 2012).

Em prosseguimento, tem-se a avaliação de metas, descrita por Costa e Castanhar (2003) como a mais tradicional entre as metodologias de avaliação. Ela visa a “medir o grau de êxito que um programa obtém com relação ao alcance de metas previamente estabelecidas” (COSTA e CASTANHAR, 2003, p. 979). Pode-se considerar esse tipo de avaliação como uma avaliação *ex post*, pois requer que o programa tenha sido concluído para avaliá-lo (COSTA e CASTANHAR, 2003). Ramos e Schabbach (2012) apresentam uma conceituação diferente, denominada *avaliação de resultados*, porém, com objetivos similares à avaliação de metas; nesse contexto, medem-se os efeitos e as consequências de dada política, determinando seu sucesso ou fracasso com base em mudanças efetivas nas populações beneficiadas.

A avaliação, uma vez considerada como forma de mensurar o desempenho de políticas ou programas, sendo tal desempenho o alcance de metas (COSTA e CASTANHAR, 2003) ou a avaliação dos resultados (RAMOS e SCHABBACH, 2012), necessita de critérios para se estruturar e, por fim, mensurar a consequência contraída (COSTA e CASTANHAR, 2003). Desse modo, prevalecem três dimensões básicas para mensurar os resultados obtidos (DRAIBE, 2001; COSTA e CASTANHAR, 2003; ALVES e PASSADOR, 2011):

- **Eficiência:** de origem econômica, significa atingir os objetivos do programa, priorizando as normas estabelecidas, com a menor relação custo-benefício possível;
- **Eficácia:** mede o grau em que as metas e os objetivos foram alcançados, traduzindo de modo simplificado o resultado atingido;
- **Efetividade:** também tratado na literatura como medida de impacto, indica os efeitos positivos relacionados ao público-alvo do programa. É uma dimensão mais ampla, pois analisa os aspectos econômicos, socioculturais, ambientais e institucionais, ou seja, a efetividade mede tanto a quantidade quanto a qualidade das metas alcançadas pelo programa.

No encadeamento de ideias, Costa e Castanhar (2003) e Ramos e Schabbach (2012) concordam na última definição proposta: a *avaliação de impacto*. Infere-se, portanto, que avaliação de impacto abrange não apenas os resultados alcançados em termos de eficiência, efetividade e eficácia, mas também as alterações ocorridas na população-alvo em decorrência da implantação da política ou programa avaliado. Cotta (1998, p. 113) acrescenta que:

A diferença entre a avaliação de resultados e avaliação de impacto, portanto, depende eminentemente do escopo da análise: se o objetivo é inquirir [acerca dos] efeitos de uma intervenção sobre a clientela atendida, então, trata-se de uma avaliação de resultados; se a intenção é captar os reflexos desta mesma intervenção em contexto mais amplo, então, trata-se de uma avaliação de impacto. Ou, dito de outra forma, a avaliação de resultados visa [a] aferir os resultados intermediários da intervenção, e a avaliação de impacto, seus resultados finais.

Pode-se relacionar a avaliação de impacto com a avaliação somativa proposta por Scriven (1991), que é realizada *ex post* à realização do programa ou projeto. Em síntese, esse modelo avaliativo recorre a estruturas metodológicas que forneçam relações de causa e efeito entre as ações desempenhadas pelo programa e os efeitos promovidos na comunidade-alvo (COSTA e CASTANHAR, 2003; RAMOS e SCHABBACH, 2012). Para tanto, é necessário o acompanhamento de duas diretrizes apontadas por Cotta (1998): a) os objetivos devem ser definidos para comportar a identificação de metas mensuráveis; e b) a implementação da política ou programa deve ser no mínimo satisfatória, para suportar uma avaliação de impacto sem vieses ou até inviabilizá-la.

Vale ressaltar, ainda, a importância dos desenhos e desfechos investigativos para cada um dos tipos de avaliação: de processo, metas/resultados ou de impactos, em função do problema que ela responde. Diante dessa necessidade, o Quadro 2 descreve, sem a pretensão de esgotar, cada um dos desenhos e desfechos apresentados na literatura (SCRIVEN, 1991; WEISS, 1998; COTTA, 1998; OECD, 2002; COSTA e CASTANHAR, 2003; VEDUNG, 2009; CAMELO JUNIOR, FERNANDES, JORGE et al., 2011; RAMOS e SCHABBACH, 2012; SILVEIRA, VIEIRA, CAPOBIANGO et al., 2013).

Quadro 2

Descrição dos desenhos e desfechos investigativos

Desenho/desfecho	Breve descrição	Tipo de avaliação
Custo-benefício	Compara os benefícios e custos de um programa/projeto e busca a melhor relação de benefícios.	Avaliação de processo; Avaliação de meta/resultados.
Custo-efetividade	Comparação entre formas alternativas da ação para ser selecionado aquele programa/projeto que atenda aos objetivos com o menor custo.	Avaliação de processo; Avaliação de meta/resultados.
Custo-utilidade	Compara não só os custos e benefícios, mas a percepção de qualidade de vida ou impacto social da população-alvo.	Avaliação de impacto.
Estudos experimentais	Implicam a conformação de duas populações: o grupo com o programa/projeto (tratamento) e outro grupo sem o programa/projeto (controle). O objetivo é verificar as mudanças geradas pelo programa/projeto.	Avaliação de impacto.
Estudos quase experimentais	Similares ao modelo experimental, porém, os grupos não são selecionados aleatoriamente, trata-se da escolha entre os pertencentes à população-alvo que se encaixam em um dos grupos (tratamento ou controle). O objetivo também é verificar as mudanças ocorridas nessas populações.	Avaliação de impacto. Avaliação de meta/resultados.
Métodos documentais	Método de recolha de dados envolvendo a utilização das informações para outros fins, tais como estatísticas ou dados produzidos pelo próprio órgão e/ou outros documentos escritos.	Avaliação de processo;
Métodos interrogativos	Estratégia de acumular dados por meio do questionamento de pessoas, apropriados com a realização de entrevistas, questionários e afins.	Avaliação de processo; Avaliação de meta/resultados; Avaliação de impacto.
Métodos observacionais	Estratégia de coleta de dados no qual o investigador, por meio de visita ao local, faz observações com seus próprios olhos. Em observação direta, a participação no processo a ser julgado é um pretexto para permitir a recolha de dados imparciais.	Avaliação de processo; Avaliação de meta/resultados; Avaliação de impacto.
Meta-avaliação	O termo é usado para as avaliações concebidas que agregam resultados de uma série de avaliações. Também pode ser usada para denotar a avaliação de uma avaliação, para avaliar sua qualidade e/ou avaliar o desempenho dos avaliadores.	Avaliação de meta/resultados; Avaliação de impacto.
Triangulação	É o uso de várias medidas ou instrumentos de medição diferentes para obter informações sobre a mesma ação, condição ou coisa.	Avaliação de processo; Avaliação de meta/resultados; Avaliação de impacto.

Fonte: Elaborado pelas autoras.

Em conclusão, diante da variedade e complexidade das aplicações, métodos, ferramentas, abordagens, momentos e atores envolvidos no processo de avaliação, a próxima seção explora as tendências internacionais acerca da evolução dos estudos de avaliação, com o intuito de relacioná-los sob o ponto de vista dos tipos, metodologias e desfechos.

TENDÊNCIAS NACIONAIS E INTERNACIONAIS RECENTES

Diante de cenários cada vez mais complexos e turbulentos, as tendências mundiais convergem para a estruturação e o fortalecimento de ações internacionais de avaliação, apoiadas por órgãos governamentais, agências independentes, organizações multilaterais e organizações da sociedade civil. É claro o entendimento das partes que atividades avaliativas são importantes para a boa governação e podem contribuir com o avanço de políticas públicas sociais mais efetivas, eficazes e eficientes (RAMOS e SCHABBACH, 2012; EVALPARTNERS, 2015). Para tanto, este tópico mapeia, sem a pretensão de esgotamento, as ações desenvolvidas e apresenta ações levadas a cabo por organismos internacionais.

As organizações internacionais, como Banco Mundial, ONU e OCDE, participam e desempenham um papel fundamental na construção e no fortalecimento de práticas avaliativas institucionalizadas. Tem-se como ação o desenvolvimento de arcabouços metodológicos e ferramentas destinadas a aprimorar os desenhos, a execução e a avaliação de programas ou projetos ao redor do mundo (RAMOS e SCHABBACH, 2012).

Nesse contexto, 2015 foi estabelecido pela EvalPartners e aprovado pela ONU como o ano internacional da avaliação. A organização, construída por meio de parceiras entre a Organização Internacional para a Cooperação em Avaliação (IOCE) e o Fundo das Nações Unidas para a Infância (Unicef), em parceria com várias organizações importantes, tem como meta a criação de uma rede mundial de avaliadores voluntários que possam contribuir com as avaliações de políticas públicas em seus países (EVALPARTNERS, 2016). De acordo com as diretrizes da organização, a iniciativa visa a contribuir com o melhoramento das capacidades das organizações da sociedade civil (OSC), bem como envolvê-las de modo estratégico e expressivo nos processos de avaliação nacionais, com o intuito de desenvolver sistemas de avaliação e garantir a equidade nas ações (EVALPARTNERS, 2016). Para isso, desenvolveram a Agenda Global de Avaliação (ou EvalAgenda2020). O documento, aprovado pelos delegados do Fórum Global de Avaliação, realizado no Nepal, em 2015, viabiliza um roteiro de ações para o futuro da avaliação e o fortalecimento dos avaliadores de diferentes países e organizações (EVALPARTNERS, 2016).

A União Europeia também desenvolveu um projeto para avaliação das políticas estruturais do bloco, bem como a identificação e utilização de práticas inovadoras de metodologias de avaliação, além de envolvimento multissetorial, com o intuito de reduzir a pobreza (SOPHIE PROJECT, 2016a).

No contexto da América Latina, a Rede de Monitoramento, Avaliação e Sistematização da América Latina e do Caribe (ReLAC), em conjunto com o Projeto Fomento das Capacidades em Avaliação (Foceval), do Ministério de Planejamento e Política Econômica da Costa Rica (Mideplan), e o Instituto Alemão de Avaliação da Cooperação para o Desenvolvimento (DEval) realizaram uma série de consultas públicas a especialistas e profissionais de avaliação durante os anos de 2014 e 2015, e apresentaram os primeiros esboços na IV Conferência ReLAC, no Peru, em 2015. Recentemente, em agosto de 2016, lançaram um documento com o resumo dessa proposta intitulado *Diretrizes para avaliação para a América Latina e o Caribe*, cujo objetivo é enriquecer o marco comum de orientação para o alcance de avaliações com padrão superior de qualidade, promover a formação teórica e prática dos avaliadores na região e instigar a cultura da avaliação no continente (BILELLA, VALENCIA, ALVAREZ et al., 2016).

Constata-se a atuação do Centro de Aprendizagem em Avaliação e Resultados para o Brasil e a África Lusófona (FGV EESP-Clear), com sede na Fundação Getulio Vargas (FGV) em São Paulo, que se beneficia do ambiente multidisciplinar da FGV e é um dos seis centros regionais que compõem a iniciativa Clear – que envolve diversos países e visa a aperfeiçoar políticas e programas com o desenvolvimento de capacidades e sistemas de monitoramento e avaliação. Dentre os eixos de pesquisa de monitoramento e avaliação, sobressaem-se: a) capacitação; b) assistência técnica; c) geração de novas evidências; e d) difusão de evidências e conhecimentos (FGV CLEAR, 2018).

Ainda no Brasil, outra iniciativa ganha notoriedade entre a comunidade científica: o Centro de Integração de Dados e Conhecimentos para Saúde (Cidacs), criado em 2016 e vinculado à Fundação Oswaldo Cruz da Bahia (Fiocruz Bahia), que promove pesquisas e projetos interdisciplinares com base no processamento, na integração e na análise de grandes volumes de dados (*big data*). Com isso, almeja-se contribuir com a utilização de metodologias inovadoras, proporcionar a capacitação profissional e científica e apoiar os tomadores de decisão em relação às políticas públicas sociais, especialmente as determinantes de saúde e as condições ambientais (CIDACS, 2018).

Vale ressaltar, ainda, que há um esforço global voltado à construção de uma agenda baseada na integração entre todos os organismos internacionais, guiada principalmente pela consecução dos 17 Objetivos de Desenvolvimento Sustentável (ODS) do Programa das Nações Unidas para o Desenvolvimento (PNUD, 2016). Os primeiros cinco anos vislumbrados pela EvalAgenda2020 tem como prioridade a criação e o fortalecimento de mecanismos para avaliação dos ODS. O Quadro 3 sintetiza as principais ações (PNUD, 2016; ONU, 2016; EVALPARTNERS, 2015; OECD, 2016; SOPHIE PROJECT, 2016b; BILELLA, VALENCIA, ALVAREZ et al., 2016; CLEAR, 2018; CIDACS, 2018; IBGE, 2018).

Quadro 3

Ações de estruturação de indicadores e avaliação desenvolvidas por organismos nacionais e internacionais

Organização	Ação desenvolvida	Breve descrição
ONU (2015)	Transformando nosso mundo: a Agenda 2030 para o desenvolvimento sustentável.	Documento que estabelece os 17 Objetivos e as 169 metas do Desenvolvimento Sustentável aprovados na Assembleia Geral da ONU em 2015.
ONU (2016)	Primeiro relatório de acompanhamento dos ODS.	Relatório inaugural que dá um panorama dos 17 ODS, utilizando dados disponíveis para enfatizar alguns pontos críticos e desafios.
ONU (PNUD, 2016)	Acompanhamento da Agenda 2030.	Documento com subsídios iniciais para a identificação dos indicadores nacionais referentes aos ODS, cujo objetivo é apresentar os índices disponíveis para acompanhamento e avaliação do cumprimento das metas estabelecidas.
ODS no Instituto Brasileiro de Geografia e Estatística (IBGE)	Plataforma Digital ODS.	Desenvolvida por equipes da Diretoria de Informática do IBGE, fornece um sistema de informações para o acompanhamento da Agenda 2030 no Brasil. Constitui um primeiro conjunto de indicadores globais, elaborado de modo colaborativo entre instituições globais, com a finalidade de apresentar os dados nacionais produzidos de maneira regular, a metodologia e os padrões internacionalmente estabelecidos.
EvalPartners (2015)	EvalAgenda2020.	Estabelece prioridades e oferece treinamentos e fóruns, objetivando criar um ambiente propício às avaliações e tornar viável o alcance das metas estabelecidas na Agenda 2030.
OECD (2016)	Measuring distance to the SDGs targets: a pilot assessment of where OECD countries stand.	Estudo-piloto das posições de partida, em relação aos ODS, de alguns países da OCDE. Tem por objetivo ajudar os governos a desenvolver planos de ação nacionais. Utiliza a riqueza de dados estatísticos obtidos por membros da OCDE, harmonizada por meio de ferramentas e processos da organização.
SOPHIE PROJECT (2016b)	Evaluating the impact of structural policies on health inequalities and their social determinants, and fostering change (SOPHIE).	Desigualdades na saúde são fortemente afetadas pelas circunstâncias em que as pessoas nascem, crescem, vivem, trabalham e envelhecem, assim como pelas políticas que influenciam essas circunstâncias. O SOPHIE Project tem por objetivo gerar novas evidências acerca do impacto das políticas estruturais sobre as desigualdades na saúde e desenvolver metodologias inovadoras para a avaliação dessas políticas na Europa. É coordenado pela Agência de Saúde Pública de Barcelona e financiado pelo Fundo de Pesquisa da União Europeia.
ReLAC	Diretrizes para Avaliação para a América Latina e o Caribe.	Instrumento para fomentar a cultura da avaliação nas instituições públicas e privadas, orientando a prática e a formação dos avaliadores, garantindo a credibilidade, a transparência e o caráter integral da avaliação (BILELLA, VALENCIA, ALVAREZ et al., 2016). As diretrizes foram divididas em 5 dimensões de análise: a) avaliação rigorosa; b) avaliabilidade adequada; c) avaliação conduzida de acordo com princípios éticos e jurídicos; d) compreensão cultural adequada; e e) relevância e utilidade.
FGV Clear	Desenvolvimento de projetos variados em conjunto com agentes governamentais, instituições privadas, sociedade civil e academia.	Os projetos se destinam à formulação e/ou ao aperfeiçoamento da gestão de políticas públicas e programas sociais a partir de resultados em 4 eixos de atuação: a) diagnóstico; b) monitoramento: sistemas de monitoramento, construção de indicadores; c) avaliação: marco lógico, avaliação de processo, análise de custo-benefício, avaliação de impacto; e d) capacitação institucional.
Cidacs – Fundação Oswaldo Cruz da Bahia (Fiocruz Bahia)	Desenvolvimento de estudos com grandes volumes de informações, integrando bases de dados nacionais e disponibilizando produtos (bancos de dados) para pesquisas.	O Cidacs atua com as seguintes plataformas de pesquisa: Coorte de 100 Milhões de Brasileiros; Plataforma Zika; Tecnologias e Inovações para o Sistema Único de Saúde (SUS); Equidade e Sustentabilidade Urbana; Bioinformática e Epidemiologia Genética (Epigen). Para esse estudo, ressalta-se a plataforma Coorte de 100 Milhões de Brasileiros, constituída por 114 milhões de indivíduos pertencentes a famílias de baixa renda (com renda <i>per capita</i> de menos de 3 salários-mínimos) elegíveis para programas sociais do Governo Federal. Foram vinculadas, para essa população, informações socioeconômicas, demográficas, de moradia, saneamento, remuneração e trabalho, despesas, escolaridade, composição familiar, nascimento e mortalidade constantes nos bancos de dados do Cadastro Único (CadÚnico), dos sistemas informacionais de saúde, do sistema sobre condicionalidades do Programa Bolsa Família e do sistema de informações do Programa Minha Casa Minha Vida (CIDACS, 2018).

Fonte: Elaborado pelas autoras.

É claro o entendimento de que atividades avaliativas são importantes para a boa governação e podem contribuir com o avanço de políticas públicas mais efetivas, eficazes e eficientes (RAMOS e SCHABBACH, 2012; EVALPARTNERS, 2015). Cabe destaque, como se pode deduzir, no meio acadêmico e nas organizações governamentais de pesquisa, para o fortalecimento do estudo da avaliação para buscar as melhores práticas e teorias, contribuindo sobremaneira com a criação de uma “cultura” avaliativa capaz de abarcar a complexidade e as especificidades de cada política ou programa em questão (CARVALHO, 1999; PNUD, 2016).

Com base no referencial teórico construído e nas avaliações realizadas pelos organismos multilaterais, as principais lacunas teóricas observadas são: a) estudos que envolvam múltiplos países; b) utilização da abordagem de métodos mistos; c) estudos avaliativos que combinam duas ou mais políticas públicas; d) fortalecimento, capacitação e estruturação das práticas avaliativas nos países desenvolvidos e em desenvolvimento; e) utilização dos resultados das avaliações para informar e melhorar as políticas públicas; e f) avaliações participativas – contemplação de todos os atores envolvidos (gestores, tomadores de decisão, formuladores, implementadores, executores e beneficiários).

Diante dos esforços apresentados, fica evidente a importância de estudos avaliativos robustos, imparciais e que contemplem abordagens multimétodos. De acordo com o exposto, observa-se que os países desenvolvidos já incorporaram na agenda política a avaliação como processo, diferente do que ocorre no Brasil e em outros países em desenvolvimento, onde a avaliação de políticas e programas públicos ainda sofre com descontinuidade e indisponibilidade de dados confiáveis e completos, além de sistemas informacionais integrados que facilitem o acesso à informação, o que contribuiria sobremaneira com a tomada de decisão.

CONSIDERAÇÕES FINAIS

Em consonância com a busca de modelos articulados para o desenvolvimento de estratégias governamentais de combate à pobreza e a promoção da cidadania no Brasil, diversos organismos internacionais vêm desenvolvendo propostas integrativas e intersetoriais, com o intuito de melhorar a qualidade de vida da população mundial. Ações como o estabelecimento dos ODS do PNUD e estudos comparativos e compartilhados entre os países pertencentes à OCDE constituem exemplos de tais esforços.

Na América Latina, percebe-se uma busca crescente por ações avaliativas para o fortalecimento e a melhoria das políticas públicas adotadas na região, principalmente relacionadas ao enfrentamento das vulnerabilidades sociais. Entretanto, o cenário demanda desenvolvimento de capacidades profissionais dos governos e aumento da atuação da sociedade civil no acompanhamento e controle. A evolução e a internalização de processos avaliativos contribuem com o desenvolvimento de metodologias robustas e abrangentes. Observa-se, a partir da década de 1980, uma busca crescente por modelos capazes de conjugar informações quantitativas e qualitativas (métodos mistos), com vistas a aperfeiçoar e tornar a avaliação mais realista no que se refere à complexidade das políticas analisadas. A busca por novas metodologias de avaliação robustas e abrangentes se mostra essencial para o desenvolvimento e a evolução das atividades de avaliação, proporcionando aos gestores dados concretos, reais e aplicáveis. Portanto, almeja-se o uso de metodologias inovadoras, proporcionando capacitação profissional e científica e apoiando os tomadores de decisão em relação às políticas públicas sociais – especialmente as de saúde.

Entretanto, vale destacar que, no caso brasileiro, as bases de dados públicas disponíveis apresentam problemas de confiabilidade, dados indisponíveis e descontinuidade na produção de indicadores; tais fatores podem contribuir com a baixa adesão a estudos avaliativos quantitativos. Assim, as avaliações no Brasil estariam com maior quantidade de avaliações qualitativas de caráter local e com análise de políticas específicas, voltadas principalmente a prestação de contas e medição de desempenho e em busca da qualidade dos serviços prestados. Também há experiências que promovem pesquisas e projetos interdisciplinares com base no processamento, na integração e na análise de grandes volumes de dados (*big data*).

Felizmente, vale destacar que os dados demonstram o fortalecimento do estudo da avaliação no Brasil, para buscar as melhores práticas e teorias, contribuindo sobremaneira com a criação de uma “cultura” avaliativa capaz de abarcar a complexidade e as especificidades de cada política ou programa em questão. Não obstante, também se mostra necessário pontuar a importância de ações avaliativas contínuas, principalmente em cenários com escassez de recursos financeiros, humanos e materiais. Tais avaliações podem contribuir com a destinação adequada dos recursos e o fortalecimento das ações de controle por parte dos órgãos responsáveis.

REFERÊNCIAS

- ALA-HARJA, M.; HELGASON, S. Em direção às melhores práticas de avaliação. **Revista do Serviço Público**, v. 51, n. 4, p. 5-60, 2000.
- ALVES, T.; PASSADOR, C. S. **Educação pública no Brasil: condições de oferta, nível socioeconômico dos alunos e avaliação**. São Paulo: Annablume, 2011.
- ARRETCHE, M. T. S. Tendências no estudo sobre avaliação. In: RICO, E. M. (Org.). **Avaliação de políticas sociais: uma questão de debate**. São Paulo: Cortez, 2001.
- BILELLA, P. D. R. et al. **Diretrizes para avaliação para a América Latina e o Caribe**. Buenos Aires: Akian, 2016.
- BLOCH, C. et al. Developing a methodology to assess the impact of research grant funding: A mixed methods approach. **Evaluation and Program Planning**, v. 43, p. 105-117, 2014.
- BOBBIO, N.; MATTEUCCI, N.; PASQUINO, G. **Dicionário de política**. 1. ed. Brasília: Editora da UnB, 1998.
- CAMELO JUNIOR, J. S. et al. Avaliação econômica em saúde: triagem neonatal da galactosemia. **Cadernos de Saúde Pública**, v. 27, n. 4, p. 666-676, 2011.
- CARVALHO, M. C. B. Avaliação participativa: uma escolha metodológica. In: RICO, E. M. (Org.). **Avaliação de políticas sociais: uma questão de debate**. 2. ed. São Paulo: Cortez, 1999. p. 87-94.
- CENEVIVA, R.; FARAH, M. F. S. Avaliação, informação e responsabilização no setor público. **Revista de Administração Pública**, Rio de Janeiro, v. 46, n. 4, p. 993-1016, 2012.
- CIDACS. Centro de Integração de Dados e Conhecimentos para Saúde. **Quem somos**. 2018. Disponível em: <<https://cidacs.bahia.fiocruz.br/sobre/quem-somos/>>. Acesso em: 18 jul. 2018.
- COHEN, E.; FRANCO, R. **Avaliação de projetos sociais**. 8. ed. Rio de Janeiro: Vozes, 2008.
- COSTA, A. F. et al. Mixed-methods evaluation in complex programmes: the national reading plan in Portugal. **Evaluation and Program Planning**, v. 39, p. 1-9, 2013.
- COSTA, F. L.; CASTANHAR, J. C. Avaliação de programas públicos: desafios conceituais e metodológicos. **Revista de Administração Pública**, Rio de Janeiro, v. 37, n. 5, p. 969-992, 2003.
- COTTA, T. C. Metodologias de avaliação de programas e projetos sociais: análise de resultados e de impacto. **Revista do Serviço Público**, v. 49, n. 2, p. 103-124, 1998.
- COTTA, T. C. Avaliação educacional e políticas públicas: a experiência do Sistema Nacional de Avaliação da Educação Básica (Saeb). **Revista do Serviço Público**, v. 52, n. 4, p. 89-111, 2001.
- CRUMPTON, C. D. et al. Avaliação de políticas públicas no Brasil e nos Estados Unidos: análise da pesquisa nos últimos 10 anos. **Revista de Administração Pública**, Rio de Janeiro, v. 50, n. 6, p. 981-1001, 2016.
- CRUZ, M. M. Avaliação de políticas e programas de saúde: contribuições para o debate. In: MATTOS, R. A.; BAPTISTA, T. W. F. (Org.). **Caminhos para análise das políticas de saúde**. Porto Alegre: Rede Unida, 2015. p. 285-317.
- DERLIEN, H.U. Una comparación internacional en la evaluación de las políticas públicas. **Revista do Serviço Público**, v. 52, n. 1, p. 105-124, 2001.
- DRAIBE, S. M. Avaliação de implementação: esboço de uma metodologia de trabalho em políticas públicas. In: BARREIRA, M. C. R. N.; CARVALHO, M. C. B. (Org.). **Tendências e perspectivas na avaliação de políticas e programas sociais**. São Paulo: IEE/PUC-SP, 2001. p. 13-42.
- EVALPARTNERS. **Global Evaluation Agenda 2016-2020**. 2015. Disponível em: <<http://www.evalpartners.org/sites/default/files/documents/EvalAgenda2020.pdf>>. Acesso em: 15 jul. 2016.
- EVALPARTNERS. **Global evaluation agenda**. [2016]. Disponível em: <<http://www.evalpartners.org/global-evaluation-agenda>>. Acesso em: 27 jul. 2016.
- FARIA, C. A. P. A política da avaliação de políticas públicas. **Revista Brasileira de Ciências Sociais**, v. 20, n. 59, p. 97-109, 2005.
- FGV CLEAR. **Centro de Aprendizagem em Avaliação e Resultados para o Brasil e a África Lusófona (FGV EAESP Clear)**. 2018. Disponível em: <<http://fgvclear.org/pt/sobre-o-fgv-clear/>>. Acesso em: 17 jul. 2018.
- FREY, K. Políticas públicas: um debate conceitual e reflexões referentes à prática da análise de políticas públicas no Brasil. **Planejamento e Políticas Públicas**, n. 21, [n.p.], 2000.
- GARCIA, R. C. Subsídios para organizar avaliações da ação governamental. **Planejamento e Políticas Públicas**, n. 23, p. 7-70, 2001.
- HOWLETT, M.; RAMESH, M.; PERL, A. **Política pública – seus ciclos e subsistemas: uma abordagem integradora**. Rio de Janeiro: Elsevier, 2013.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. **Indicadores ODS**. 2018a. Disponível em: <<https://ods.ibge.gov.br/>>. Acesso em: 25 abr. 2018.
- MOKATE, K. M. Convirtiendo el “monstruo” en aliado: la evaluación como herramienta de la gerencia social. **Revista do Serviço Público**, v. 53, n. 1, p. 89-134, 2002.
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT – OECD. **Glossary of key terms in evaluation and results-based management**. Paris: OECD, 2002.
- ORGANIZAÇÃO DAS NAÇÕES UNIDAS – ONU. **Transformando nosso mundo: a agenda 2030 para o desenvolvimento sustentável**. 2015. Disponível em: <<https://nacoesunidas.org/pos2015/agenda2030/>>. Acesso em: 27 jul. 2016.
- ORGANIZAÇÃO DAS NAÇÕES UNIDAS – ONU. **Global Evaluation Agenda: Eval2016-2020**. [s.l.]: EvalPartners, 2016.
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT – OECD. **Measuring distance to the SDGs targets: a pilot assessment of where OECD countries stand**. [2016]. Disponível em: <<http://www.oecd.org/std/oecd-measuring-distance-to-the%20sdgs-target-pilot-study-web.pdf>>. Acesso em: 27 jul. 2016.
- PATTON, M. Q. **Utilization-focused evaluation in Africa**. 1999. Disponível em: <http://gametlibrary.worldbank.org/FILES/284_Utilisation%20Focused%20evaluation%20in%20Africa.pdf>. Acesso em: 25 jul. 2016.
- PAULA, A. P. P. **Por uma nova gestão pública: limites e potencialidades da experiência contemporânea**. Rio de Janeiro: FGV, 2005.

- PICCIOTTO, R. Democratic evaluation for the 21st century. **Evaluation**, v. 21, n. 2, p. 150-166, 2015.
- PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO – PNUD. **Objetivos de desenvolvimento sustentável**. [2016]. Disponível em: <<http://www.pnud.org.br/ods.aspx>>. Acesso em: 27 jul. 2016.
- RAMOS, M. P.; SCHABBACH, L. M. O estado da arte da avaliação de políticas públicas: conceituação e exemplos de avaliação no Brasil. **Revista de Administração Pública**, Rio de Janeiro, v. 46, n. 5, p. 1271-1294, 2012.
- ROSSI, P. H.; WRIGHT, J. D. Evaluation research: an assessment. **Annual Review of Sociology**, v. 10, p. 331-352, 1984.
- SCRIVEN, M. **Evaluation thesaurus**. 4. ed. Newbury Park: Sage, 1991.
- SCRIVEN, M. Types of evaluation and types of evaluator. **Evaluation Practice**, v. 17, n. 2, p. 151-162, 1996.
- SECCHI, L. **Políticas públicas: conceitos, esquemas de análise, casos práticos**. 2. ed. São Paulo: Cengage Learning, 2014.
- SILVEIRA, S. F. R. et al. Políticas públicas: monitorar e avaliar para quê? In: FERREIRA, M. A. M.; ABRANTES, L. A. (Org.). **Políticas públicas, gestão e sociedade**. Viçosa: Triunfal, 2013. p. 301-327.
- SOPHIE PROJECT. **Conclusions**. [2015]. Disponível em: <<http://www.sophie-project.eu/pdf/conclusions.pdf>>. Acesso em: 08 ago. 2016a.
- SOPHIE PROJECT. **Evaluating the impact of structural policies on health inequalities and their social determinants, and fostering change (SOPHIE)**. [2012]. Disponível em: <<http://www.sophie-project.eu/project.htm>>. Acesso em: 13 set. 2016b.
- SOUZA, C. Políticas públicas: uma revisão da literatura. **Sociologias**, v. 8, n. 16, p. 20-45, 2006.
- SPINK, P. Avaliação democrática: propostas e prática. **Fundamentos de Avaliação**, n. 3, p. 7-25, 2001.
- THOENIG, J. C. A avaliação como conhecimento utilizável para reformas de gestão pública. **Revista do Serviço Público**, v. 51, n. 2, p. 54-71, 2000.
- TREVISAN, A. P.; BELLEN, H. M. V. Avaliação de políticas públicas: uma revisão teórica de um campo em construção. **Revista de Administração Pública**, Rio de Janeiro, v. 42, n. 3, p. 529-550, 2008.
- VEDUNG, E. **Public policy and program evaluation**. New Brunswick: Transaction, 2009.
- VEDUNG, E. Four waves of evaluation diffusion. **Evaluation**, v. 16, n. 3, p. 263-277, 2010.
- WEISS, C. H. Have we learned anything new about the use of evaluation? **American Journal of Evaluation**, v. 19, n. 1, p. 21-33, 1998.
- WEISS, C. H. The interface between evaluation and public policy. **Evaluation**, v. 5, n. 4, p. 468-486, 1999.

Lilian Ribeiro de Oliveira

ORCID: <https://orcid.org/0000-0002-3506-9011>

Doutora em Administração das Organizações na Faculdade de Economia Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo (USP); Membro do Centro de Estudos em Gestão e Políticas Públicas Contemporâneas da Universidade de São Paulo (GPublic); Professora de graduação em Administração na Fundação Educacional de Ituverava (FEI/FFCL) e Centro Universitário Barão de Mauá (CBM), Ribeirão Preto – SP, Brasil. E-mail: lilianroliveira@fearp.usp.br

Claudia Souza Passador

ORCID: <https://orcid.org/0000-0002-9333-563X>

Doutora em Educação pela Universidade de São Paulo; Livre-Docência em Administração Pública pela Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo (USP) (2012); Professora Associada do Departamento de Administração da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo (USP); Coordenadora do Centro de Estudos em Gestão e Políticas Públicas Contemporâneas da Universidade de São Paulo (GPUBLIC/USP); Pesquisadora do Núcleo de Apoio à Pesquisa Centro de Biologia Sistêmica Integrada (NAP/CISBi- FMRP/USP) e do Instituto de Estudos Avançados da Universidade de São Paulo (IEA/USP). E-mail: cspassador@usp.br