



As Avaliações Vão Para a Gaveta? Uma revisão teórica dos fatores relacionados ao uso de avaliações de políticas públicas

What Happens to Policy Evaluations? A review of the factors related to evaluation use

Cinthia Helena de Oliveira Bechelaine¹, Ivan Beck Ckagnazaroff²

RESUMO

As avaliações de políticas públicas são realizadas com base na premissa de que os resultados serão direta e automaticamente utilizados, o que nem sempre ocorre. Mesmo que exista um consenso inegável sobre o valor das avaliações para a efetividade da administração pública, o tema da utilização das avaliações ainda é um tema pouco discutido no Brasil. O objetivo deste trabalho é construir um arcabouço conceitual dos fatores que favorecem ou bloqueiam o devido uso das avaliações, baseado numa aprofundada revisão bibliográfica. Destacam-se três estudos principais, que contribuíram para a pesquisa sobre as variáveis que influenciam a utilização das avaliações: (1) o estudo de Cousins e Leithwood (1986), que tiveram por objetivo identificar os estudos empíricos sobre uso das avaliações conduzidos entre 1971 e 1985, destacando-se o trabalho de Leviton e Hughes (1981); (2) o trabalho de Shulha e Cousins (1997), que conduziram um trabalho semelhante, no período entre 1986 a 1997; e, por fim, (3) o estudo de um grupo de pesquisadores (Johnson *et al.*, 2009) que se dedicou à pesquisa sobre o uso das avaliações nos principais periódicos de avaliação, em um período de 25 anos, compreendido entre 1986 a 2005. Como implicações da discussão teórica, conclui-se que a utilidade das avaliações envolve uma série processos que, do início ao fim, irão afetar seu uso e aplicação, tais como o contexto em que as avaliações estão inseridas, a estrutura organizacional predominante e o envolvimento dos stakeholders no processo.

Palavras-chave: Administração pública. Políticas públicas. Avaliação de políticas públicas. Utilização dos resultados de avaliações.

ABSTRACT

Evaluations of public policy are made based on the premise that the results will be directly and automatically used to promote the necessary improvements; but this does not always occur. Although there is a consensus about the undeniable value of evaluations to the effectiveness of public administration, the issue of evaluation use is rarely discussed in Brazil. The purpose of this study, based on a detailed literature review, is to develop an overview of the factors that favor or block the utilization of policy evaluations. The paper highlights three main studies that have contributed to the research on the factors influencing the use of evaluations: the study of Cousins and Leithwood (1986), which reviewed empirical studies of evaluation use conducted between 1971 and 1985, giving major attention to the work of Leviton and Hughes (1981); the work of Shulha and Cousins (1997), who conducted a similar study of research published between 1986 and 1997; and finally, the study of a group of

¹ Universidade Federal de Minas Gerais - Belo Horizonte, Minas Gerais - Brasil. cinthia.bechelaine@gmail.com

² Universidade Federal de Minas Gerais - Belo Horizonte, Minas Gerais - Brasil. ivanbeck00@gmail.com

researchers (Johnson et al., 2009) who looked at research on the use of evaluation in leading evaluation journals of evaluation over the 25 year period covering 1986 to 2005. The review carried out suggests that a series of processes is involved that affect the use, application and usefulness of policy evaluations, such as the context in which the evaluation is carried out, organizational structure and the degree of stakeholder involvement in the process.

Keywords: Public policy. Policy evaluation. Policy evaluation use.

1. INTRODUÇÃO

Atualmente, como afirmado por Bamberger, Rugh e Mabry (2011), existe uma demanda crescente por avaliações sistemáticas de ações, programas e projetos governamentais. Segundo os autores, cada vez mais acende o interesse de agências financiadoras, governos, sociedade civil, e até dos próprios beneficiários, por respostas como: o projeto alcançou seus objetivos? O programa teve impacto sobre a população? Quem se beneficiou ou não com a ocorrência do programa? O programa deve continuar? Respostas essas que só podem ser comprovadas por avaliações das políticas públicas, estudos necessários à tarefa de comparar e escolher as melhores alternativas para a implementação de programas governamentais e para aqueles que já estão em andamento, buscando implementar intervenções com maior benefício social para a população.

Apesar de o tema “avaliação” de políticas públicas acomodar muitas definições, com uma profusão de conceitos e significados (Patton, 2008), certos autores (Arretche, 1998; Weiss, 1998; Patton, 2008; Ala-Harja & Helgason, 2000; Aguilar & Ander-Egg, 1995) concordam que a finalidade das avaliações não é outra senão contribuir para o aprimoramento da política ou do programa avaliado, com base em análises sistemáticas dos aspectos de um programas e seu valor. Assim, a avaliação é um instrumento de mudança que não deve apenas demonstrar os problemas ou propor soluções para os problemas detectados, mas, produzir mudanças apropriadas que conduzam ao cumprimento de padrões de qualidade em diferentes serviços. Para Worthen, Sanders e Fitzpatrick (2004, p. 594), a avaliação, devidamente praticada, deve levar “a melhorias diretas e incontestáveis de sistemas, programas e práticas – melhorias que não teriam ocorrido de nenhuma outra forma”. Faria (2005) argumenta que sistemas institucionalizados de avaliação de programas e políticas públicas constituem elementos profícuos na criação e no aperfeiçoamento de mecanismos de prestação de contas do desempenho das burocracias e representantes políticos.

Contudo, embora seja clara a percepção da importância das avaliações de políticas públicas, o crescimento do campo e o aumento dos investimentos realizados nessa área de avaliações, observa-se uma questão que merece destaque nesse debate: a utilização. Mesmo que exista um consenso inegável sobre o valor das avaliações de programas governamentais para a efetividade da administração pública, em suas várias instâncias governamentais, pouca atenção tem sido dada à apropriação, pelos gestores dos programas, políticos e formuladores de políticas, dos resultados das avaliações.

De acordo com Patton (2008), o modo como as avaliações são utilizadas afeta o gasto de bilhões de dólares investidos em programas para a redução da pobreza, melhoria das condições de saúde da população, políticas públicas de emprego e segurança pública, bem como outras áreas. Segundo Faria (2005, p. 101), existe uma ingenuidade nas expectativas de se produzir, apenas com a realização de avaliações, “melhores e mais sábias decisões”, tendo em vista inúmeras evidências de que o efeito esperado de *feedback* das avaliações é falho. Trevisan e van Bellen (2008) declaram que as informações provindas da avaliação são

utilizadas de forma insuficiente. De forma semelhante, Ala-Harja e Helgason (2000) argumentam que o uso apropriado dos resultados das avaliações não constitui tarefa fácil. Segundo os autores, as avaliações representam, historicamente, um universo de promessas que não chegam a se realizar.

Ainda que a utilização de avaliações tenha sido um campo explorado internacionalmente, com pesquisas que se iniciaram nos anos 1970 – quando alguns autores (Weiss, 1972; Patton *et al.*, 1977; Alkin, 2004; Leviton & Hughes, 1981; Cousins & Leithwood, 1986; Guba & Lincoln, 1989) passaram a questionar a efetiva utilidade das avaliações para influenciar decisões e contribuir para o aperfeiçoamento de programas e políticas – no Brasil, ainda não é perceptível tal preocupação. A ausência de estudos teóricos e empíricos sobre a utilização das avaliações de programas, bem como a lacuna existente em modelos de meta-avaliações e a carência por organizações que prezem por melhores padrões nas avaliações são alguns fatos que provam insuficiência brasileira nesse assunto (Serpa, 2010).

O processo de institucionalização da avaliação na América Latina ocorreu tardiamente. A perspectiva prevalecente foi a de instrumentalização da pesquisa avaliativa para o sucesso da reforma do Estado. E, mesmo acompanhando as atividades referentes à modernização do governo, a função avaliação no Brasil é ainda bastante incipiente, com institucionalização precária, como comenta Faria (2005). De acordo com Lobo (2001), a administração pública brasileira, à semelhança de outros países latino-americanos, não desenvolveu ao longo do tempo uma cultura de efetividade em relação ao gasto público com os programas sociais. E quando existem, nota-se que a maior parte dos estudos desenvolvidos no Brasil sobre avaliação concentram-se no desenvolvimento de metodologias adequadas para avaliar políticas sociais, sendo escassos os autores brasileiros que ousaram tratar do tema da utilização dos resultados das avaliações. Ressalta-se, no entanto, que mesmo que sejam desenvolvidos métodos sofisticados, de nada esses servirão se não forem utilizados para gerar modificações, aprimoramentos nos programas avaliados. Ainda que os estudos avaliatórios cresçam e se desenvolvam, com métodos sofisticados, com dados confiáveis sobre a realidade na qual se intervém, alocando esforços físicos, intelectuais e financeiros, se o *modus operandi* da intervenção pouco se alterar com base nas recomendações das avaliações, os estudos podem ser “engavetados” e o investimento feito será perdido.

Nesse sentido, dada a insuficiência de literatura em língua portuguesa que trata o tema, atenta-se para a necessidade de formar um referencial coeso e completo sobre a utilização dos resultados das avaliações, seja para inserir o país no debate internacional, seja para evitar o desperdício de gastos com avaliação sem a apropriada utilização de seus resultados. Apesar do crescente interesse em avaliações de programas governamentais no Brasil, é possível observar certa escassez de estudos teóricos nacionais que permitam verificar a utilização de tais avaliações como instrumento de correção de rumos e aprimoramento dos serviços públicos.

A presente pesquisa tem por objetivo evidenciar os principais fatores que favorecem ou bloqueiam o devido uso das avaliações na finalidade de aprimorar os programas governamentais, fundamentados em uma revisão bibliográfica sobre o assunto da utilização dos resultados de estudos avaliatórios. Trata-se, portanto, de um estudo baseado em pesquisa bibliográfica densa, cuja coleta de dados foi realizada por intermédio da revisão dos principais autores que trataram do tema da utilização.

Para cumprir esse objetivo, este trabalho está compreendido em cinco seções principais, a começar por esta introdução. Na seção seguinte apresentam-se aspectos conceituais gerais do campo da utilização das avaliações de programas, discute-se o conceito de utilização dos

resultados das avaliações. Na terceira seção, apresentam-se os diferentes usos presentes nas avaliações de programas, ou seja, as formas como as mesmas podem ser utilizadas. A seção quatro trata dos diferentes autores que abordam o “calcanhar de Aquiles” das avaliações – as variáveis que influenciam (ou não) a utilização das avaliações na melhoria de programas públicos. Destacam-se três estudos que contribuíram para a pesquisa sobre as variáveis que influenciam a utilização das avaliações, quais sejam, o estudo de Cousins e Leithwood (1986), o trabalho de Shulha e Cousins (1997), e o estudo de Johnson *et al.* (2009). Por fim, como consideração final deste trabalho sintetiza-se a revisão teórica realizada, apontando a necessidade de integração dessas abordagens para analisar a utilização das avaliações face ao ambiente e à capacidade institucional de organizações públicas.

2. DISCUTINDO A UTILIZAÇÃO DOS RESULTADOS DAS AVALIAÇÕES

A questão da utilização dos resultados das avaliações tem sido discutida por uma série de autores a partir da década de 1970 (Weiss, 1972; Patton *et al.*, 1977; Alkin, 2004; Leviton & Hughes, 1981; Cousins & Leithwood, 1986; Guba & Lincoln, 1989). De acordo com Shulha e Cousins (1997), esses trabalhos foram responsáveis por inaugurar “o uso das avaliações” como um campo de pesquisa em potencial. Com base no esforço desses autores, uma série de outros trabalhos, datados da década de 1980, aprofundaram a pesquisa com apontamentos teóricos e empíricos.

De acordo com Patton (2008), não existe uma definição ideal e absoluta para o conceito de utilização, dado que a mesma é inerente aos valores e objetivos daqueles que farão uso dos resultados das avaliações. Por resultados das avaliações, entende-se qualquer informação associada aos achados de uma avaliação, sejam informações, interpretações, dados ou recomendações comunicados pelos avaliadores na conclusão de avaliações ou até mesmo no decorrer das mesmas (Cousins & Leithwood, 1986).

Para Alkin (2004, p. 19), “*evaluation use*” refere-se à proposta, às consequências planejadas que resultam da aplicação de informações de uma avaliação a um problema, uma questão ou preocupação. Em outras palavras, como afirmado pelo autor, a utilização de avaliação alude à aplicação das informações geradas por uma avaliação para alcançar fins pretendidos.

Para Shulla e Cousins (1997), a utilização das avaliações, ao contrário de um conceito unitário, representa um fenômeno multidimensional, caracterizado por fatores como relevância da avaliação, credibilidade do avaliador, envolvimento com o usuário, comunicação, atendimento às necessidades de informação, percepção da avaliação como instrumento de gerenciamento e, principalmente, características do contexto da instituição. Assim, como afirmado pelos autores, a utilização pode ser vista como um fenômeno complexo aliado ao contexto, em que as pessoas, o ambiente e as circunstâncias estão fortemente conectados, influenciando o uso.

Para Leviton e Hughes (1981), o termo utilização refere-se ao uso dos resultados das avaliações para os programas e as políticas unicamente, excluindo a utilização por acadêmicos e pela imprensa, por exemplo. Para serem consideradas utilizadas, como declarado pelos autores, é necessário, primeiro, que os resultados das avaliações possam ser traduzidos em recomendações para as políticas públicas avaliadas, como atesta Weiss (1978 como citado em Leviton & Hughes, 1981, p. 527) e, segundo, que existam evidências de que, na ausência da informação da avaliação, os envolvidos nas atividades do programa avaliado teriam pensado ou agido de forma diferente, o que torna difícil de ser verificado por meio de pesquisas. Segundo Leviton e Hughes (1981), é importante também diferenciar o conceito de

impacto de uma avaliação do conceito de utilização. Por impacto entende-se a modificação nos programas e políticas advindas de contribuições das avaliações e, por outro lado, na utilização considera-se a relevância das descobertas da avaliação para questões que preocupam os envolvidos nos programas e políticas.

Neste trabalho, entende-se por utilização das avaliações o emprego dos resultados das mesmas – sejam informações, interpretações, dados ou recomendações – para a tomada de decisão governamental. Como afirmado por diferentes autores (Leviton & Hughes, 1981; Shulla & Cousins, 1997; Patton, 2008), acredita-se que a utilização é um tema complexo, relacionado a diversas variáveis, que podem influenciar de forma positiva ou negativa o uso.

Discutidos os conceitos de utilização, foca-se nas diferentes dimensões da utilização. A utilização, como será vista a seguir, admite diferentes formas de usos, desde os mais tímidos, até aqueles que causam verdadeiras mudanças nos programas e pessoas envolvidos em uma avaliação.

3. AS DIMENSÕES DO USO DAS AVALIAÇÕES

Quando os primeiros artigos sobre utilização das avaliações foram escritos no início da década de 1970 entendia-se, unicamente, que o uso das avaliações estaria relacionado com a tomada de decisão. Nesse sentido, como afirmado por Weiss (1972), entendia-se por utilização o uso dos resultados das avaliações para a tomada de decisão nos programas avaliados. Para a autora, como as avaliações são responsáveis por apontar recomendações sobre os processos ou os resultados de um programa, os avaliadores esperam que essas recomendações sejam aplicadas ao programa avaliado para modificá-lo, estendê-lo ou até finalizá-lo.

Contudo, conforme afirmado por Shulha e Cousins (1997), na década de 1980, já havia um consenso de que o uso, ao contrário de um conceito unitário, constitui um fenômeno com variadas extensões, que podem ser descritas por diferentes dimensões apresentadas por Leviton e Hughes, em 1981. Os autores distinguem três tipos de uso das avaliações, quais sejam, instrumental, conceitual e simbólico (ou político).

A dimensão instrumental diz respeito aos casos nos quais as conclusões das avaliações são utilizadas fundamentalmente para a tomada de decisão ou para a solução de problemas dos programas (Leviton & Hughes, 1981). Dessa forma, o conceito de utilização estabelecido por Weiss em 1972 (utilização das avaliações para a tomada de decisão nos programas) constituía, assim, apenas uma das dimensões do uso, no caso, a dimensão do uso instrumental.

O uso instrumental é comum quando o avaliador entende bem o funcionamento do programa avaliado e suas questões, conduz a avaliação de forma adequada e comunica os resultados de forma satisfatória (Weiss, 1998). Para a mesma autora, no uso instrumental, observam-se três condições: (i) as conclusões das avaliações não são controversas, nem provocam conflitos de interesse na estrutura da organização; (ii) as mudanças a serem implementadas encontram-se dentro do escopo de atuação do programa e não provocam grandes rupturas no funcionamento do programa; e (iii) o ambiente do programa é relativamente estável, não apresentando dúvidas quanto ao seu orçamento, público-alvo, implementação, entre outros.

Uma quarta condição que favorece o uso instrumental trata da circunstância, do momento, vivido pelo programa quando ele recebe uma avaliação. Para Weiss (1998), quando o programa se encontra em uma situação de crise e ninguém sabe qual a melhor maneira para agir, a utilização instrumental de uma avaliação pode ser favorecida.

Uma segunda dimensão da utilização é o uso conceitual (Leviton & Hughes, 1981). O uso conceitual é circunscrito aos técnicos dos programas, cujo poder de decisão é limitado. Para Weiss (1998), mesmo que os técnicos sejam incapazes de utilizar os resultados das avaliações para a tomada de decisão no momento em que as conclusões das avaliações são apresentadas, as conclusões das avaliações podem modificar o entendimento desses técnicos sobre o que constitui o programa e como ele deveria funcionar. Dessa maneira, o uso conceitual, embora não se tenha uma decisão imediata decorrente de uma avaliação, possibilita a geração de novas ideias e direciona formas diferentes de ação para o programa.

Dito de outra maneira, a dimensão conceitual caracteriza-se pela capacidade de influenciar a opinião dos técnicos a respeito de um determinado programa, sem que exista qualquer ação em decorrência dos resultados de uma avaliação. Contudo, como afirmado por Weiss (1998), quando as condições organizacionais se tornam favoráveis, os técnicos podem utilizar as novas ideias e conceitos apreendidos para introduzir mudanças nos programas de maneira instrumental.

Quando a avaliação é utilizada na dimensão simbólica ou política, os tomadores de decisão já possuem conhecimento sobre o que seria necessário mudar em um determinado programa, e, nesse sentido, as avaliações servem para justificar o processo de mudança (Leviton & Hughes, 1981). Nessa dimensão, os coordenadores de programas já possuem conhecimento das necessidades de seus programas, suas reformulações e mudanças, que seriam implementadas de qualquer maneira, e as avaliações têm a função de legitimar o processo e ganhar adeptos, servindo como instrumento de persuasão (Weiss, 1998).

Além das três dimensões do uso citadas, colocadas por Leviton e Hughes, em 1981, (instrumental, conceitual e simbólica ou política), um debate extenso aponta a existência de mais dimensões. Patton (2008) coloca que, na medida em que o campo da avaliação foi sendo ampliado e novas discussões sobre utilização surgiram, distinções e tipologias adicionais emergiram.

Weiss (1998), por exemplo, aponta uma quarta dimensão do uso das avaliações, conceituada como “*enlightenment*”. Para a autora, as avaliações podem ser utilizadas para influenciar outras instituições ou eventos que vão além do escopo de determinado programa avaliado. A quarta dimensão justifica-se, uma vez que as conclusões de uma avaliação podem ultrapassar seus limites originais e pautar, não somente os coordenadores de programas avaliados, mas uma determinada comunidade acadêmica interessada nos resultados de dada política pública e coalizões políticas. Também é possível que os resultados das avaliações mudem a maneira de pensar sobre determinada política pública ou política social, alterando paradigmas sobre os benefícios da mesma (Weiss, 1980). Nesse caso, a avaliação acarreta impacto sobre os formadores de opinião, bem como alterações nas crenças e na forma de ação de instituições modificando, assim, conhecimentos sobre a formulação de políticas e da agenda governamental.

A quarta dimensão da avaliação, contudo, não ocorre de forma banal (Weiss, 1998). Mas, como argumentado pela autora, a acumulação de conhecimento é uma importante força na formulação de políticas. Quando uma avaliação contribui para o acúmulo de conhecimento, pode contribuir para grandes escalas de modificação de pensamento e, em última instância, para mudanças nos cursos de ação das políticas. Nem todos os avaliadores aspiram a esse tipo de influência para os resultados das avaliações, uma vez que representa uma modificação maior que a instrumental. De acordo com Patton (2008), o termo influência chama a atenção para os efeitos das avaliações ao longo do tempo e além da especificidade do programa avaliado.

Patton (2008) adiciona mais uma forma de utilização, denominada pelo autor de uso processual. Para o autor, a utilização pelo processo ocorre quando pessoas envolvidas nas avaliações (com exceção dos avaliadores) demonstram alterações comportamentais e cognitivas, como resultado da participação no processo de avaliação. Essas mudanças são reflexo, portanto, da aprendizagem experimental adquirida ao longo da realização de uma avaliação e da convivência com avaliadores, apreendendo seus valores e raciocínio. Em síntese, o uso processual envolve “aprender a pensar como avaliador” (Johnson, 1998, p. 93) e pode ter retornos positivos no longo prazo como, por exemplo, por meio de habilidades aprimoradas, entendimento da linguagem de avaliação, maior aproveitamento dos resultados das avaliações, melhor tomada de decisão e propriedade e segurança no acompanhamento de novas avaliações. O uso pelo processo, dessa forma, se sobrepõe ao uso instrumental e conceitual.

Além do uso processual, Patton (2008), adiciona mais quatro tipologias de utilização: legitimativo, persuasivo, imposto e mecânico. O uso legitimativo ocorre quando os resultados da avaliação são usados para suportar e justificar uma decisão que seria implementada antes mesmo de a avaliação ser realizada. O uso persuasivo refere-se à utilização seletiva dos resultados da avaliação para suportar decisões específicas sobre financiamento e debates políticos. O uso imposto acontece quando indivíduos do alto escalão, com maior nível de autoridade, obrigam a utilização àqueles situados em um nível hierárquico menor. Por fim, o uso mecânico, ou uso por conformidade, refere-se ao simples cumprimento das exigências de uma avaliação, uma vez que a implementação dos resultados de uma avaliação ocorre de forma forçosa, condicionada, por exemplo, a aumentos de orçamento para o programa que o fizer.

Assim, Patton (2008), ao reunir as tipologias clássicas às mais atuais, cria uma classificação para as possíveis dimensões do uso, quais sejam: (1) uso direto – instrumental, conceitual e processual; (2) uso no longo prazo – *enlightenment* e influência; e (3) uso político – simbólico, legitimativo, persuasivo, imposto e mecânico, como afirmado no quadro abaixo.

Quadro 1 – Diferentes usos das avaliações

Categoria	Tipos de Uso	Descrição
Uso direto	Instrumental	Quando os tomadores de decisão utilizam os resultados da avaliação para modificar o objeto da avaliação, de alguma forma.
	Conceitual	A avaliação possui capacidade de influenciar a opinião dos técnicos a respeito de um determinado programa, sem que exista qualquer ação em decorrência dos resultados de uma avaliação.
	Processual	A avaliação possui a capacidade de impor mudanças advindas do engajamento no processo de avaliação e no aprendizado de “pensar como avaliador”.
Uso no longo prazo	<i>Enlightenment</i>	A avaliação é utilizada para influenciar outras instituições ou eventos que vão além do escopo de determinado programa avaliado.
	Influência	A avaliação possui <i>unaware/unintended</i> impactos, ou seja, impactos não previstos ou planejados antecipadamente pelos avaliadores.
Uso político	Simbólico	A avaliação é utilizada de forma simbólica ou política; os tomadores de decisão já possuem conhecimento sobre o que seria necessário mudar em um determinado programa.
	Legitimativo	Os resultados de uma avaliação são utilizados como forma de justificar uma decisão que seria tomada.
	Persuasivo	Utilização seletiva dos resultados da avaliação para suportar decisões específicas sobre financiamento e debates políticos.
	Imposto	Indivíduos do alto escalão, com maior nível de autoridade, obrigam a utilização das avaliações àqueles situados em um nível hierárquico menor.
	Mecânico	Simple cumprimento das exigências de uma avaliação, uma vez que a

		implementação dos resultados de uma avaliação ocorre de forma forçosa, condicionada, por exemplo, a aumentos de orçamento para o programa que o fizer.
--	--	--

Fonte: Adaptado de Patton (2008).

Torna-se importante ressaltar que uma dimensão raramente é encontrada sozinha, como afirmado por Weiss (1998). Muitas vezes é difícil identificar a dimensão de uso em que uma avaliação é encontrada, já que os diferentes tipos de uso podem ser permutados em um mesmo estudo avaliatório.

Descritos os conceitos de utilização de avaliação e mencionadas as diferentes dimensões do uso, a seção seguinte apresenta os fatores que influenciam a utilização das avaliações nos programas e instituições. Nesse sentido, como se segue, serão apresentados os critérios e as circunstâncias que afetam a utilização das avaliações, na visão de diferentes autores.

4. OS FATORES QUE INFLUENCIAM A UTILIZAÇÃO DAS AVALIAÇÕES

O debate sobre utilização das avaliações, inaugurado na década de 1970, expôs o *calcanhar de Aquiles* das avaliações de políticas públicas – as avaliações não estavam sendo utilizadas como deveriam (Henry, 2000). Dahler-Larsen (2005) argumenta que o crescimento, a difusão e a institucionalização dos procedimentos de avaliação foram realizados antes mesmo que o problema mais fundamental sobre a utilização das avaliações fosse resolvido. Weiss (1972), ao listar uma série de casos em que as avaliações não estavam sendo utilizadas para a tomada de decisões, já declarava na época que “[...] a experiência no campo sugere que até os melhores estudos de avaliação de programas acabavam como lixo de uma usina burocrática” (Weiss, 1972, p. 11, tradução da autora da dissertação). Nesse período, buscavam-se modelos para a mensuração dos diferentes fatores e variáveis que influenciavam a utilização das avaliações. Como consequência, inúmeros estudos começaram a desenvolver teorias e desenhar estratégias na tentativa de sanar a não utilização, a má utilização e o abuso das avaliações.

Neste trabalho destacam-se três estudos principais que contribuíram para a pesquisa sobre as variáveis que influenciam a utilização das avaliações: (1) o estudo de Cousins e Leithwood (1986), que tiveram por objetivo identificar os estudos empíricos sobre uso das avaliações conduzidos entre 1971 e 1985; (2) o trabalho de Shulha e Cousins (1997), que conduziram um trabalho semelhante ao já realizado em um período diferente, dessa vez de 1986 à 1997; e, por fim, (3) o estudo de um grupo de pesquisadores (Johnson *et al.*, 2009), que se dedicou a uma pesquisa sobre a utilização das avaliações nos principais periódicos de avaliação em um período de 25 anos, compreendido entre 1986 a 2005.

4.1. Cousins e Leithwood (1986): A Importância do Contexto

Para Cousins e Leithwood (1986), o conhecimento advindo da experiência prática fornece a base para a ação e a decisão na maioria das organizações. O conhecimento derivado de estudos e métodos aplicados às ciências sociais, por outro lado, tende a ser seletivo, pouco compartilhado e, muitas vezes, apenas serve de complemento ao conhecimento prático. Motivados por esse pensamento, os autores buscaram revisar as variáveis (dependentes e independentes) que influenciavam a utilização de informações de uma avaliação.

Como definição de utilização das avaliações, Cousins e Leithwood (1986) consideraram duas variáveis dependentes: uso de avaliações como apoio para decisões (uso instrumental) e

uso educativo das avaliações para o aprendizado dos tomadores de decisão (uso conceitual). Como variáveis independentes, por sua vez, os autores consideraram os fatores que influenciam os tomadores de decisão a utilizar as avaliações. Dos 65 estudos identificados, Cousins e Leithwood (1989) observaram que, independentemente do tipo de uso da avaliação (instrumental ou conceitual), diversas foram as orientações coerentes com a visão de que o uso corrente das avaliações era limitado.

O exame realizado pelos autores resultou em duas categorias de fatores relacionados à utilização das avaliações: (i) características de implementação e (ii) características da decisão ou do *policy setting*. As características de implementação foram desdobradas em seis: qualidade da avaliação, credibilidade, relevância, qualidade da comunicação, resultados e tempestividade. As características da decisão ou do *policy setting*, por sua vez, também foram desdobradas em seis: necessidade da informação, características da decisão, clima político, informação competitiva, características pessoais e comprometimento e receptividade à avaliação.

De acordo com Cousins e Leithwood (1986), por qualidade das avaliações entende-se a sofisticação metodológica, o tipo de abordagem utilizada ao problema da avaliação, ou a intensidade das atividades avaliativas. No entanto, as pesquisas do período não chegam a um consenso. Enquanto alguns autores (Yeh, 1980) apontavam que métodos menos sofisticados proporcionam um uso melhor das avaliações, outras linhas apontavam que a robustez metodológica tem uma relação positiva com o uso (Patton *et al.*, 1977). Outros autores, como Leviton e Boruch (1983), argumentaram que estudos avaliatórios focados em desenho ou processos seriam mais bem utilizados que estudos focados em resultados dos programas. Por fim, os autores mostram que os estudos avaliatórios que fizeram uma pesquisa mais ampla, com maior número de testes e/ou observações, apresentaram melhor utilização.

Quanto à variável credibilidade, os autores observaram que quando a avaliação era entendida pelos tomadores de decisão como um instrumento válido e de singular importância para as atividades desempenhadas pela organização, o uso, ou o potencial para a utilização parecia maior (Cousins & Leithwood, 1986).

A relevância das avaliações, ou seja, o grau de importância da avaliação para os públicos que a receberam (Cousins & Leithwood, 1986), mostrou-se um fator que influencia fortemente o uso das avaliações. Os estudos analisados pelos autores demonstraram que aquelas avaliações que refletem o conhecimento do contexto em que as mesmas estavam inseridas, seriam mais bem utilizadas que aquelas afastadas do contexto da organização em que se inserem. Como medida da relevância da avaliação, a posição organizacional do avaliador (interna ou externa), tem direta relação com a utilização. A presença de avaliadores internos sugeriria uma maior utilização, para os mesmos apreenderem maior conhecimento sobre as características dos programas avaliados e a cultura da organização em que estão inseridos.

Embora um terço dos estudos indicasse que a qualidade da comunicação na apresentação dos resultados das avaliações e a amplitude da divulgação apresentam alguma relação com o uso, os mesmos não se mostraram conclusivos (Cousins & Leithwood, 1986). Observou-se que mesmo que houvesse apresentações orais dos resultados e recomendações das avaliações (juntamente com os relatórios), com uma linguagem não-técnica; esse fato, no entanto, pouco contribuiria para a tomada de decisão em um programa. Em contraste, a utilização pareceu estar fortemente relacionada à comunicação permanente e à proximidade (geográfica) entre avaliador e tomador de decisão.

Quanto à variável resultados, os estudos analisados destacaram que resultados esperados estariam mais relacionadas à aceitação e à utilização por parte das equipes dos programas, enquanto resultados surpreendentes das avaliações estariam fadados ao não uso (Cousins &

Leithwood, 1986). Recomendações mais práticas e conclusivas também se mostraram mais receptivas ao uso que resultados gerais e pouco diretos.

Por fim, como última variável relacionada às características de implementação, a questão tempo, ou seja, o fornecimento oportuno e tempestivo dos resultados das avaliações foi positivamente relacionado à utilização em menos de 15% dos estudos analisados (Cousins & Leithwood, 1986). Patton *et al.* (1977), no entanto, mencionaram que o atraso em prazos e na entrega dos relatórios de avaliações pouco influencia a utilização das mesmas.

Na parte direcionada às características da decisão ou do *policy setting*, as variáveis categorizadas de acordo com as necessidades de informação foram variadas (Cousins & Leithwood, 1986). Elas incluíram tanto a intensidade da necessidade, como o tipo de informação necessária e variâncias em relação ao público solicitador das informações. Quanto à intensidade da necessidade, observou-se que quanto maior precisão, melhor a utilização. Quanto ao tipo de necessidade, verificou-se que necessidades de informação voltada para a ampliação de programas apresentam uso mais positivo das informações. E quanto ao público, a maior variedade de audiências quanto aos resultados das avaliações mostrou influenciar de forma negativa os resultados das avaliações.

Mais de um terço dos estudos relatou relação positiva entre a utilização e as características da decisão, como, por exemplo, área, contexto e significado da decisão. Cousins e Leithwood (1986) relataram que a maior parte dos trabalhos analisou a utilização para diferentes áreas da decisão, sendo que nas áreas em que as avaliações apontavam necessidades como capacitação, pontos fracos do programas e questões ligadas à administração do programa, o uso foi caracterizado como elevado.

A importância do clima político foi considerada tanto em nível organizacional, como extra-organizacional (Cousins & Leithwood, 1986). De acordo com os autores, a utilização diminui caso seus resultados sejam percebidos como uma ameaça pelos funcionários da organização. Como exemplo de influência extraorganizacional, os autores citam o exemplo de exigências de realização de avaliações para o financiamento de projetos como um dos fatores que influenciam positivamente o uso.

Quando confrontados com um problema, os tomadores de decisão confiam em várias fontes de informações diferenciadas, além dos resultados das avaliações. Por isso, Cousins e Leithwood (1986) consideraram a variável informação competitiva, indicando que fontes alternativas de informação competindo com os resultados da avaliação para a tomada de decisão também influenciam o uso. O conhecimento prático de pessoas envolvidas nos programas, crenças, valores e interesses, por exemplo, mostraram-se concorrentes poderosos com os resultados das avaliações.

Mais de um terço dos estudos analisados por Cousins e Leithwood (1986) relataram relações entre as características pessoais do tomador de decisão, como liderança, experiência, treinamento e papel na organização, e utilização das avaliações. No entanto, de acordo com os autores, as informações são conflitantes. Poucos estudos apoiaram a relação entre posição do tomador de decisão na hierarquia da organização e a adequada utilização das avaliações, ou seja, a ideia de que uma posição mais alta na hierarquia seria favorável ao uso. A relação entre fatores pessoais, como experiência, também se mostrou ambígua; a experiência no campo contribui mais para diferentes formas de processamento das informações. Características de liderança mostraram-se positivamente associadas ao uso das avaliações.

O comprometimento e a receptividade à informação da avaliação foram definidos em termos do grau em que o tomador de decisão estava envolvido na avaliação, sendo que 25% dos estudos analisados afirmaram que existe uma relação direta entre a participação e o

acompanhamento do processo avaliatório e a adequada utilização das avaliações (Cousins & Leithwood, 1986).

Cousins e Leithwood (1986) revisaram as pesquisas sobre utilidade das avaliações no período de 1971 a 1986. Posteriormente, outros autores dedicaram-se a trabalho semelhante, porém, em períodos diferenciados e com outro conjunto de teóricos, o que culminou na discussão e no surgimento de informações sobre outras variáveis que também poderiam influenciar a utilidade. O trabalho de Shulha e Cousins (1997) é um deles.

4.2. Shulha e Cousins (1997): O Papel da Política e das Estruturas e Processos Organizacionais

De acordo com Shulha e Cousins (1997), em 1996, a pesquisa desses autores foi responsável não somente por marcar o 10º aniversário da Associação Americana de Avaliação (American Evaluation Association – AEA), como também prestigiar uma década de atividades de estudos e trabalhos interessados na natureza, causas e consequências da utilização das avaliações. Nesse momento, na seção nomeada “*The past, present and future of evaluation use: the continuing evolution of evaluation practice and theory*”, apresentou-se a revisão dos trabalhos de utilização, no período de 1986 e 1996, feito pelos autores.

A pesquisa de utilização após 1986, para Shulha e Cousins (1997), já havia conseguido provar a complexidade do tema do uso das avaliações, dos programas e dos contextos políticos em que a avaliação ocorreu. Nesse ponto de desenvolvimento da pesquisa, os teóricos do campo já haviam reconhecido que o “projeto de avaliações úteis” exigia muito mais do que simplesmente a combinação de perguntas dos usuários das avaliações sobre os métodos adequados, abordagens ou modelos de avaliação.

Shulha e Cousins (1997) argumentam que os estudos que tentaram catalogar as variáveis que influenciariam a utilização, no período de 1986 e 1996, podem ser divididos em três fases principais.

A primeira delas leva em conta o papel do contexto em que as avaliações estão inseridas. Alguns teóricos de destaque, como Guba e Lincoln (1989), colocaram que problemas comuns podem comprometer a utilização das avaliações. Entre esses problemas, os autores citam a hegemonia do paradigma positivista, a incapacidade em acomodar o pluralismo de atores envolvidos nos programas públicos e a desconsideração do contexto, das circunstâncias e propósito que envolve a implantação de programas. Os autores desafiaram a confecção de avaliações tradicionais e propuseram um método construtivista para a realização das mesmas, uma vez que, na percepção dos autores, a construção de conhecimento é parte integrante do contexto em que ele é produzido.

Uma segunda perspectiva colocada por Shulha e Cousins (1997), que marcou a pesquisa de utilização de avaliações, enfatiza que a atividade política também estaria relacionada ao contexto. Nesse veio, a tomada de decisão alinhada às recomendações de uma avaliação não dependeria somente de um único ator, mas de outros grupos de interesse e circunstâncias do próprio setor público.

Uma terceira abordagem dos estudos de utilização de avaliação de políticas relaciona-se à estrutura e aos processos organizacionais em que as avaliações estão inseridas (Shulha & Cousins, 1997). Por meio dessa perspectiva, consideram-se os canais de comunicação da organização e a burocracia de processos que, em menor ou maior grau, influenciam, diretamente, a transmissão das informações geradas pelas avaliações e a inserção de mudanças no nível organizacional.

Outros pesquisadores focaram o papel da interação entre avaliadores e o contexto do programa em questão. Um dos marcos mais importantes para o campo, de acordo com Shulha e Cousins (1997), consiste no debate entre dois autores centrais: Carol Weiss e Michael Patton. Com base na discussão desses autores, o papel dos avaliadores resume-se em assegurar a utilização das avaliações.

O debate entre Weiss e Patton (*Weiss-Patton debate*) contribuiu de forma significativa para o desenvolvimento da teoria de utilização, como afirmado por Shulha e Cousins (1997). As discussões entre os autores, iniciadas em 1987 na AEA, posteriormente traduziram-se em uma série de artigos do *American Journal of Evaluation* e se tornou um marco no campo. O ponto de partida do debate foi o clássico trabalho de Weiss (1988) “*Is anybody there? Does anybody care?*”, cujo foco central aborda a fraca utilização dos resultados das avaliações em mudanças concretas nos programas e ações governamentais, ou seja, a ausência de tomada de decisão advinda de estudos avaliativos contratados para esse fim. A constatação da autora baseava-se em cinco argumentos centrais: (i) natureza política dos processos de decisão; (ii) as avaliações não são capazes de dar respostas pragmáticas alinhadas aos anseios dos tomadores de decisão; (iii) as decisões não são tomadas com base em um processo racional; (iv) muitas vezes não se conhece o tipo de informação necessária para se tomar decisões; e (v) interesses diversos protegem a continuidade do *status quo* do programa avaliado.

De acordo com Shulha e Cousins (1998), Patton que, no mesmo momento, discordou dos argumentos da autora, ressaltou que, embora as conclusões de Weiss fossem válidas para análise do processo decisório, em se tratando de avaliações, outros argumentos deveriam ser levados em conta. Na visão do autor, a perspectiva de Weiss (1988) seria limitada, uma vez que a autora se restringia à utilização das avaliações unicamente direcionadas à tomada de decisão.

O debate entre os autores Patton e Weiss continuou de forma mais aprofundada, centrado na questão do papel dos avaliadores (Shulha & Cousins, 1997). Para ela, como os processos decisórios no setor público são caracterizados pela irracionalidade e pelo pluralismo, seria papel dos avaliadores encontrarem a melhor maneira – método e abordagem avaliativa – de auxiliar os usuários das avaliações. Para ele, os avaliadores teriam um papel mais ativo na promoção e no cultivo do uso, além de auxiliar os decisores quanto à abordagem da avaliação. Enquanto o impasse não era resolvido, a importância do contexto em que os avaliadores estavam inseridos era cada vez mais valorizada pelos teóricos do campo, como afirmado por Shulha e Cousins (1997).

Após Cousins e Leithwood (1986) e Shulha e Cousins (1997), Johnson *et al.* (2009) revisam as pesquisas sobre a utilidade das avaliações, no período de 1986 a 2005, como será apresentado na seção seguinte.

4.3. Johnson *et al.* (2009): A Importância do Envolvimento dos Stakeholders

Johnson *et al.* (2009) conduziram uma pesquisa sobre a utilização das avaliações em um período de 25 anos, compreendido entre 1986 a 2005. Sua pesquisa constituiu na revisão de periódicos de avaliação – *American Journal of Evaluation*, *Canadian Journal of Program Evaluation*, *Evaluation*, *Evaluation Practice*, *Evaluation and Program Planning*, *Evaluation Review*, *New Directions for Evaluation*, and *Studies in Educational Evaluation* – dissertações e capítulos de livros, baseados em termos como “*evaluation use*”, “*evaluation utilization*” e “*evaluation influence*”. Após a revisão da pesquisa, a equipe de pesquisadores identificou 98 publicações que obedeciam aos critérios de análise e, destes, apenas 41 estudos foram considerados pelos autores como adequados.

Como sugerido por Shulha e Cousins (1997), mudanças nas conceptualizações sobre o uso das avaliações ocorreram e, assim, novas categorias apareceram. No estudo de Johnson *et al.* (2009), uma nova característica – competência do avaliador – foi adicionada à categoria de implementação das avaliações de Cousins e Leithwood (1986). O estudo também foi responsável por criar uma nova categoria completa – envolvimento dos *stakeholders* – para acomodar 25 estudos examinados que não correspondiam aos critérios observados por Cousins e Leithwood (1986).

A característica de competência do avaliador foi adicionada, de acordo com os autores, pois seis estudos apontaram que os profissionais da avaliação desempenham importante papel na condução da utilização das mesmas, por diferentes razões. Embora o estudo de Cousins e Leithwood (1986) apontasse a reputação dos avaliadores na característica de credibilidade, a definição não compreendia a natureza influenciadora do avaliador, sua competência ou liderança no momento de utilizar os resultados de uma avaliação.

A nova categoria envolvimento dos *stakeholders* foi adicionada devido a estudos mais recentes, como apontado por Johnson *et al.* (2009). Dos estudos analisados, mais da metade (23 de 41) apresentaram alguma preocupação com o envolvimento de atores.

5. CONSIDERAÇÕES FINAIS

Apesar de ser clara a percepção de que a prática da avaliação de programas pode levar à transformação nas estruturas de conhecimento de organizações públicas, ou seja, na forma como ela detecta e corrige seus erros, pouca atenção tem sido dada à utilização das informações e resultados de programas governamentais, principalmente, no Brasil. Tal fato merece atenção, uma vez que, as avaliações constituem, antes de tudo, referências complexas, que podem ser utilizadas para mudanças de atitudes ou para o aprimoramento de programas e projetos.

Embora tenham ganhado importância considerável nos últimos tempos, torna-se importante considerar que as avaliações de políticas públicas, normalmente, são realizadas com base na premissa de que os resultados serão direta e automaticamente utilizados para promover as melhorias necessárias – o que nem sempre ocorre. A existência de avaliações por si só não garante o melhoramento da gestão pública, uma vez que a viabilidade de um sistema para promover uma cultura de verificação de resultados e fortalecimento de instituições depende, em grande parte, da utilização da informação gerada.

Várias questões importantes podem ser inferidas com base neste breve mapeamento da literatura que trata do uso das avaliações de políticas públicas. Em primeiro lugar, infere-se que a utilidade real das avaliações envolve uma série de processos e fatores, do início ao fim, que irão afetar o uso e a aplicação dos resultados de estudos avaliatórios, sejam esses fatores relacionados ao contexto em que as avaliações estão inseridas, à estrutura organizacional predominante ou ao envolvimento dos diversos *stakeholders* no processo.

Tendo como referência os estudos de Cousins e Leithwood (1986), Shulha e Cousins (1997) e Johnson *et al.* (2009), a utilização das avaliações parece ser mais fortemente evidente quando:

- Os tipos e abordagens avaliatórias foram adequadamente selecionados, juntamente com a metodologia e a intensidade em que a avaliação ocorreu;
- As decisões a serem tomadas tinham relevância para os usuários e necessitavam de coleta de dados analisadas por avaliações;
- Os resultados das avaliações foram condizentes com as crenças e expectativas dos usuários;

- Os usuários e diferentes *stakeholders* foram envolvidos no processo de avaliação e possuíam um compromisso com os resultados anteriores ao processo avaliatório;
- Os usuários consideraram as informações relatadas pelas avaliações relevantes para suas preocupações e problemas;
- Os avaliadores possuem uma interação contínua com os usuários das avaliações, com papel ativo na promoção e no cultivo do uso;
- As estruturas e os processos organizacionais em que as avaliações estão inseridas, incluindo canais livres de comunicação, são receptíveis a mudanças.

Assim, observa-se que o uso apropriado dos resultados das avaliações não constitui tarefa fácil. A questão da utilização das avaliações deve tornar-se uma preocupação relevante de avaliadores, formuladores de políticas públicas e cidadãos, pois, do contrário, as avaliações constituirão apenas um universo de promessas que não chegam a se realizar, como afirmado por Ala-Harja e Helgason (2000).

Outra questão que merece destaque diz respeito à necessidade de se criarem organizações, que prezem por padrões para a realização de avaliações de políticas públicas e meta-avaliações (avaliações das avaliações), no intuito de ajudar a avaliação a realizar o seu potencial. Apesar do crescente interesse em avaliações de programas governamentais no Brasil, o país ainda mostra-se carente nesse aspecto, sendo escassos os estudos nacionais que permitam verificar a utilização das avaliações como instrumento de correção de rumos e aprimoramento dos serviços públicos.

Por fim, a expectativa é de que o arcabouço conceitual aqui construído, longe de ser definitivo, ajudará a analisar como as avaliações tendem a modificar as instituições públicas e a dificuldade existente em se utilizar, de forma apropriada, seus resultados. É possível perceber, à primeira vista, uma necessidade de integração dessas abordagens para analisar a utilização das avaliações face ao ambiente e à capacidade institucional de organizações públicas.

REFERÊNCIAS

Aguilar, M. J., & Ander-Egg, E. (1995). *Avaliação de serviços e programas sociais*. Rio de Janeiro: Vozes.

Ala-Harja, M., & Helgason, S. (2000). Em direção às melhores práticas de avaliação. *Revista do Serviço Público*, Brasília, 51(4), 5-59.

Alkin, M. C. (2004). *Evaluation Roots: tracing theorists views and influences*. California: Sage.

Arretche, M. T. S. (1998). Tendências no Estudo sobre Avaliação. In: E. M. RICO (Orgs.). *Avaliação de Políticas Sociais: uma questão em debate*. São Paulo: Cortez.

Bamberger, M., Rugh, J., & Mabry, L. (2011). *Real World Evaluation: Working Under Budget, Time, Data, and Political Constraints*. Thousand Oaks, California: Sage.

Cousins, J. B., & Leithwood, K. A. (1986). Current Empirical Research on Evaluation Utilization. *Review of Educational research*, 56(3), 331-364.

- Dahler-Larsen, P. (2005). Evaluation and Public Management. In: E. Ferlie, L. Lynn Jr, , & C. Pollitt. *The Oxford Handbook of Public Management*. Oxford: Oxford University Press.
- Faria, C. A. P. (2005). A Política da avaliação de políticas públicas. *Revista Brasileira de Ciências Sociais*, 20(59), 97-169.
- Guba, E. G. (1975). Problems in utilizing the results of evaluation. *Journal of Research and Development in Education*, 8(3), 42-54.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Newbury Park: Sage.
- Yeh, J. (1980). A reanalysis of test use data (CSE Report No. 143). Los Angeles: University of California, Center for the Study of Evaluation.
- Henry, G. T. (2000). Why not use? In: V. J. Caracelli, & H. Preskill (Org.) *The expanding scope of evaluation use. New Directions for Evaluation*. San Francisco: Jossey-Bass, 85-98.
- Johnson, K., Greenesid, L. G., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on Evaluation Use: A Review of the Empirical Literature From 1986 to 2005. *American Journal of Evaluation*, 30(3), 377-410.
- Johnson, R. B. (1998). Toward a theoretical model of evaluation utilization. *Evaluation and Planning*, 21, 93-110.
- Leviton, L. C., & Boruch, R. F. (1983). Contributions of evaluation to education programs and policy. *Evaluation Review*, 7(5), 563-598.
- Leviton, L. C., & Hughes, E. F. X. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review*, 5(4), 525-548.
- Lobo, T. (2001) A avaliação de processos e impactos em programas sociais: algumas questões para reflexão. In: E. M. RICO (Orgs.). *Avaliação de políticas sociais: uma questão em debate*. São Paulo: Cortez.
- Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1977). In search of impact: an analysis of the utilization of federal health evaluation research. In: C. H. Weiss (ed.). *Using Social Research in Public Policy Making*. Lexington: Lexington Books.
- Patton, M. Q. (2008) *Utilization-Focused Evaluation*. California: Sage.
- Serpa, S. M. H. C. (2010) *Para que avaliar? Identificando a Tipologia, os Propósitos e a Utilização das Avaliações de Programas Governamentais no Brasil*. Dissertação de mestrado em Administração, Universidade de Brasília, Brasília, BsB, Brasil.
- Shulha, L. M., & Cousins, J. (1997). Evaluation use: theory, research, and practice since 1986. *EvaluationPractice*, 18(3), 195-209.

Trevisan, A. P., & Van Bellen, H. M. (2008). Avaliação de políticas públicas: uma revisão teórica de um campo em construção. *Revista de Administração Pública*, Rio de Janeiro, 42(3), 529-550.

Weiss, C. H. (1972). *Utilization of evaluation: toward comparative study*. Boston, MA: Allyn and Bacon.

Weiss, C. H. (1978). Improving the linkage between social research and public policy. In: L. E., LYNN (ed.). *Knowledge and Policy: The Uncertain Connection*. Washington, DC: National Research Council.

Weiss, C. H. (1980). Knowledge creep and decision accretion. *Knowledge: creation, diffusion, utilization*, 1(3), 381-404.

Weiss, C. H. (1988). Evaluation for decisions: Is anybody there? Does anybody care? *Evaluation Practice*, 9(1), 5-19.

Weiss, C. H. *Evaluation*. (1998). Upper Saddle River: Prentice-Hall.

Worthen B. R.; Sanders J.R.; Fitzpatrick J. L. (2004). *Avaliação de programas: concepções e práticas*. São Paulo: Gente.

Yeh, J. (1980). *A reanalysis of test use data* (CSE Report n. 143). Los Angeles: University of California, Center for the Study of Evaluation.