

Conceitos e métodos para a avaliação de programas sociais e políticas públicas

Mauro Serapioni
Centro de Estudos Sociais

Resumo

A avaliação de políticas públicas difundiu-se nos anos de 1960 e passa hoje por um processo de rápida disseminação, diversificação teórico-metodológica e controvérsias entre as diversas perspetivas epistemológicas. Este artigo analisa as questões-chave que deveriam nortear as avaliações, apresenta as principais abordagens avaliativas desenvolvidas nos últimos 50 anos (positivista experimental, pragmatista da qualidade e construtivista) e os modelos que resultaram da contaminação entre elas. Na conclusão apresentam-se alguns pontos de tensão persistentes nas discussões sobre avaliação.

Palavras-chave: Avaliação, Políticas públicas, Pluralidade metodológica

Concepts and methods for the assessment of social programs and public policies

Abstracts

The evaluation of public policies has spread in the 1960s and it is going through a process of rapid dissemination, methodological diversification and disputes between different epistemological perspectives. The present article analyses key issues that should guide the evaluation exercises, it presents three evaluation approaches developed in the last 50 years (the experimental positivist, the pragmatist of quality and the constructivist) and the models resulting from their mutual influence. In the conclusion some persistent points of tension in the discussions about evaluation are presented.

Keywords: Evaluation, Public policies, Methodological plurality.

Concepts et méthodes pour la évaluation des programmes sociaux y politiques publiques

Resumé

L'évaluation des politiques publiques est répandue dans les années 1960 et il passe par un processus de diffusion rapide, de diversification méthodologique et de conflits entre les différentes perspectives épistémologiques. Cet article analyse les questions clés qui devraient guider les évaluations et présente trois approches de l'évaluation développée au cours des 50 dernières années (positiviste expérimentale, pragmatique de la qualité et constructiviste) et les modèles résultant de leur influence réciproque. En conclusion sont présentés quelques points de tension persistants dans les discussions sur l'évaluation

Mots-clés: Évaluation, Politiques publiques, Pluralité méthodologique.

Conceptos y métodos para la evaluación de los programas sociales y políticas públicas

Resumen

La evaluación de las políticas públicas se desarrolló en la década de 1960, encontrándose actualmente en un proceso de difusión rápida, de diversificación metodológica y de disputas entre diferentes perspectivas epistemológicas. Este artículo analiza las cuestiones clave que deberían guiar las evaluaciones; posteriormente presenta los tres enfoques de evaluación que se han desarrollado en los últimos 50 años (positivista experimental, pragmatista de la calidad y constructivista), así como los modelos resultantes. Como conclusión se presentan los puntos de tensión existentes en el debate sobre la evaluación.

Palabras clave: Evaluación, Políticas públicas, Pluralidad metodológica.

1. Introdução

O interesse pela investigação em políticas públicas, área também designada como *policy science*, começou na década de 1950, nos Estados Unidos da América (EUA) e, a partir dos anos de 1970, difundiu-se também pela Europa, em particular na Alemanha e nos países escandinavos. Nesse período, a principal preocupação recaiu na definição do conceito de políticas públicas e na análise dos distintos atores intervenientes nesses processos, estatais e não estatais (Faria, 2005). Atualmente, num contexto de crescentes incertezas e complexidade das questões que se colocam neste âmbito, muitas abordagens e teorizações tentam compreender as diversas formas e processos de formação e de gestão das políticas públicas (Trevisan e van Bellen, 2008). Nesse sentido, para melhor analisar a ação política, vários autores têm recorrido à analogia do ciclo político (*policy cycle*), o qual pode ser dividido em diferentes fases, tais como: i) inserção de um tema na agenda política, ii) fase de formulação da política, iii) tomada de decisão, iv) implementação e v) avaliação e controle dos efeitos das políticas públicas (Garcia,

2014; Howlett *et al.* 2009; Frey, 2000). No âmbito deste artigo realça-se a importância da última fase do ciclo político, ou seja, a avaliação dos programas e ações implementadas, no sentido de avaliar o alcance dos objetivos definidos e seus impactos efetivos.

A prática da avaliação de programas e políticas públicas começou a difundir-se nos anos 1960 do século passado e hoje passa por um processo de rápida disseminação e diversificação teórica e metodológica. Entre as várias razões de afirmação da avaliação nos países ocidentais cabe assinalar as seguintes: i) progressiva erosão das bases tradicionais de legitimação dos poderes públicos; ii) crescente exigência, por parte dos cidadãos, de transparência e participação na definição e apreciação da qualidade dos serviços prestados; iii) crise fiscal do Estado, que requer uma maior capacidade de alocar recursos de forma otimizada, tanto do ponto de vista da sua eficiência como da sua efetividade; iv) pressão da União Europeia para a avaliação dos recursos comunitários destinados a áreas importantes de atividade, tais como trabalho, formação profissional, desenvolvimento local, políticas agrícolas, etc.; v) crescente complexidade social que torna sempre mais difícil dominar os fatores relevantes para o sucesso de uma intervenção e para a previsão de resultados positivos.

Hoje em dia existe uma multiplicidade de abordagens avaliativas e diversas controvérsias entre estudiosos que se baseiam em diferentes perspectivas concetuais e epistemológicas. Alguns autores, por exemplo, questionam-se se a avaliação pode ser considerada um setor da ciência ou se se trata apenas de uma aplicação da metodologia da investigação (Cohen e Franco, 1994). Entretanto, há um consenso sobre o fato de que a avaliação utiliza a metodologia da investigação como um meio para determinar em que medida as políticas, os programas e os projetos sociais alcançam os seus objetivos. Porém, a avaliação não se restringe somente aos métodos e instrumentos para conduzir com sucesso um processo de avaliação; ela desenvolveu também uma própria teoria, quer sobre os aspetos a serem avaliados, quer sobre como obter conhecimentos válidos de tais aspetos (Øvretveit, 1998).

Este artigo, após discutir os principais conceitos e definições de avaliação formulados por alguns dos mais reconhecidos estudiosos e depois de apresentar um histórico da avaliação, analisa três questões-chave que deveriam nortear as avaliações das políticas públicas. Em seguida, são ilustradas as três abordagens de avaliação desenvolvidas nos últimos 50 anos (positivista-experimental, pragmatista da qualidade e construtivista) e os modelos que resultaram da contaminação entre elas.

2. Conceitos e definições

A avaliação é uma disciplina ainda jovem, embora se relacione com uma prática muito antiga que atingiu consideráveis níveis de formalização já nas dinastias imperiais da China, Egito e Japão, há mais de mil anos (Scriven, 1991; Hartz, 2009). Entretanto, como disciplina e campo de estudos, ou seja, como área de produção de conhecimento - e não simplesmente como atividade espontânea e não sistematizada - ela surge somente na metade dos anos 60. Scriven (1991: 9-10) tem enfatizado o “status paradoxal da avaliação”, que apesar de ser considerada “o mais importante ingrediente de todas as atividades práticas e intelectuais (...), não tem sido tratada seriamente por qualquer disciplina acadêmica até o ultimo terço do século XX”. Somente a partir do último quarto do século, acrescenta Scriven (1991:11), a avaliação tem sido objeto de interesse e atenção em vários campos, embora com discussões que ainda “não chegam a uma profundidade suficiente para relacionar os diferentes campos e transformá-los numa disciplina”. Neste sentido, há um amplo consenso entre os especialistas deste campo sobre o facto de que a avaliação é uma área ainda em construção concetual e metodológica, que precisa consolidar-se no plano epistemológico, teórico e metodológico (Novaes, 2000)..

Analisando as diversas definições encontradas na literatura internacional, referentes à avaliação de programas e políticas sociais e educacionais, podemos observar um consenso entre os maiores estudiosos sobre alguns aspetos que fundamentam o processo de avaliação. Pelo menos três elementos aproximam a maioria das definições apresentadas: a) a avaliação surge no sentido de formular juízos sobre o valor ou mérito de uma intervenção que visa modificar a realidade social das comunidades; b) é colocada ênfase na sistematicidade e rigorosidade dos procedimentos de recolha de dados que suportam o julgamento do mérito e valor das ações; c) confere-se destaque à avaliação como ferramenta indispensável para a tomada de decisões, já que oferece aos gestores todas as informações necessárias para aprimorar o processo de planeamento e de gestão dos programas, serviços e políticas.

3. Breve histórico da avaliação

Como já foi apontado, a avaliação no sentido de julgar o valor das ações com o propósito de melhorá-las é tão antiga quanto a consciência humana. Porém, a avaliação tal

como a conhecemos hoje começou a desenvolver-se com a revolução científica registada nos séculos XVI e XVII, na qual estabeleceu uma distinção entre factos e valores. A ciência era assim responsável por recolher e utilizar factos e não por julgar valores. A divisão do ato de recolha de informação do ato de julgar essa informação representou a base do desenvolvimento da avaliação entendida como atividade sistemática. Porém, a avaliação só se tornou uma atividade especializada logo após a Segunda Guerra Mundial. Sucessivamente, alguns avaliadores têm desafiado a separação entre o ato de recolher fatos, tradicionalmente da responsabilidade dos avaliadores, e o ato de avaliar (de formular juízos), uma função dos que solicitaram/contrataram o estudo (*users* ou utilizadores dos resultados da avaliação). Como observa Øvretveit (1998), hoje em dia há uma interação entre esses dois atores distintos.

A avaliação de programas sociais, embora iniciada nos EUA, no início do século XX, entre a primeira e a segunda guerra mundial, no campo da educação, somente a partir dos anos 60 foi desenvolvida em larga escala no âmbito do processo de reforma da *Great Society* e da luta contra a pobreza lançada pelos governos de Kennedy e Johnson (Moro, 2009). De facto, neste período, os governos dos EUA investiram enormes recursos para combater o desemprego, a delinquência, a degradação das áreas urbanas e para oferecer serviços públicos na área da saúde e da educação. As expectativas acerca das mudanças sociais provocadas por tais programas levaram o governo a financiar uma série de estudos avaliativos para verificar a efetividade do investimento e, ao mesmo tempo, para introduzir uma mais eficiente distribuição dos recursos. No mesmo período, outros países, tais como Canadá, Suécia e Alemanha do Oeste, compartilharam estas primeiras experiências de análise e de avaliação.

Neste prisma, desenvolveu-se também a avaliação das políticas públicas, em particular no campo da saúde e da assistência social, com a colaboração dos cientistas sociais e de outras unidades académicas. Nesta primeira fase, os avaliadores adotaram abordagens experimentais ou quase-experimentais para aferir os resultados das políticas e dos programas implementados. Naqueles anos, Campbell e Stanley (1966), pioneiros da avaliação de programas sociais, propuseram estudos avaliativos sobre as inovações geradas por tais programas utilizando métodos quase-experimentais que adotavam princípios e condições experimentais sem a randomização e o controle. Entretanto, a partir de 1970, a perspetiva construtivista entrou em conflito com a perspetiva positivista-experimental, propondo a análise qualitativa e o envolvimento dos atores. Nesse contexto, Stake (1980) propôs a avaliação responsável (*responsive*) e Guba e Lincoln (1989) a avaliação

naturalística (*naturalistic*). Patton (1986) tentou domar este conflito metodológico, o qual definiu, aliás, como “dragão metodológico”, propondo o paradigma da escolha (*paradigme of choice*) como estratégia de intermediação entre as duas perspetivas conflitantes. Sucessivamente, durante os anos de 1980 e 1990, outros autores tentaram resolver esta oposição metodológica introduzindo abordagens multimétodos (Greene e Caracelli, 1997).

No início dos anos 1980, a crise petrolífera e a sucessiva grande crise económica dos países ocidentais reduziu a expansão das políticas públicas, colocando em primeiro plano a necessidade de dar prioridade à redução do défice público. Neste contexto, emergiu uma mudança no papel do estado, teorizada pela *New Public Management* e aplicada em Austrália, Nova Zelanda e alguns países europeus, cuja responsabilidade não era mais a de executar, mas a de liderar e orientar. Assim, mudou a finalidade da avaliação que, a partir daquele momento, se tornou um instrumento para racionalizar a despesa pública dos Estados e para nortear os governos e os ministérios na contenção dos gastos (Moro, 2009). De facto, o aumento da pressão sobre os escassos recursos destinados aos programas sociais tem estimulado um crescente interesse pela avaliação económica e pelas diferentes técnicas que permitem aferir a eficiência dos programas e serviços sociais. Neste âmbito, os economistas desenvolveram métodos para a análise de custos e benefícios dos programas públicos. Assim, o número e o tipo de atividades de avaliação e monitoramento aumentaram consideravelmente durante os anos de 1980, por forma a responder a diferentes propósitos: i) respaldar as exigências dos governos no controlo dos gastos; ii) incrementar a responsabilidades dos gestores (*accountability*) e iii) obter maiores informações sobre o impacto das novas tecnologias (Øvretveit, 1998). Porém, as abordagens dos economistas revelaram-se insuficientes para dar conta da complexidade das dimensões não económicas da avaliação de programas e serviços sociais (saúde, educação, assistência social, etc.).

Nos anos de 1990, a avaliação difundiu-se em todos os países da Europa como resultado do impulso fundamental e da pressão exercida pela União Europeia (Fundos Estruturais e Fundos de Desenvolvimento Regional), que implicaram uma grande mobilização de recursos financeiros e humanos e exigiram o monitoramento e a avaliação dos seus resultados.

Nesse período, cresceu também a preocupação pelos aspetos metodológicos da avaliação e a sensibilidade para promover uma perspetiva de avaliação interdisciplinar. As abordagens qualitativas começaram assim a ocupar um espaço mais relevante e tornaram-se métodos imprescindíveis para a avaliação de políticas públicas e serviços sociais, de saúde

e de educação. Neste prisma, o pluralismo metodológico baseado numa conceção integrada e multidisciplinar assumiu crescente importância entre os estudiosos e os avaliadores profissionais.

4. Pressupostos teóricos e metodológicos da avaliação

A avaliação é uma área de conhecimento ainda jovem, o que justifica uma maior existência de desconincidências e desacordos entre as abordagens dos diferentes estudiosos e avaliadores. No entanto, apesar da discordância e dos desentendimentos entre as diferentes escolas e perspectivas sobre a avaliação, nos últimos anos, a reflexão conceitual e metodológica tende a convergir sobre alguns denominadores comuns fundamentados em abordagens mais pragmáticas e menos ideológicas. Analisando a literatura científica mais relevante sobre esta matéria (Glasser, 1972; Cronbach, 1996; Guba e Lincoln, 1989; Scriven, 1991; Cohen e Franco, 1994; Aguilar e Ander Egg, 1995; Stake, 1996; Madaus *et al.*, 1996; Stufflebeam, 1996; Patton, 1997; Chelimsky e Shadish, 1997; Øvretveit, 1998; Weiss, 1998; Rossi *et al.*, 1999; Hartz, 1999; Novaes, 2000; Bezzi, 2003), identificam-se três questões-chave que permeiam as diversas definições sobre a avaliação: 1) atenção conferida às questões metodológicas; 2) preocupação com a finalidade e utilidade da avaliação e com a necessidade de aumentar o seu valor de uso no âmbito dos processos de tomada de decisões; 3) reconhecimento do pluralismo de valores e da importância de incluir diversas perspectivas e grupos de interesses no processo avaliativo. A seguir, identificam-se e analisam-se os três pressupostos que deveriam nortear as avaliações de programas e políticas públicas.

Em relação à tónica nas questões metodológicas, a avaliação caracteriza-se como um campo que utiliza uma ampla gama de ferramentas das ciências sociais (tanto quantitativas como qualitativas) para analisar programas sociais. Nesse sentido, utiliza os mesmos métodos e técnicas geralmente adotadas pela investigação social e apresenta a mesma “rigoriedade” e “debilidade”. Porém, como apontam vários autores, a avaliação é mais problemática que a investigação social, pois não pretende somente compreender a ação social, visando, também, formular um juízo (Glasser, 1972). Para Bezzi (2003: 29), aliás, a avaliação é “um conjunto de atividades que permite expressar um juízo argumentado” sobre políticas e programas e a “argumentação” representa o alicerce sobre o qual é preciso construir um processo de avaliação sistemático, profissional e não-

improvisado ou informal. Argumentar significa, portanto, apresentar todos os elementos a partir dos quais se formulam os juízos, assim como os procedimentos metodológicos através dos quais aqueles elementos foram analisados e interpretados. Nesse sentido, a pesquisa avaliativa torna-se o coração da avaliação.

Ainda em relação à questão metodológica, cabe assinalar que a tradicional polarização entre as diversas abordagens e métodos de aproximação da realidade (construtivismo *versus* realismo, estrutura *versus* sujeito e qualitativo *versus* quantitativo) tem vindo a assumir posições e tons menos radicais (Minayo e Sanches, 1993; Patton, 1999; Serapioni, 2000; Hartz, 2008). Hoje em dia, a literatura sobre métodos de investigação e de avaliação já não é mais caracterizada pela aspereza e pelo estridente conflito entre as duas perspetivas de análise, tal como aconteceu no passado. No âmbito da sociologia, por exemplo, regista-se um crescente reconhecimento de que a distinção entre qualitativo e quantitativo veio gerar uma rutura desnecessária. Para Santos (2003: 26), a sociologia preocupa-se em demasia “com discussões teóricas estêreis como, por exemplo, a relação entre estrutura e ação ou entre a análise macro e a análise micro”. Para o autor, a “verdadeira distinção e relação fundamental a fazer era entre ação conformista e ação rebelde”. Também o sociólogo italiano Ardigò (1988: 288) põe em primeiro plano o “ambivalente ponto de vista do observador científico” ao privilegiar, quer o sistema em prejuízo da pessoa, quer a pessoa sem se preocupar com os aspetos sistémicos. É portanto criticável, observa o autor, tanto a interpretação da vida social baseada exclusivamente nas categorias relativas ao mundo da vida como a análise ‘funcionalista’ que enfatiza somente o sistema social ignorando os “limites do processo de objetivação” (Ardigò, 1988: 288). Nesta mesma linha, são igualmente esclarecedoras as afirmações de Beck (2000: 326): “(...) não se deve jurar lealdade a qualquer determinado ponto de vista ou perspetiva teórica. A decisão de adotar uma abordagem realista ou construtivista, tem para mim mais um aspeto pragmático, o facto de escolher os meios adequados para atingir o objetivo pretendido”.

Gradualmente, foi emergindo o consenso de que o grande desafio é combinar, de forma apropriada, os métodos, as perguntas e as questões empíricas e não defender uma única abordagem metodológica para todos os problemas. Nesse prisma, Patton propõe um paradigma – que ele define como “*paradigm of choice*” – que deve dispor de um amplo repertório de métodos e de técnicas a serem utilizadas na variedade dos problemas. Esse paradigma, acrescenta o autor, reconhece que diferentes métodos são apropriados para diversas situações e propósitos de avaliação (Patton, 1997: 297). Há, portanto, um consenso

(construtivismo *versus* realismo, estrutura *versus* sujeito e qualitativo *versus* quantitativo) tem vindo a assumir posições e tons menos radicais (Minayo e Sanches, 1993; Patton, 1999; Serapioni, 2000; Hartz, 2008). Hoje em dia, a literatura sobre métodos de investigação e de avaliação já não é mais caracterizada pela aspereza e pelo estridente conflito entre as duas perspetivas de análise, tal como aconteceu no passado. No âmbito da sociologia, por exemplo, regista-se um crescente reconhecimento de que a distinção entre qualitativo e quantitativo veio gerar uma rutura desnecessária. Para Santos (2003: 26), a sociologia preocupa-se em demasia “com discussões teóricas estéreis como, por exemplo, a relação entre estrutura e ação ou entre a análise macro e a análise micro”. Para o autor, a “verdadeira distinção e relação fundamental a fazer era entre ação conformista e ação rebelde”. Também o sociólogo italiano Ardigò (1988: 288) põe em primeiro plano o “ambivalente ponto de vista do observador científico” ao privilegiar, quer o sistema em prejuízo da pessoa, quer a pessoa sem se preocupar com os aspetos sistémicos. É portanto criticável, observa o autor, tanto a interpretação da vida social baseada exclusivamente nas categorias relativas ao mundo da vida como a análise ‘funcionalista’ que enfatiza somente o sistema social ignorando os “limites do processo de objetivação” (Ardigò, 1988: 288). Nesta mesma linha, são igualmente esclarecedoras as afirmações de Beck (2000: 326): “(...) não se deve jurar lealdade a qualquer determinado ponto de vista ou perspetiva teórica. A decisão de adotar uma abordagem realista ou construtivista, tem para mim mais um aspeto pragmático, o facto de escolher os meios adequados para atingir o objetivo pretendido”.

Gradualmente, foi emergindo o consenso de que o grande desafio é combinar, de forma apropriada, os métodos, as perguntas e as questões empíricas e não defender uma única abordagem metodológica para todos os problemas. Nesse prisma, Patton propõe um paradigma – que ele define como “*paradigm of choice*” – que deve dispor de um amplo repertório de métodos e de técnicas a serem utilizadas na variedade dos problemas. Esse paradigma, acrescenta o autor, reconhece que diferentes métodos são apropriados para diversas situações e propósitos de avaliação (Patton, 1997: 297). Há, portanto, um consenso sobre a importância de adotar uma pluralidade metodológica para responder às diversas questões colocadas pelos processos de avaliação e pelas necessidades dos diferentes atores neles envolvidos.

No que diz respeito à preocupação pela utilidade dos resultados, as diferentes definições analisadas enfatizam que a avaliação não deve ser considerada como uma atividade separada do processo de tomada de decisões. Nesse sentido, Stufflebeam

(1996:118) afirma que o mais importante propósito da avaliação “não é provar, mas melhorar” (*is not to prove but to improve*). Trata-se de uma ferramenta para aprimorar os programas com vista a responder às necessidades dos beneficiários. Ela nasce num contexto específico – o contexto decisional ou gerencial – que justifica e explica a necessidade de um processo cognitivo (a avaliação) voltado para a redução da “complexidade” (Bezzi, 2003) e da “incerteza” (Weiss, 1998). Como afirmam Aguilar e Ander-Egg (1995:132), não se avalia por interesses académicos ou por curiosidade intelectual, mas a partir de um claro sentido utilitário e prático. A pesquisa avaliativa deve ser efetuada para que seus resultados e recomendações sejam realmente aplicados. Entretanto, todas as experiências de avaliação realizadas nos EUA, a partir dos de 1960, demonstraram uma escassa capacidade de influência sobre as decisões e os programas sociais do governo. Essa constatação levou Patton (1997) a escrever um livro específico sobre o assunto: *Utilization-Focused Evaluation* (Avaliação Focada na Utilização), no qual argumenta que “os resultados das avaliações deveriam ser julgados pela sua utilidade” (Patton, 1997: 20). O autor recomenda que os avaliadores mantenham uma estreita relação com os promotores e financiadores da avaliação no sentido de os ajudar a identificar os pontos críticos do programa ou serviço e assim escolher o tipo de avaliação de que eles necessitam.

Na mesma linha se inserem algumas experiências internacionais que visam identificar a melhor forma de transferir os resultados dos estudos académicos para as políticas públicas e para a organização dos serviços. Neste contexto é interessante a experiência canadense dos coletivos de pesquisa (*research collective*) que desenvolvem atividades de interação e de intercâmbio entre investigadores e decisores políticos como estratégias promissoras para aumentar a possibilidade aplicar de os resultados das avaliações na gestão e a formulação de políticas (Pineault *et al.*, 2007; Brousselle *et al.*, 2009).

Em relação ao reconhecimento do pluralismo de valores, é importante analisar o modelo de avaliação proposto por Guba e Lincoln (1989), no famoso livro *Fourth Generation Evaluation* (Avaliação da quarta geração), que visa valorizar as demandas, as preocupações e os assuntos postos pelos diversos atores e grupos de interesses. Para os autores, é a interação constante entre avaliador e implicados que cria o produto da avaliação através da adoção da abordagem hermenêutico-dialética. Esse modelo que os autores denominam *responsive constructivist evaluation* (avaliação responsável e construtivista) pretende superar os limites das anteriores gerações sobre processos de avaliação, mais precisamente, i) o excessivo poder dos gestores no processo avaliativo, ii) a incapacidade

de aceitar e conciliar o pluralismo de valores e iii) o excessivo envolvimento com o paradigma científico de tipo positivista. Com a designação de *responsive*, os autores visam delinear, a partir das elaborações de Stake (1996), uma maneira diferente de abordar a avaliação, ou seja, focando-a num processo interativo, participativo e negociado que envolve todos os sujeitos, incluindo: a) as pessoas empenhadas na produção, no uso e na implementação da avaliação; b) os beneficiários do processo avaliativo; e c) as ‘vítimas’, ou seja, as pessoas que podem ser prejudicadas ou afetadas negativamente por esse processo (Guba e Lincoln, 1989: 201). Já com o termo *constructivist*, os autores referem-se a uma metodologia fundamentada no assunto em que a verdade não corresponde a uma realidade objetiva, mas que resulta do consenso entre construtores informados. A realidade, sublinham os autores, é assim uma construção social, podendo existir tantas construções quantas são as pessoas envolvidas.

5. Abordagens à avaliação

Como já foi observado na introdução deste artigo, existem muitas perspectivas e escolas de avaliação que promovem um intenso debate sobre os temas centrais e fundamentos da avaliação. Nesta secção apresentamos os principais modelos de avaliação desenvolvidos nos últimos 50 anos e a contribuição teórica e metodológica dos mais reconhecidos estudiosos sector desta matéria. Para esse fim, adotaremos a tipologia de Nicoletta Stame (2001), que tem reclassificado os modelos de avaliação em três grandes famílias ou abordagens: i) positivista-experimental; ii) pragmatista-da qualidade; iii) construtivista. Importa realçar que a autora prefere utilizar o termo ‘abordagem’ em vez de ‘paradigma’ por duas razões: a) “uma abordagem é um conjunto de modelos diferentes, portanto, apresenta fronteiras menos rígidas dos paradigmas”; b) “as abordagens coexistem, enquanto os paradigmas – pelo menos na definição de Kuhn (2006) – substituem-se um ao outro no curso das revoluções científicas” (Stame, 2001: 25).

5.1. Abordagem positivista-experimental

Desenvolve-se nos anos em que iniciaram os estudos avaliativos das políticas públicas dirigidas ao combate da pobreza nos EUA. Entre os principais autores que trabalharam nesta perspectiva Stame (2007:26) inclui Hyman, Suchman e Weiss com seus primeiros trabalhos, Campbell, o teórico da experimentação, e os autores mais ecléticos,

como Rossi, Freeman e Lipsey e Chen. De acordo com esta abordagem, a tarefa principal da avaliação é mensurar os efeitos de um programa, o que implica a capacidade de definir com precisão os seus objetivos e estabelecer os indicadores aptos para efetuar tais mensurações. Para superar as dificuldades advindas da não clareza dos objetivos, foi introduzida a distinção entre finalidades (*goals*) ou as aspirações não quantificáveis e os objetivos (*objectives*), ou seja, as metas a serem alcançadas e mensuradas através dos indicadores (Rossi *et al.*, 1999: 94). Esta abordagem recorre a desenhos experimentais ou quase-experimentais, a técnicas de pesquisa quantitativas e a avaliadores externos ao programa para atender ao requisito da objetividade na avaliação. O foco desta abordagem é a avaliação *ex-post* (ou somativa), sem a mínima preocupação em analisar o processo de implementação do programa (avaliação formativa). De facto, um dos maiores limites deste modelo é a própria dificuldade em aceder à “caixa preta” do programa, ou seja, compreender o seu funcionamento e os mecanismos intermediários que relacionam as causas com os efeitos esperados (Moro 2009: 48). Sucessivamente, esta rígida posição inicial evoluiu, reconhecendo que um programa pode ser implementado diferentemente do seu desenho inicial, ser gerenciado de forma ineficaz ou ficar comprometido por ingerência política (Rossi *et al.*, 1999). Assim, foram desenvolvidas diversas técnicas de monitoramento e de aferição intermediária dos resultados. Outro limite deste modelo é sua incapacidade de identificar os mecanismos que promovem a mudança. Para superar estas limitações, várias vozes críticas se levantaram (incluindo a partir da própria abordagem experimental) contra a centralidade da metodologia e em favor de uma maior consideração da teoria (Stame (2007). Esta perspectiva – definida como ‘avaliação orientada pela teoria’ (*theory-driven evaluation*) (Chen (1990, *apud* Stame, 2001) – afirma que a tarefa da avaliação é a de explicar por que um programa deveria funcionar como planeado e esclarecer, portanto, a teoria que o fundamenta.

5.2. Abordagem pragmatista da qualidade

Contrariamente à abordagem experimental, em que o critério de avaliação é o alcance dos objetivos, esta abordagem – “nascida em ambiente ‘pragmatista’ em oposição ao positivismo” (Stame, 2001: 29) – pretende avaliar os programas com base nos valores que podem ser internos ou externos ao mesmo. Neste sentido, enfatiza Moro (2009: 55), não deve ser o programa a “estabelecer *a priori* a validade de um objetivo”, mas deve ser o resultado de uma pesquisa avaliativa. Para Michel Scriven (1991), considerado o fundador desta abordagem, a avaliação deve ser livre dos objetivos (*goal free evaluation*). A

perspetiva orientada para os objetivos, comenta Stame (2001:29), poderia representar “um alibi para uma pretensa neutralidade de valores do avaliador”. O avaliador não deve portanto deixar-se influenciar pelos objetivos, mas deve basear-se nos seu valores e competências.

Scriven (1991) desagrega o juízo de valor em duas dimensões: i) mérito (*merit*) ou valor intrínseco de uma atividade, mensurado com padrões de qualidade específicos para aquela atividade; ii) valor extrínseco (*worth*) de uma intervenção que responde às necessidades dos beneficiários que vivem naquele contexto.

Este modelo é adequado para avaliar a efetividade e a eficiência de serviços que desenvolvem regularmente as suas atividades em resposta às necessidades da população. De facto, os princípios deste modelo encontram-se incorporados em todas as estratégias de avaliação das instituições que se inspiram na Gestão da Qualidade e na Melhoria Contínua da Qualidade. Nesse sentido, a qualidade torna-se a propriedade positiva a ser levantada. Entretanto, sendo a qualidade um conceito polissémico e multidimensional (Serapioni, 2009), é preciso desagregá-la em diferentes dimensões e identificar indicadores e padrões que nos informem sobre os níveis de alcance da qualidade. Trata-se do processo que Scriven (1995) define como ‘lógica da avaliação’. Uma vez estabelecidos os padrões de qualidade de um determinado serviço social, a equipe de avaliação é responsável por: i) mensurar o desempenho (*performance*) do serviço; ii) compará-lo com os padrões estabelecidos na fase da programação; iii) expressar um juízo sobre as variações identificadas. Para este tipo de atividade, os avaliadores podem recorrer ao juízo de especialistas (*experts*), mas podem também envolver especialistas leigos ou os mesmos beneficiários (Stame, 2001).

No campo da saúde, Donadedian (1980) sugere desagregar a qualidade nas três grandes categorias que compõem um serviço de saúde: estrutura, processo e resultado.

5.3. Abordagem construtivista

Trata-se de um conjunto de modelos de avaliação que partilham algumas características comuns, nomeadamente a de valorizar a contribuição dos diferentes atores e a importância atribuída à fase de implementação de um programa. O interesse dos avaliadores é observar o desenvolvimento das intervenções, mais do que aferir o alcance dos objetivos ou dos padrões previamente estabelecidos. Nesse sentido, são considerados também os efeitos positivos ou negativos inesperados (Moro, 2009). Esta abordagem inclui diferentes modelos de avaliação: a ‘Avaliação da quarta geração’, de Egon C. Guba e

Yvonne S. Lincoln (1989); a ‘Avaliação focada na utilização’, de Michael Q. Patton (1997); a ‘Avaliação como processo social e político’, de Lee J. Cronbach (1996); a ‘Avaliação sensível’ (*Responsive evaluation*), de Robert E. Stake (2007) e a ‘Avaliação para o empoderamento’ (*Empowerment Evaluation*), de David M. Fetterman (1994).

Relativamente a avaliação da quarta geração, trata-se, na opinião de Guba e Lincoln (1989), de um espaço democrático e dialético onde os sujeitos podem interagir e participar livremente no processo deliberativo, apresentando os seus próprios pontos de vista e reivindicações. O avaliador participa, junto com os outros atores, na construção do consenso, desempenhando o difícil papel de mediador, o que requer capacidades e habilidades de negociação (Guba e Lincoln, 1989). De facto, nesse processo de negociação, como advertem Aguilar e Ander Egg (1995: 90), “é preciso conciliar ou compatibilizar as demandas de todos os atores implicados, com as exigências do método científico”.

A avaliação focada na utilização atribui muita importância ao contexto político e organizativo em que são tomadas as decisões. Para Patton (1997), os que encomendam a avaliação são mais motivados e propensos a aplicar as recomendações do estudo se forem envolvidos nas diferentes etapas da avaliação e se forem considerados atores principais desse processo. Para isso é importante que o avaliador prepare o terreno para que os resultados da avaliação sejam implementados. O envolvimento desses atores, enfatiza Patton (1998: 226), tem um impacto positivo porque ajuda os decisores a adotar a perspetiva dos avaliadores e a “aprender a pensar avaliativamente” (*to learn to think evaluatively*).

Para Cronbach (1996) – teórico da relação entre avaliação e política – os avaliadores devem atentamente considerar a influência que o contexto político exerce nos programas sociais. Neste prisma, Cronbach (1996: 405) salienta que a verdadeira missão da avaliação não é a de eliminar a possibilidade de erros dos decisores políticos, mas a de “facilitar o processo democrático e pluralista, contribuindo para a sensibilização de todos os participantes”.

A ‘avaliação sensível’ aos valores dos *stakeholders*, desenvolvida por Stake (2007), estuda em profundidade casos específicos, adotando diversas estratégias avaliativas em cada caso. O mesmo programa, acrescenta o autor, desenvolve-se de forma diferente nas diversas situações. Na visão de Stake (2007:162), a avaliação sensível “sacrifica um certo rigor das medições em troca de uma maior utilidade dos resultados”.

Finalmente, a ‘avaliação para o empoderamento’, cujos objetivos, de acordo com Fetterman (1997:382), remetem para o uso de conceitos e técnicas da avaliação para

“fomentar o melhoramento e a autodeterminação” e “ajudar as pessoas a ajudarem-se a si mesmas e a aperfeiçoar os seus programas”. Os avaliadores, atuando como facilitadores, ensinam as pessoas a conduzir uma avaliação dos próprios programas para que possam tornar-se autossuficientes (Fetterman, 1994). A par de outras tipologias incluídas na abordagem construtivista, este modelo enfatiza o seu valor educativo.

5.4. Modelos mistos e pluralismo

Ao equiparar as diferentes abordagens, sem reconhecer qualquer primazia a uma delas, optou-se claramente por favorecer aquelas que geralmente são consideradas mais débeis. Faz-se aqui referência não tanto à abordagem pragmática da qualidade, que é bastante utilizada nas avaliações dos serviços sócio assistenciais, educativos, de saúde e, no geral, no setor dos serviços públicos, mas à abordagem construtivista que continua a ser considerada menos importante que a abordagem positivista-experimental. Esta última, de facto, ainda é preferida tanto por promotores e financiadores de avaliações que querem alcançar dados “objetivos”, como por avaliadores formados nos pressupostos conceituais e metodológicos desta abordagem.

Após ter diferenciado as três abordagens à avaliação e ter realçado as suas características, a lógica e a coerência interna de cada uma delas, torna-se agora preciso gerenciar a complexidade decorrente desta distinção epistemológica e metodológica. Neste sentido, advoga-se a necessidade de tirar proveito da variedade de abordagens, combinando métodos e ferramentas teóricas de diferentes origens. Tal acontece espontaneamente ao formular projetos de avaliação que combinam a análise de impacto, questionários de satisfação dos utentes e entrevistas com informadores privilegiados. Importa, todavia, realçar que tais estratégias metodológicas não devem ser consideradas intercambiáveis, devendo ser trazidas de volta à lógica que as sustentam.

Cada vez mais, estes arranjos e combinações têm-se tornado questão central de debate, conceitualizada na ideia de pluralismo de abordagens na avaliação. Entre as várias formas de pluralismo possíveis, as vertentes mais interessantes e promissoras são as dos métodos mistos e a da contaminação entre diversas abordagens, as quais têm apresentado interessantes desenvolvimentos teóricos. A ideia dos métodos mistos nasce da observação das vantagens e desvantagens de cada um deles e advoga a complementaridade metodológica, ou seja assente na triangulação de métodos, seja qual é aceite tanto pelos experimentalistas, como Campbell e Russo (1999), como pelos construtivistas, como Greene e Caracelli (1997). Neste prisma, importa assinalar a difusão de métodos de

avaliação participativa, entendida não somente como estratégia de empoderamento dos diferentes *stakeholders* e beneficiários, mas também como estratégia para reforçar o alcance dos próprios programas. Contudo, a estratégia mais promissora ainda é, sem dúvida, essa da contaminação entre diversas abordagens, utilizando e combinando diversos aspetos de cada uma das abordagens e que poderia levar a desenvolvimentos fecundos (Stame, 2001).

Nesse sentido, a seguir serão apresentados dois modelos teóricos - a ‘Avaliação baseada na teoria’ e a ‘Avaliação realista’ - que Stame (2001: 40) considera serem o resultado da contaminação entre as abordagens acima apresentadas e, em particular, da “contaminação entre a abordagem positivista e construtivista”. Nestes modelos, o pluralismo não consiste somente na utilização de métodos advindos de diferentes perspetivas (experimentais, participativos, quantitativos ou qualitativos), mas na convicção de que em cada situação deve ser identificada a forma mais específica - entre uma multiplicidade de possíveis alternativas - através da qual pode operar um programa.

a) Avaliação baseada na teoria

Este tipo de avaliação, desenvolvido por Carol Weiss (1997), aponta para a necessidade de compreender os pressupostos teóricos do programa para se poder aferir a sua efetividade. De facto, a avaliação deve responder não somente à pergunta “O programa funciona?”, mas a questões como “o que o faz funcionar?”, “porque teve êxito?” e “como pode funcionar melhor?”. Conhecer exclusivamente os resultados não é suficiente para “aprimorar o programa ou para revisar uma política”. A avaliação, acrescenta a autora, deve “entrar na caixa negra” do programa. Na opinião de Stame (2001: 41), Carol Weiss “contamina uma abordagem positivista centrada na ideia de causalidade com uma abordagem construtivista, atenta à forma em que os atores reagem ao programa e o interpretam”. A avaliação baseada na teoria é desdobrada em duas dimensões: i) a teoria da implementação, que analisa a forma como são realizadas as atividades do programa, pressupondo que, se realizadas com qualidade e de acordo com o plano inicial, os resultados desejados serão atingidos; ii) a teoria do programa, que se concentra nos mecanismos que intervêm entre a prestação das atividades do programa e o alcance dos resultados. Por outras palavras, a teoria do programa ocupa-se dos mecanismos que intermediam processos e resultados. O mecanismo da mudança – enfatiza Weiss (1997: 46) – não é constituído pelas atividades do programa em si, mas pela resposta que tais atividades geram.

b) Avaliação realista

Este modelo desenvolvido por Pawson e Tilly (1997) critica a “causalidade sequencial” da abordagem positivista - segundo a qual o resultado é obtido após uma determinada intervenção - e propõe o conceito de “causalidade generativa”, que busca compreender como aquela intervenção obteve aquele resultado. Neste prisma, os autores deslocam o foco da avaliação da efetividade de um programa para a necessidade de compreender o que, deste programa, precisamente, possibilita o seu funcionamento. De acordo com Pawson (2002: 342), não são os programas que funcionam, mas os “recursos que eles oferecem para habilitar os sujeitos a fazê-los funcionar”. Este processo sobre como os sujeitos interpretam as estratégias do programa é designado pelo autor como ‘mecanismo do programa’ e representa o eixo em torno ao qual gira a avaliação realista. Na perspectiva realista, a avaliação deve, portanto, responder às seguintes perguntas: quais são os mecanismos de mudança desencadeados pelo programa? Por que em alguns contextos alguns mecanismos funcionam e outros não? Por que o mesmo programa gera resultados diversos em contextos diferentes? Partindo dos pressupostos acima mencionados, o modelo realista articula o processo de avaliação em três elementos essenciais: os mecanismos, o contexto e os resultados. Na análise de Pawson e Tilly (1997) são claramente identificáveis as influências do construtivismo quando reconhecem que em cada contexto são possíveis diversos mecanismos. Nesse sentido, a avaliação realista consiste na compreensão do contexto (as pessoas, os processos, as instituições) e dos mecanismos que podem funcionar nesse determinado contexto. Esta atividade certamente não é preordenada – como na abordagem positivista. Finalizando, este modelo reconhece a importância da relação entre o avaliador e os *stakeholders*, embora o avaliador mantenha o seu papel de cientista social, contrariamente aos outros modelos construtivistas.

Conclusão

A prática da avaliação tem vindo a assumir peso e visibilidade crescentes a nível internacional, disseminando-se, cada vez mais, entre os países do sul da Europa, graças ao impulso de vários fatores convergentes que a têm tornado cada vez mais independente das disciplinas académicas e das áreas profissionais em que surgiu. A avaliação representa uma fase importantíssima do ciclo político e é imprescindível para apreciar os programas e as políticas públicas implementadas, os efeitos desejados e indesejados e para monitorar o desenvolvimento e a adaptação constante da ação pública. Além destas importantes

funções, a avaliação representa, como realçado por vários autores citados ao longo deste artigo, um pilar central a favor da governação e da democracia na sociedade pós-moderna, sendo que cada avaliação, como afirma Patton (2002: 127), – “é uma oportunidade para fortalecer a democracia, ensinando as pessoas a pensar avaliativamente”. A avaliação, como acrescenta Chelimsky (1997), permite beneficiar aqueles que tomam decisões sobre as políticas públicas, e isso, por seu lado, beneficia os cidadãos que têm de lidar com essas decisões e suas consequências.

Neste artigo foram analisados vários aspetos referentes ao processo de avaliação: o seu papel e os seus diversos usos, os pressupostos teórico-metodológicos, as suas abordagens, assim como a situação atual da temática a nível internacional. Com o intuito de reiterar a importância da avaliação no que toca à sua capacidade de retroalimentar o ciclo das políticas públicas, como a sua proeminência no campo das ciências sociais e políticas, finaliza-se esta análise apresentando alguns pontos de tensão persistentes nas discussões sobre esta matéria. A primeira questão refere-se ao debate quantitativo-qualitativo que tem contribuído para o reconhecimento dos métodos qualitativos, tais como estudos de caso e observação participante, hoje parte integrante do repertório dos avaliadores. A continuação deste debate, porém, desvia muita da energia intelectual das questões mais urgentes e dos desafios que esta questão coloca, considerando-se, por isso, urgente sair dessa “zona de conflito”. Outro ponto controverso diz respeito à forma como os resultados das avaliações devem ser utilizados. Nos anos de 1970, muitos avaliadores acreditavam que os seus resultados podiam retroalimentar os processos de tomada de decisões. Porém, a experiência ensinou que as decisões não são tão facilmente tomadas no mundo da política. Tais resultados empíricos levaram assim a reconhecer que o uso instrumental (*instrumental perspective*) da avaliação é raro e que o tipo de uso mais difuso da avaliação é chamado de “iluminação” ou de “esclarecimento” (*enlightenment perspective*), cujos resultados podem mudar a maneira como as questões sociais e os programas são enquadrados e a forma como os *stakeholders* pensam acerca de problemas e soluções (Cook, 1997). Mas será que os financiadores e promotores se contentam apenas com a função de esclarecimento não exigindo também uma avaliação que se retroalimente dos processos de decisão?

Importa, por fim, assinalar o debate referente ao papel do avaliador, que varia de acordo com a perspectiva sobre a avaliação que é adotada. Na perspectiva orientada para o desenvolvimento institucional (*Development perspective*), a avaliação é considerada uma ferramenta flexível que visa aprimorar o desempenho das instituições e promover a mudança organizacional. De acordo com esta perspectiva, o avaliador torna-se “parceiro”

(*partner*) ou “amigo crítico” (*critical friend*), segundo Fetterman (1994), mas também desenvolve funções de advocacia, segundo Stake (2007), podendo ainda ser encarado como facilitador, de acordo com a proposta de Guba e Lincoln (1989). Já na perspectiva da avaliação orientada para a análise da eficiência e da efetividade (*accountability perspective*), o avaliador deve manter a independência e uma certa distância para poder aferir de forma objetiva o valor ou mérito do programa (Scriven, 1995). Hoje em dia, contudo, esta visão sobre o papel do avaliador é objeto de várias críticas.

Referências bibliográficas

- Aguilar, M.J., Ander-Egg, E. (1995), *Avaliação de serviços e programas sociais*. Petrópolis, Editora Vozes.
- Beck, U. (2000), *La società del rischio*, Roma, Carocci.
- Bezzi, C. (2003), *Il disegno della ricerca valutativa*. Milão: Angeli.
- Brousselle, A.; Contandriopulos, D; Lemire, M (2009), “Using Logic Analysis to Evaluate Knowledge Transfer Initiatives: The case of the research collective on the organization of Primary Health Care”, *Evaluation*,15, pp. 165-183
- Campbell D.; Russo J. (1999), *Social experimentation*, Thousand Oaks: Sage Publications.
- Campbell, D.; Stanley, J. (1966), *Experimental and Quasi-Experimental Design for Research*, Chicago, Rand McNally.
- Cerqueira Prata A.C. (2012), *Metodologia da avaliação das ações sociais*. CEPAL, Comissão Económica para América Latina e o Caribe, Março de 2007; pp. 43. Site visitado: 20 Fevereiro 2012:<http://www.eclac.org/publicaciones/xml/4/29014/LSBRSR182AnaCarolinaAiresCerqueiraPrata.pdf>
- Chelimsky, E.; Shadish, W. (eds.) (1997), *Evaluation for the 21st century. A handbook*, Thousand Oaks: Sage Publications.
- Cohen, E., Franco, R. (1993), *Avaliação de projetos sociais*, Petrópolis, Vozes.
- Cook, T.D. (1997), Lessons learned in Evaluation over the past 25 years. In; Chelimsky, E.; Shadish, W. (eds.), *Evaluation for the 21st century. A handbook*, Thousand Oaks: Sage Publications.
- Cronbach, L.J. (1996), Ninety-five theses for reforming program evaluation. In: G.F. Madaus, Scriven M.S., Stufflebean D.L. (Orgs.), *Evaluation models. Viewpoints on educational and human service evaluation*, Boston:Kluwer-Nijhoff, 11^a Edição; pp. 405-412.
- Donabedian, A. (1980), *Exploring in Quality Assessment and Monitoring, Definitions of Quality and Approaches to its Assessment*. Ann Arbor, Health Administration Press, University of Michigan, v.1.
- Faria, C.A.P. (2005), “A política da avaliação de políticas públicas”, *Revista Brasileira de Ciências Sociais*, 20 (59), pp. 97-109.

- Fetterman, D. (1994), “Empowerment evaluation”, *Evaluation Practice*, 1994 15(1), pp. 1-15.
- Fetterman, D. (1997), Empowerment Evaluation and Accreditation in Higher Education. In: Chelmsky E., Shadish W.R. *Evaluation for the 21st Century*. Thousand Oaks, Sage, pp. 381-394.
- Frey, K. (2000), “Políticas Públicas: um debate conceitual e reflexões referentes à prática da análise de políticas no Brasil”, *Planejamento e Políticas Públicas*, 21, pp. 211-259.
- Garcia, A.T. (2014), “A implementação de políticas públicas de ambiente – o caso da qualidade da água para consumo humano”, *Análise Social*, 211, XLIX (2.º), 2014.
- Glasser, P.H. *La ricerca valutativa*. Padova, Fondazione Cancan, 1972.
- Greene, J. E., Caracelli, V. (1997), Advances in mixed-method evaluation: the challenges and benefits of integrating paradigms, *New Direction for Programme Evaluation*, n. 74, San Francisco: Jossey-Bass.
- Guba, E.G.; Lincoln, Y.S. (1981), *Effective evaluation: improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco: Jossey-Bass.
- Guba, E., Lincoln, Y. (1989), *Fourth Generation Evaluation*. Newbury Park, Sage Publications.
- Hartz, Z.M.A. “Undisciplinary comments from evaluative research in health”, *Ciência & Saúde Coletiva* 2008, 13(6), pp. 1711-1717.
- Hartz, Z.M.A. (1999), “Avaliação dos programas de saúde: perspectivas teórico metodológicas e políticas institucionais”, *Ciência & Saúde Coletiva*, 4(2), pp. 341-353.
- Howlett, M.; Ramesh, M.; Perl, A. (2009), *Studying Public Policy*, Oxford: Oxford University Press.
- Kuhn T.S. (2006), *A estrutura das revoluções científicas*, São Paulo, Perspectiva, 2006.
- Madaus G.F., Scriven M. e Stufflebeam D.L. (1996), *Evaluation Models. Viewpoints on Educational and Human Services Evaluation*. Dordrecht:Kluwer Nijhoff Publishing, 11ª Edição.
- Minayo, M.C., Sanches, O. Quantitativo-Qualitativo: Oposição ou Complementaridade? *Caderno de Saúde Pública* 1993, 9 (3), pp. 239-262.
- Moro, G. (2009), *La valutazione delle politiche pubbliche*. Roma: Carrocci.
- Novaes, H.M.D. (2000), “Avaliação de programas, serviços e tecnologias em saúde”, *Revista Saúde Pública*, 34(5), pp. 547-59
- Øvretveit, J (1998), *Evaluating Health Interventions*. Buckingham, Open University Press.
- Palumbo, M. (2001), *Il processo di valutazione. Decidere, programmare, valutare*. Milao, Angeli.
- Patton, M.Q. (1998), “Discovering Process Use”, *Evaluation*, 4(2), pp. 225-233.
- Patton, M.Q. (1999), “Enhancing the quality and credibility of qualitative analyses”, *Health Services Research*, 34, 5, pp. 1189-1208.
- Patton, M.Q. (1997), *Utilization-Focused Evaluation. The new century text*. Thousand Oaks, Sage Publications.

- Pawson, R., Tilley, N. (1997), *An Introduction to Scientific Realist Evaluation*. In: Chelimsky, E. and Shadish, W.R. (eds.), *Evaluation for the 21st Century*, Thousand Oaks, Sage, 1997.
- Pawson, R. (2002), “Evidenced-based Policy: The Promise of ‘Realist Synthesis’”, *Evaluation*, 8(3), pp. 340-358.
- Pineault, R.; Tousignant, P.; Roberrg, D.; Lamarche, P.; Rheinarz, D.; Larouche, D.; Beaulne, G.; Lesage, D. (2007), “Involving decision-makers in producing research-syntheses. The case of Research Collective on Primary Health Care in Quebec”, *Healthcare Policy*, 2,4, pp. 11-17.
- Rossi, P.H.; Freeman, H.E.; Lipsey, M.W. (1999), *Evaluation. A systematic approach*, Thousand Oaks, Sage Publications.
- Rutman, L. (1984), *Evaluation research methods: a historical guide*, Beverly Hills, Sage Publications.
- Scriven, M. (1995), “The Logic of Evaluation and Evaluation Practice”, *New Directions for Evaluation*, 68, pp. 49-70.
- Scriven, M. (1991), *Evaluation Thesaurus*, Newbury Park, Sage Publications.
- Serapioni, M. (2000), “Métodos qualitativos e quantitativos: algumas estratégias para a integração”, *Ciência & Saúde Coletiva* 2000, (5)1: 187-192.
- Serapioni, M. (2009), “Avaliação da qualidade em saúde. Reflexões teórico-metodológicas para uma abordagem multidimensional”, *Revista Crítica de Ciências Sociais*, 85, 65-82.
- Shortell, S.M. (1999), “The Emergence of Qualitative Methods in Health Services Research”, *Health Services Research*, 34(5), pp. 1083-1090.
- Stake, R.E. (1980), Programme evaluation, particularly responsive evaluation. In: W.B. Dockrell & D. Hamilton (eds.), *Rethinking educational research*: London: Hodder & Stoughton.
- Stake, R.E. (1996), Program Evaluation, Particularly Responsive Evaluation. In: G.F. Madaus, Scriven M.S., Stufflebean D.L. (Orgs.), *Evaluation models. Viewpoints on Educational and Human Service Evaluation*, Boston, Kluwer-Nijhoff, 11^a Edição, pp. 287-310.
- Stake, R.E. (2007), La valutazione di programmi, con particolare riferimento alla valutazione sensibile. In Stame N. (org.) *Classici della valutazione*, Milão, Angeli, 2007; pp. 156-177.
- Stame, N. (org.)(2007), *Classici della Valutazione*. Milão, Angeli.
- Stame, N. (2001), Tre approcci principali alla valutazione: distinguere e combinare. In: Palumbo M. (org.) *Il processo di valutazione. Decidere, programmare, valutare*. Milão, Angeli, pp. 21-46.
- Stufflebeam, D.L. (1996), The CIPP Model for Program Evaluation. In: Madaus, G.F., Scriven, M. e Stufflebeam, D.L. (eds.), *Evaluation Models. Viewpoints on Educational and Human Services Evaluation*. Dordrecht, Kluwer Nijhoff Publishing, 11^a Edição, 1996.
- Trevisan, A.P.; van Bellen, H.M. (2008), “Avaliação de políticas públicas: uma revisão teórica de um campo em construção”, *Revista de Administração Pública*, 42 (3), pp. 529-550.
- Weiss, C.H. (1997), “Theory-Based Evaluation: Past, Present, and Future”, *New Directions for Evaluations*, 76, pp. 41-55.
- Weiss, C.H. (1998), *Evaluation*. Second Edition. Uppler Saddle River: Prentice Hall.

Mauro Serapioni. Investigador do Centro de Estudos Sociais (CES), Universidade de Coimbra (Coimbra, Portugal). Endereço de Correspondência: Centro de Estudos Sociais Colégio S. Jerónimo, Apartado 3087 3000-995, Coimbra, Portugal. *E-mail:* mauroserapioni@ces.uc.pt

Artigo recebido a 20 de maio de 2015. Publicação aprovada a 15 de novembro de 2015