Cadernos EBAPE.BR
ISSN: 1679-3951
Fundação Getulio Vargas, Escola Brasileira de Administração Pública e de Empresas

Ribeiro de Oliveira, Lilian; Souza Passador, Claudia
Ensaio teórico sobre as avaliações de políticas públicas
Cadernos EBAPE.BR, vol. 17, núm. 2, 2019, Abril-Junho, pp. 324-337
Fundação Getulio Vargas, Escola Brasileira de Administração Pública e de Empresas

DOI: 10.1590/1679-395169657

# Theoretical essay on public policy evaluations

Lilian Ribeiro de Oliveira [1]
Claudia Souza Passador [1]

[1] Universidade de São Paulo (USP) / Faculdade de Economia, Administração e Contabilidade, Ribeirão Preto – SP, Brazil

*Abstract*

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

**Keywords**: Theoretical Framework. Evaluation. Public Policies Evaluation.

## *Ensaio teórico sobre as avaliações de políticas públicas*

*Resumo*

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

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

## *Ensayo teórico sobre las evaluaciones de políticas publicas*

*Resumen*

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

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

Article submitted on June 20, 2017, and accepted for publication on November 13, 2018.
[Translated version] Note: All quotes in English translated by this article's translator..
DOI: http://dx.doi.org/10.1590/1679-395169657

## INTRODUCTION

Although widely discussed in Brazil and in the world, the term "public policy" may adopt several meanings. According to Secchi (2014), latin-speaking countries find it difficult to distinguish two terms from the English language: politics and policy. For Bobbio, Matteucci and Pasquino (1998), the first refers to a set of human activities pertaining to the *polis* or to the State; and, according to Secchi (2014), the second denotes something more concrete, that has relation with actions, and decisions. Thus, "public policy" is aligned with the meaning of policy, i.e., it refers to the process of political construction, acting, and decision. It may be said that public policy is closely linked to state actions, however, there are other currents of thought that indicate a multi-centric scenario (SECCHI, 2014). In the multi-centered approach, there appears a network of public policy made up of government actors, private and non-governmental organizations, and multilateral organizations (FREY, 2000; SECCHI, 2014). From the conceptual point of view, besides involving several actors in the formulation or implementation, public policies are multidisciplinary and have an impact on the economy, and societies (SOUZA, 2006; HOWLETT, RAMESH and PERL, 2013). In view of the aforementioned definitions, we may infer that the study of public policy is a complex phenomenon with a multiplicity of actors internal or external to the State (HOWLETT, RAMESH and PERL, 2013). Consequently, the search for analytical processes is a way to reduce the complexity of the object and to simplify the analysis. In order to do so, the division into stages was established, under the aegis of the policy cycle and in relation to the various phases of the political process arranged in a sequential and interdependent way (FREY, 2000; HOWLETT, RAMESH and PERL, 2013; SECCHI, 2014). Among the formats observed in the literature, we have chosen to describe the applied resolution of problems and the stages corresponding to the political cycle, based on Howlett, Ramesh and Perl (2013) and Secchi (2014) – Box 1 summarizes these connections.

**Box 1**

**Applied resolution of problems and its relation with the political cycle**

| Applied problem resolution | Political cycle stage according to Howlett, Ramesh e Perl (2013) | Political cycle stage according to Secchi (2014) |
|---|---|---|
| Problem recognition | Agenda set-up | Problem identification |
| Solution proposal | Policy formulation | Agenda formation/alternatives formation |
| Choice of solution | Political decision taking | Decision taking |
| Solution effectuation | Policy implementation | Public policy implementation |
| Results monitoring | Policy evaluation | Public policy evaluation |
| End or extinction[1] | – | Public policy extinction |

[1]Studies on the extinction of public policies began to be prominent in developed countries in the 1970s. Those countries performed various social actions linked to the welfare state, which were put in check due to the resolution of a specific problem: inefficiency of action or loss importance, and exit from the political agenda (SECCHI, 2014).

Source: Elaborated by the authors.

Since it is a holistic field with several unfolding events, this study proposes to specifically analyze the historical process of structuring one stage of the political cycle, namely evaluation. Given its conceptual, methodological, and developmental importance, the evaluation of public policies, public programs, or public processes is included in the contemporary research agenda as an essential tool for the improvement of public policies in all spheres of government, and in the global scenario. The United Nations (UN), the World Bank, the Inter-American Development Bank (IDB), the Organization for Economic Co-operation and Development (OECD), the Economic Commission for Latin America and the Caribbean (ECLAC), and the Latin American Center for Public Administration and Development (Clad) emerge as protagonists in the generation of methodologies, and evaluative projects (RAMOS and SCHABBACH, 2012).

The effective use of evaluative processes can contribute to the transparency of public acts, as well as to present forms of control and monitoring of governmental actions to citizens, thus guaranteeing the legitimacy of the policies or programs developed (WEISS, 1999; ALA-HARJA and HELGASON, 2000; MOKATE, 2002; VEDUNG, 2009; RAMOS and SCHABBACH, 2012). Nevertheless, it is worth noting that, because of its political nature, evaluations can contribute to the maintenance or not of a certain policy, sometimes linked to the propositions and strategies of the parties involved (RAMOS and SCHABBACH, 2012; ARRETCHE, 2001; CRUZ, 2015).

Carvalho (1999) notes that the evaluation presents a high degree of complexity and specificity, factors that increase the difficulty in incorporating the evaluative activity in the day-to-day of the administrators, since there are multiple actors, objectives, causes, effects, and outcomes. In this sense, in order to contribute to the debate on public policy evaluation, this article devotes the next topics to the historical rescue of the study of the evaluation, as well as to the details of the existing models, types, and methodologies in the literature and recent international trends inserted in this discussion. For the construction of this theoretical essay, we opted for advanced search in the following databases: SCOPUS, Web of Science, Google Scholar, SciELO, USP Bank of Theses and Dissertations, VHL, and Capes Journal Portal. The main keywords of the research were: “public policies”; “evaluation”; and “public policy evaluation”. Nevertheless, this search also resulted in an amplification of the literature through the articles analyzed.

## THEORETICAL REVIEW

The literature considers the end of the 1950s as a cornerstone of the evaluative studies. According to Vedung (2010), the historical process of studies on evaluation in the global context can be divided into four diffusion waves. Each wave is indicated for a period of time and coined according to social, political, and economic issues. The construction of this metaphor refers to the constant renewal of the scenario resulting from the repetitive movement of the waves and the constant sediment deposits for the creation of a base (VEDUNG, 2010). Using this theoretical framework, we opt for describing the evolution of the evaluative studies using Vedung’s diffusion wave conception (2010).

The first wave, or diffusion phase of the evaluation studies, dates back to the late 1950s, consolidated in the mid-1960s (DERLIEN, 2001; TREVISAN and BELLEN, 2008; VEDUNG, 2010; PICCIOTTO, 2015), and is characterized as “the wave oriented to science” (VEDUNG, 2010). The studies belonging to this period in time are related to the emergence of the Welfare State, especially in the countries with reformist aspirations in that period, especially United States of America (USA), France, Germany, Sweden, and Canada (DERLIEN, 2001; VERDUNG, 2010; PICCIOTTO, 2015).

The main objectives of the studies of this phase were following up and analyzing the results of the policies implemented; Derlien (2001) attributes to this set of studies the information function. These are evaluations with the purpose of assessing the effects and results of the policy or program, as well as identifying possible positive or negative consequences. These studies also aimed to verify points to be improved and to indicate recurrent failures (DERLIEN, 2001). “The focus was on the improvement of the program, and administrators were interested in using evaluation as a feedback mechanism” (TREVISAN and BELLEN, 2008, p. 537). For Vedung (2010), it was during this period that evaluations emerged from a context of radical rationality; the structure of evaluative thinking was guided by the notion of “scientification” of processes of public policy, and public administration. The evaluation would bring to the government an air of rationality with a scientific and fact-based scaffold (ROSSI and WRIGHT, 1984; WEISS, 1999; VEDUNG, 2010). In his paper on public sector reform, Thoenig (2000) reinforces the rationality described by Vedung (2010) and confirms the information function proposed by Derlien (2001), stating that if the use of evaluation is oriented to action, it must concentrate on providing usable knowledge. For the author, the evaluation is much more likely to be accepted if its informational aspect is emphasized (THOENIG, 2000).

The second time period, or the second wave described by Vedung (2010), named “dialogue-oriented wave”, dates back to the mid-1970s, in a context in which researchers argued that assessments should be more pluralistic, driven by progressive ideals (VEDUNG, 2010; PICCIOTTO, 2015).

It is discussed the insertion of all the actors involved in the intervention to be analyzed, from researchers and politicians to operators of the policy or program in question (VEDUNG, 2010). Rossi and Wright (1984) emphasize the growth in the use of qualitative research in this period, a model that is convergent to the inclusive thinking of the time, justifying its use based on the lower costs of these surveys at the local level, and with greater utility and flexibility for administrators. These evaluative models would support better the participatory comprehensiveness and would bring greater contact with reality (ROSSI and WRIGHT, 1984; VEDUNG, 2010). Summing up the characteristics of this wave, Vedung (2010, p. 270, our translation) explains that, "in contrast to the science-induced wave, the dialogue-oriented wave rested on communicative rationality. Instead of producing truths, a dialogic evaluation would generate broad agreements, consensus, political acceptability, and democratic legitimacy".

On the other hand, the third wave, named by Vedung (2010) "neoliberal wave", introduced in the early 1980s, modified once again the evaluation studies standard. Vedung (2010) and Picciotto (2015) describe this period as a neoliberal flood in the field of evaluative studies, focusing on reducing the state, and promoting free market and public-private partnerships. In the USA, the new evaluative conceptions followed the dismantling of social programs promoted by Ronald Reagan (ROSSI and WRIGHT, 1984). The emerging management model was the New Public Management (NPM), with greater focus on results in relation to the processes. Vedung (2010, p. 270, our translation) complements by arguing that "decentralization, deregulation, privatization, civil society and, in particular, customer orientation have become new slogans. Previously considered the solution to the problems, the public sector had become the problem to be solved". Derlien (2001) attributes the mapping function to this wave, and the studies focus on diagnosing the policies or programs that can be cut based on the negative results presented and on the consequences of the privatization of certain state functions or activities; financial efficiency in the conduction of public policies and programs is aimed at (DERLIEN, 2001).

In their contribution, Ala-Harja and Helgason (2000) indicate that, in this phase, evaluation was understood as a tool capable of justifying policies and reallocating financial resources. At the governmental level, the evaluators focus on the Ministry of Finance and audit organs. In conclusion, Vedung (2010) summarizes that the evaluation activities of the third wave were mainly directed to accountability, performance measurement and consumption, search for the quality of services provided, and comparative evaluations.

Following, the fourth and last wave described by Vedung (2010) is called "wave of evidence: the return of experimentation". This wave of evaluative studies took shape in the decades of 1990 and 2000, especially in the North Atlantic and Nordic countries. These are evidence-based studies and, in this sense, Picciotto (2015) emphasizes that those involved in assessment studies are "surfing" this fourth wave. Studies focused on scientific evidence aim to separate what works from what does not work (VEDUNG, 2010; PICCIOTTO, 2015).

The authors highlight that studies with experimental approaches, by using randomization (control and treatment groups) or quasi-experimental, have as main objective the verification of the impacts caused by the public policy or program (VEDUNG, 2010; PICCIOTTO, 2015). Complementarily, Silveira, Vieira, Capobiango et al. (2013) resume the concept of evaluation that seeks to understand and know the consequences of an action, as well as to anticipate its possible results. In the meanwhile, international organizations, academics, and governments began to perform systematic reviews of the literature, in order to strengthen the research bases, guaranteeing greater rationality to the process. Vedung (2010) classifies this movement of incessant search for scientific evidence as a return to the science-based model, however, with a new style. In Derlien's view (2001, p. 107, our translation), evaluation at this time has the function of legitimation: "it is believed that scientific evidence justifies political decisions, either to improve, reduce or eliminate programs."

In view of the international panorama presented, it is necessary to make a geographical breakdown for the discussion on the evaluation studies field in Brazil. Unlike the countries with a long tradition of evaluating public policies and programs, Brazil entered at a late time in this phase of the political cycle. According to Mokate (2002, p. 90, our translation):

> Despite what has been said and according to what was published about the importance of evaluation processes, it is still uncommon to find social programs or policies in Latin America that have a rigorous systematic evaluation process incorporated into daily processes of administration and decision-making.

The concern about evaluating public policies or programs is something that has recently been incorporated into the political agenda, since administrators' priority was channeled into the policy formulation process, with little emphasis on the other stages of the cycle (COSTA and CASTANHAR, 2003; ALVES and PASSADOR, 2011; CRUMPTON, MEDEIROS, FERREIRA et al., 2016). This incorporation was reported in the mid-1980s, after the experience of structural reforms in the state apparatus proposed in various parts of the world, in order to reduce government costs and make it more efficient (PAULA, 2005). Therefore, a greater and less fragmented use of evaluation in Brazil only appears during the third wave described by Vedung (2010) or during the legitimization function proposed by Derlien (2001).

On the other hand, the adoption of evaluative practices in countries such as Brazil aims at strengthening institutional arrangements and providing administrators with an analysis of the policies adopted (CRUMPTON, MEDEIROS, FERREIRA et al., 2016). In their study, Crumpton, Medeiros, Ferreira et al. (2016) compared the evaluative studies carried out in Brazil and the USA, in order to present the degree of establishment of this type of research in the countries; as a result, it is found that, unlike the USA, the area of research in evaluation is not fully established in Brazil, although there is an effort on the part of the researchers to change this picture.

It is believed that, in the Brazilian case, the public databases available present reliability problems (some have declared data), unavailable data, and discontinuity in the indicators production; such factors may contribute to the low adherence to quantitative evaluative studies. Therefore, evaluations in the country would be between the second wave proposed by Vedung (2010), with more qualitative assessments of local character and specific policies analysis, and the third wave, focused mainly on accountability, performance measurement, and search for quality in the services provided.

At any rate, it should be emphasized that there is a strengthening of the evaluation study to seek best practices and theories, contributing in a substantial way to the creation of an evaluative “culture” capable of embracing the complexity and specificities of each policy or program in question (CARVALHO, 1999; UNDP, 2016).

The evolutionary trajectory of evaluative processes in Brazil and in the world facilitates understanding that there is no single method or standard defined at the present time. There is a multiplicity of approaches and methodologies covering a wide range of options. Based on the assumption that evaluations are not the same, Cohen and Franco (2008), and Ramos and Schabbach (2012) suggest the description of the different forms of evaluation according to a series of criteria, namely use, time, objectives, who performs it, to whom it is intended for, what is the expected outcome, among others. Costa and Castanhar (2003) call the variety of concepts and methodologies of evaluation studies a conceptual entanglement. The authors indicate that the knowledge of several theoretical alternatives contributes greatly to the choice of the most appropriate method for each type of policy or program to be evaluated.

According to Faria (2005), it is possible to indicate at least four types of use for the evaluation: instrumental; conceptual; persuasive; or clarification. The instrumental use refers to the importance of quality and the adequate propagation of the results achieved, in order to turn the evaluation somewhat tangible, and feasible of employability. Weiss (1998) adds that the instrumental use of evaluation might guide decision-makers as it enables a return to the history of the program analyzed, provides feedback to evaluation practitioners, and also highlights the objectives established for the program.

The second use, called *conceptual*, is often applied to those that do not have direct contact with the program formulation (local technicians); It is, therefore, necessary to provide information about the policy or program, to present its nature and operation, and to introduce them to the impacts and results achieved (FARIA, 2005). We may state that the main objective of the conceptual use is the learning provided to the actors; in this regard, Mokate (2002, p. 127, our translation) emphasizes that “summing up, evaluation allows us to enrich management processes with dynamic learning. Far from being a devilish monster, then, the assessment may have a new image, that of a guardian ally”. Complementarily, Ramos and Schabbach (2012) indicate that learning can be enhanced when technicians actively participate in the evaluation process.

Additionally, the third use, called *persuasive*, described by Faria (2005, p. 103), “occurs when it is used to mobilize the support towards the position that decision-makers already have about the necessary changes in the policy or program”. The evaluation, at that moment, has the function of legitimizing the desired changes and winning new supporters for this movement. Finally, the use of clarification is proposed as a result of the accumulation of knowledge originated from the several already performed evaluations (FARIA, 2005) and becomes capable of guiding the governmental agenda. Mokate (2002) states that the maturation of a new management paradigm, focused on evaluative practices, may be an ally in incorporating evaluation as a natural and gradual process for the achievement of the objectives proposed.

In order to provide a foundation for the differentiation between the various moments of the evaluation, Cohen and Franco (2008) divide it into *ex ante* evaluation and *ex post* evaluation.

*Ex-ante* evaluations are those performed prior to the decision-making process and may be said to serve as an input to decision-making on the implementation of social policies or programs (COTTA, 2001; COHEN and FRANCO, 2008). Furthermore, Ramos and Schabbach (2012) affirm that, in this type of evaluation, a diagnosis is generated on the situation observed, contributing with the adequacy of the available resources to the proposed objectives (viability verification), also being able to serve as a direction tool for the maintenance and/or formulation of policies or programs.

The *ex post* classification indicates that the evaluations are performed during the execution of a project or at the end of the project. When evaluating refers to the intermediate moment (process evaluation), the conclusions may found decisions of continuity or discontinuity of a certain policy or program (COTTA, 2001; COHEN and FRANCO, 2008; TREVISAN and BELLEN, 2008; RAMOS and SCHABBACH, 2012). With regard to terminal evaluations (which measure impact and outcome, for example), *ex-post* evaluations provide information concerning the experience gained and can support future decisions on the same policy or program or contribute to the construction of an evaluation model to another objective or similar experiences (COTTA, 2001; COHEN and FRANCO, 2008; TREVISAN and BELLEN, 2008; RAMOS and SCHABBACH, 2012). Based on the literature, we may infer that *ex post* evaluations are more widespread both in academic and practical applications; thus, they present better established and more sophisticated methodologies.

Similarly, Scriven (1991) defines two terms for the evaluation moments: *formative evaluation* and *summative evaluation*.

The formative evaluation resembles the function of the *ex post* intermediate evaluation, being conducted during the process of implementation or development of the program, and has as purpose providing an analytical reading to the decision makers.

The summative evaluation is directed after the completion of the program (terminal *ex post*), providing administrators and users with a judgment of the overall value of the program (SCRIVEN, 1991; TREVISAN and BELLEN, 2008). From the same point of view, Ramos and Schabbach argue that the evaluation ratings made by Scriven (1991) also present connections with the nature of evaluation, whether it is program formation (formative assessments) or analysis and production of information (summative assessment). Weiss (1998) argues that the simplicity proposed in Scriven's definition (1991) does not diminish the reach of the evaluative purposes, but produces information that contributes to the modification or maintenance of the program.

Regarding the classification of the evaluation from the point of view of who executes it or participates in it, it can be stated that there is a consensus in the literature regarding the distinction in four types: external evaluation; internal evaluation; mixed evaluation; and participatory evaluation. Firstly, Cohen and Franco (2008) warn that evaluations performed by people outside the organization tend to emphasize the method applied to the detriment of specific knowledge of the area in which the policy or program is developed. Cotta (1998) contributes with the discussion informing that external evaluations add an exemption to the evaluative practice, however, the external evaluator has greater difficulties of accessing the data. In addition, we can cite as an advantage the greater objectivity of the external evaluators, because they do not participate in the internal processes, besides the possibility of comparison with similar programs (COHEN and FRANCO, 2008; RAMOS and SCHABBACH, 2012). In the view of Thoenig (2000), in his article on public sector reform, external evaluations may play a crucial role in decision making; the author cites the case of Greece, which performed external evaluations for decision-making in relation to the structural funds of the European Union.

With regard to internal evaluations, the advantages presented by the literature are the reduction of the conflicts generated by the insertion of external people or organizations; the possibility of review and learning regarding the processes executed in a given policy or program; the greater theoretical and technical knowledge of the specific features evaluated. However, internal evaluations risk losing objectivity and partial conclusions, since they are developed by those involved in the formulation and implementation of the analyzed program (WEISS 1998; COTTA 1998; COHEN and FRANCO 2008; VEDUNG 2009; RAMOS and SCHABBACH, 2012).

Then, we have the mixed evaluation, which combines the two models described. Basically, the goal of mixed assessment is to add efforts and ensure synergy between the actors involved in the action. The potentialities of each model are then combined and the impacts or biases which are presented individually are reduced (COHEN and FRANCO, 2008; RAMOS and SCHABBACH, 2012).

Finally, there exists the participatory evaluation, which has as a main objective reducing the distance between the organizers and producers of the evaluation and the beneficiaries of the evaluated policies or programs. It allows the participation of users in all phases of the political cycle, favoring opinion, and meeting the specific demands of these actors (COHEN and FRANCO, 2008, RAMOS and SCHABBACH, 2012). For Patton (1999), participatory evaluation should consider priorities and effects and establish common definitions for both evaluators and beneficiaries.

The author describes some steps to make participatory evaluation feasible, such as conducting interactive sessions to discuss aspects of the evaluation; discussion and mode of use of priorities, purposes, and definitions; and proposal of an ideal evaluation model that satisfies all the conditions exposed in the previous stages (PATTON, 1999). Additionally, Spink (2001) exposes the importance of participatory evaluation, already highlighted, capable of strengthening the ties between the involved ones, helping to maintain the democratic evaluation practice. However, one should take into account the dimension of the policy or program evaluated; in this regard, Ramos and Schabbach (2012) suggest that participatory evaluations work adequately only on small projects or local programs. In turn, Carvalho (1999, p. 89) argues that:

> In fact, in dealing with more circumscribed units (an institution or a program), participatory evaluation becomes a rich procedure, since its accomplishment is shared with the agents and beneficiaries involved (in the program or institution), allowing a reflective and socialized appropriation among the various subjects of moving action in addition to the evaluation.

We should emphasize, therefore, that participatory evaluation promotes the incorporation of individuals into decision-making and favors social learning, and citizenship (CARVALHO, 1999). Ceneviva and Farah (2012) emphasize that bureaucrats should not work in isolation, on the contrary, evaluation is a tool for transparency of public acts and citizens can and should inform themselves and supervise the actions promoted by political actors.

Regarding the approach, the evaluations can be classified in qualitative, quantitative, and mixed methods. In short, Weiss (1998) classifies qualitative research as those that present themselves in words and quantitative ones, in numbers. The author complements by stating that quantitative assessments collect data and transform it into numerical information, strongly supported by statistics, mathematical, and econometric methods. Quantitative approaches can use sophisticated tools in data analysis, which identify relationships between one or more variables, and can encompass a large sample universe, as well as provide information on the extent and distribution of a phenomenon (WEISS, 1998). On the other hand, qualitative evaluations tend to use less structured tools, such as interviews, observational techniques and document analysis (WEISS, 1998). Garcia (2001, p. 29) argues that:

> The evaluation has to be operated with a broad vision, guided by a judgment of value, something eminently qualitative, focused on complex processes, in which the elements in interaction do not always produce measurable manifestations, and some of these elements may not present quantifiable attributes.

Rossi and Wright (1984) argue that, for social research, as in the case of public policies study, the qualitative approach displays advantages because it has the capacity to promote approximation with reality and is flexible to contemplate the complexity entailed in such policies. Mokate (2002) alludes to the great discussions in the social sectors that place qualitative and quantitative approaches on different sides of the same arena. It is an exaggerated competition about which method is most efficient, effectual or effective. These differences are attributed to the reasons for which, in part, evaluations are not routinely institutionalized as a routine of the administrative practice.

In terms of results, a large part of the national and international literature discusses the advantages and disadvantages of each approach, however, there is growing acceptance that the integration between them is the best solution to the theoretical confrontations (SILVEIRA, VIEIRA, CAPOBIANGO et al., 2013). The differences between the approaches generate opportunities for complementation, adding the similarities and increasing the synergy between the parties (MOKATE, 2002). It is in this area that the mixed methods approach is inserted, which are nothing more than the combination of qualitative and quantitative assessments. Mokate (2002, p. 114, our translation) argues that "particularly, the evaluation may also be applied to quantitative and qualitative methods to generate different types of information which together correspond to the different questions proposed by the evaluation process."

Scriven (1996) establishes the skills of an evaluator and notes the use of mixed methods, arguing that the evaluator should present basic qualitative domains such as questionnaire use and observational methods. The control of polarization between approaches is also necessary to combine such practical tests, measurement procedures, judgment and narrative evaluation, and case study techniques. Therefore, Bloch, Soresen, Graversen et al. (2014) argue that the use of mixed methods contributes to the use of the best available data sources, with the aim of providing comprehensive and robust answers. However, the authors present the time taken, the resources, and the commitment to build the pragmatic aspects of the evaluation as limitations to the use of mixed methods, (BLOCH, SORENSEN, GRAVERSEN et al., 2014). Costa, Pegado, Ávila et al. (2013) add that quantitative research leads to a generalization of results generated by standardized information. On the other hand, qualitative research will probably provide data on social circumstances and environments, giving emphasis to the cultural and contextual dimensions.

Considering the methodology applied to the evaluation, Costa and Castanhar (2003) group the evaluations in process evaluation; goals evaluation; and impact assessment. As for the process evaluation, according to Costa and Castanhar (2003, p. 980):

> Its objective is to detect possible defects in the elaboration of procedures, to identify barriers, and obstacles to its implementation and to generate important data for its reprogramming, through the registration of events and activities. Thus, the proper use of the information produced during the development of the program allows introducing changes in its content during its execution.

With purposes similar to those of the formative evaluation described by Scriven (1991), the process evaluation systematically assesses the evolution of policies or programs, as well as follows its internal processes. Its uses allow to contemplate the content of the program, the relationship between what was planned and what has been executed, the reach of the target users, and the success or failure in the correct delivery of the benefits (COSTA and CASTANHAR, 2003; RAMOS and SCHABBACH, 2012).

Furthermore, we have the goals evaluation, described by Costa and Castanhar (2003) as the most traditional evaluation methodologies. It aims to "measure the degree of success that a program achieves in relation to the achievement of previously established goals" (COSTA and CASTANHAR, 2003, p. 997). We may consider this type of evaluation as an *ex post* evaluation since it requires the program to be completed to be evaluated (COSTA and CASTANHAR, 2003). Ramos and Schabbach (2012) present a different conceptualization, called results evaluation, but with objectives similar to the goals evaluation; In this context, the effects and consequences of a given policy are measured, determining its success or failure based on effective changes in the populations benefited.

The evaluation, once considered as a way of measuring the performance of policies or programs, being such performance the achievement of goals (COSTA and CASTANHAR, 2003) or the results evaluation (RAMOS and SCHABBACH, 2012), needs criteria to structure and, finally, to measure the contracted consequence (COSTA and CASTANHAR, 2003). In this manner, three basic dimensions prevail to measure the results obtained (DRAIBE, 2001; COSTA and CASTANHAR, 2003; ALVES and PASSADOR, 2011):

- **Efficiency:** of economic origin, signifies to achieve the objectives of the program, by prioritizing established standards, with the lowest possible cost-benefit ratio;
- **Efficacy:** measures the degree to which goals and objectives have been achieved, thus translating in a simplified way the result achieved;
- **Effectivity:** also treated in the literature as a measure of impact, indicates the positive effects related to the program target public. It is a broader dimension, because it analyzes the economic, socio-cultural, environmental, and institutional aspects, that is, effectivity measures both the quantity and the quality of the goals reached by the program.

In the chain of ideas, Costa and Castanhar (2003), and Ramos and Schabbach (2012) agree on the last definition proposed: *the impact assessment*. It is therefore inferred that impact assessment covers not only the results achieved in terms of efficiency, effectivity, and efficacy, but also changes in the target population as a result of the implementation of the evaluated policy or program. Cotta (1998, p. 113) adds that:

> The difference between the results evaluation and impact assessment, therefore, depends eminently on the scope of the analysis: if the objective is to inquire about the effects of an intervention on the public served, then it is a results evaluation; if the intention is to capture the reflexes of this same intervention in a broader context, then it is an impact assessment. Or, put another way, the results evaluation is aimed at measuring the intermediate results of the intervention, and the impact assessment, its final results.

We may compare the impact assessment to the summative evaluation proposed by Scriven (1991), which is executed *ex post* to the realization of the program or project. In summary, this evaluative model uses methodological structures that provide cause and effect relationships between the actions performed by the program and the effects promoted in the target community (COSTA and CASTANHAR, 2003; RAMOS and SCHABBACH, 2012). In order to do so, it is necessary to follow two guidelines noted by Cotta (1998): a) the objectives must be defined to include the identification of measurable goals; and b) the implementation of the policy or program must be at least satisfactory to support an impact assessment without bias or even to make it unfeasible.

It is also worth mentioning the importance of the research designs and outcomes for each type of evaluation: of a process, goals/results or impacts, depending on the problem that it responds to. In view of this need, Box 2 describes, without the pretension of exhausting them, each of the designs and outcomes presented in the literature (SCRIVEN, 1991; WEISS, 1998; COTTA, 1998; OECD, 2002; COSTA and CASTANHAR, 2003; VEDUNG, 2009; CAMELO JUNIOR, FERNANDES, JORGE et al., 2011; RAMOS and SCHABBACH, 2012; SILVEIRA, VIEIRA, CAPOBIANGO et al., 2013)

**Box 2**

**Description of de research designs and outcomes**

| Design/outcome | Brief description | Evaluation type |
|---|---|---|
| Cost-benefit | Compares the benefits and costs of a program/project and searches the best benefit relation. | Process evaluation;<br>Goal evaluation/results. |
| Cost-effectivity | Comparison between alternative ways of acting in order to select the program/project that meets the objectives with the lowest cost. | Process evaluation;<br>Goal evaluation/results. |
| Cost-usefulness | It compares not only the costs and benefits, but the perception of quality of life or social impact of the target population. | Impact evaluation |
| Experimental studies | They involve the conformation of two populations: the group with the program/project (treatment) and another group without the program/project (control). The goal is to check the changes generated by the program/project. | Impact evaluation |
| Quasi-experimental studies | Similar to the experimental model, however, the groups are not randomly selected, there exists a choice among those belonging to the target population of those who fit into one of the groups (treatment or control). The objective is also to verify the changes that occurred in these populations. | Impact evaluation |
| Documental methods | Method of data harvest involving the use of information for other purposes, such as statistics or data produced by the organ itself and/or other written documents. | Process evaluation;<br>Goal evaluation/results. |
| Interrogative methods | Strategy of accumulating data through the questioning of people, appropriate with interviews, questionnaires, and alike. | Process evaluation;<br>Goal evaluation/results.<br>Impact evaluation. |
| Observational methods | Data collection strategy in which the researcher, through a site visit, makes observations with his or her own eyes. In direct observation, the participation in the process to be evaluated is a pretext to allow the collection of unbiased data. | Process evaluation;<br>Goal evaluation/results.<br>Impact evaluation. |
| Meta-evaluation | The term is used for designed evaluations that aggregate the results of a series of evaluations. It can also be used to denote the evaluation of an evaluation, to evaluate its quality and/or to evaluate the performance of the evaluators. | Goal evaluation/results.<br>Impact evaluation. |
| Triangulation | It refers to the use of various measures or different measuring instruments to obtain information about the same action, condition or thing. | Process evaluation;<br>Goal evaluation/results.<br>Impact evaluation. |

Source: Elaborated by the authors.

In conclusion, in view of the variety and complexity of the applications, methods, tools, approaches, moments and actors involved in the evaluation process, the next section explores international trends about the evolution of evaluation studies, with the aim of comparing them under the point of view of types, methodologies, and outcomes.

## RECENT NATIONAL AND INTERNATIONAL TENDENCIES

In the face of ever more complex and turbulent scenarios, global trends are converging towards structuring and strengthening international evaluation actions, supported by government agencies, independent agencies, multilateral organizations, and civil society organizations. It is clear that the parties agree that evaluative activities are important for good governance and can contribute to the advancement of more effective, efficacious, and efficient social public policies (RAMOS and SCHABBACH, 2012; EVALPARTNERS, 2015). In order to achieve it, this topic maps, without the pretension of exhausting them, the actions developed and presents actions executed by international organizations.

International organizations such as the World Bank, UN and OECD participate and play a key role in building and strengthening institutionalized evaluative practices. The aim is to develop methodological frameworks and tools to improve the design, execution, and evaluation of programs or projects around the world (RAMOS and SCHABBACH, 2012).

In this context, 2015 was established by EvalPartners and approved by the UN as the international year of evaluation. The organization, built through partnerships between the International Organization for Evaluation Cooperation (IOCE) and the United Nations Children's Fund (UNICEF), in partnership with several major organizations, aims to create a network of global evaluators who can contribute to the evaluation of public policies in their countries (EVALPARTNERS, 2016). According to the guidelines of the organization, the initiative aims at contributing to the improvement of the capacities of civil society organizations (CSOs), as well as at involving them strategically and expressively in the national evaluation processes, with the aim of developing systems of evaluation and ensure equity in actions (EVALPARTNERS, 2016). To this end, they developed the Global Assessment Agenda (or EvalAgenda2020). The document, approved by the delegates of the Global Assessment Forum, held in Nepal in 2015, provides a guide for future assessment, and for the strengthening of evaluators from different countries and organizations (EVALPARTNERS, 2016).

The European Union has also developed a project to evaluate the structural policies of the bloc, as well as the identification and use of innovative practices for evaluation methodologies, as well as multisectoral involvement with the objective of reducing poverty (SOPHIE PROJECT, 2015).

In the context of Latin America, the Latin American and Caribbean Network of Monitoring, Evaluation and Systematization (ReLAC), in cooperation with the Foment to the Evaluation Capacities Project (Foceval Project) of the Ministry of Planning and Economic Policy of Costa Rica (Mideplan), and the German Institute for Development Cooperation Assessment (DEVAL) conducted a series of public consultations with experts and evaluation professionals during the years 2014 and 2015 and presented the first drafts at the IV ReLAC Conference in 2015. Recently, in August 2016, they released a document summarizing this proposal entitled Guidelines for Evaluation for Latin America and the Caribbean, with objective of enriching the common guideline for the theoretical and practical formation of evaluators in the region, and instigating the culture of evaluation in the continent (BILELLA, VALENCIA, ALVAREZ et al., 2016).

We may confirm the performance of the Evaluation and Results Learning Center for Brazil and Lusophone Africa (FGV EESP-Clear), based at the Getulio Vargas Foundation (FGV) in São Paulo, which benefits from the multidisciplinary environment of the FGV, which is one of the six regional centers that make up the Clear initiative - which involves several countries and aims to improve policies and programs with development of capacities, and monitoring and evaluation systems. Among the axes of monitoring and evaluation research, the following project: a) training; b) technical assistance; c) generation of new evidence; and d) diffusion of evidence and knowledge (FGV CLEAR, 2018).

Still in Brazil, another initiative is gaining notoriety among the scientific community: The Center for Data Integration and Knowledge for Health (Cidacs), created in 2016 and linked to the Oswaldo Cruz Foundation of Bahia (Fiocruz Bahia), which promotes research and interdisciplinary projects with processing, integration, and analysis of large data. The aim is to contribute to the use of innovative methodologies, to provide professional and scientific training, and to support decision-makers in relation to public social policies, especially health determinants, and environmental conditions (CIDACS, 2018).

It is also worth noting that there is a global effort aimed at building an agenda based on integration among all international organizations, guided mainly by the achievement of the 17 Sustainable Development Goals (SDGs) of the United Nations Development Program (UNDP, 2016). The first five years envisaged by EvalAgenda2020 have as priority the creation and strengthening of mechanisms for the evaluation of SDGs. Box 3 summarizes the main actions (UNDP, 2016; UNO, 2016; EVALPARTNERS, 2015; OECD, 2016; SOPHIE PROJECT, 2012; BILELLA, VALENCIA, ALVAREZ et al., 2016; CLEAR, 2018; CIDACS, 2018; IBGE, 2018).

**Box 3**

**Actions of structuration of indicators and evaluations developed by national and international organs**

| Organization | Action developted | Brief description |
|---|---|---|
| UN (2015) | Transforming our world: the 2030 Agenda for sustainable development. | Document establishing the 17 Objectives and 169 goals for Sustainable Development approved by the UN General Assembly in 2015. |
| UN (2016) | First follow-up report on the SDGs. | Inaugural report giving an overview of the 17 SDGs, using available data to emphasize some critical points and challenges. |
| UN (UNDP, 2016) | 2030 agenda follow-up | Document with initial subsidies for the identification of the national indicators related to the SDGs, from which the objective is to present the indices available for monitoring and evaluating the achievement of established goals. |
| SDGs at the Brazilian Institute of Geography and Statistics (IBGE) | SDGs Digital Platform. | Developed by teams of the Informatics Board of IBGE, it provides an information system to follow the 2030 Agenda in Brazil. It is a first set of global indicators, developed in a collaborative way among global institutions, in order to present national data produced on a regular basis, with internationally established methodology and standards. |
| EvalPartners (2015) | EvalAgenda2020. | It establishes priorities and offers training and forums to create an environment conducive to evaluations and make the achievement of the goals set out in the 2030 Agenda feasible. |
| OECD (2016) | Measuring distance to the SDG targets: a pilot assessment of where OECD countries stand. | Pilot study of starting positions for the SDGs in some OECD countries. It aims to help governments develop national action plans. It uses the abundance of statistical data obtained by OECD members, harmonized through the organization's tools and processes. |
| SOPHIE PROJECT (2012) | Evaluating the impact of structural policies on health inequalities and their social determinants, and fostering change (SOPHIE). | Inequalities in health are strongly affected by the circumstances in which people are born, grow, live, work and grow old, as well as by the policies that influence those circumstances. The purpose of the SOPHIE Project is to generate new evidence on the impact of structural policies on health inequalities and to develop innovative methodologies for the evaluation of these policies in Europe. It is coordinated by the Public Health Agency of Barcelona and funded by the European Union Research Fund. |
| ReLAC | Evaluation Guidelines for Latin America and the Caribbean. | An instrument to foster the culture of evaluation in public and private institutions, guiding the practice and training of evaluators, guaranteeing the credibility, transparency and integral character of the evaluation (BILELLA, VALENCIA, ALVAREZ et al., 2016). The guidelines were divided into 5 dimensions of analysis: a) rigorous evaluation; b) adequate evaluation; c) evaluation carried out in accordance with ethical and legal principles; d) adequate cultural understanding; and e) relevance and utility. |
| FGV Clear | Development of various projects in conjunction with government agents, private institutions, civil society and academia. | The projects are designed to formulate and/or improve the administration of public policies and social programs based on results in four areas: a) diagnosis; b) monitoring: monitoring systems, construction of indicator; c) evaluation: logical framework, process evaluation, cost-benefit analysis, impact assessment; and d) institutional capacity building. |
| Cidacs –Bahia Oswaldo Cruz Foundation (Fiocruz Bahia) | Development of studies with large volumes of information, integrating national databases and providing products (databases) for research. | Cidacs operates with the following research platforms: Cohort of 100 Million Brazilians; Zika Platform; Technologies and Innovations for the Unified Health System; Urban Equity and Sustainability; Bioinformatics, and Genetic Epidemiology (Epigen). For this study, the Cohort platform of 100 Million Brazilians, made up of 114 million individuals belonging to low-income families (with a *per capita* income of less than 3x the minimum wage) eligible for social programs of the Federal Government, is highlighted. Socioeconomic, demographic, housing, sanitation, remuneration and labor information, expenditures, schooling, family composition, birth, and mortality in the Unified Register (CadÚnico) databases, health information systems, system on conditionalities of the *Bolsa Família* Program, and the information system of the housing program, *Minha Casa Minha Vida* (CIDACS, 2018). |

Source: Elaborated by the authors.

It is clear to understand that evaluative activities are important for good governance and can contribute to the advancement of more effective, efficacious, and effective public policies (RAMOS and SCHABBACH, 2012; EVALPARTNERS, 2015). We should emphasize that, as can be deduced from the academic and governmental research organizations, it is important to strengthen the study of evaluation in order to seek best practices and theories, greatly contributing to the creation of an evaluative "culture" capable of encompassing the complexity and the specificities of each policy or program in question (CARVALHO, 1999; UNDP, 2016).

Based on the theoretical framework built and the assessments made by multilateral organizations, the main theoretical gaps observed are: a) studies involving multiple countries; b) use of the mixed methods approach; c) evaluative studies combining two or more public policies; d) strengthening, training, and structuring of evaluation practices in developed and developing countries; e) use of the results of evaluations to inform and improve public policies; and f) participatory evaluations - contemplation of all the actors involved (administrators, decision makers, formulators, implementers, executors, and beneficiaries).

In view of the efforts presented, the importance of robust, unbiased, and multi-method approaches is evident. According to the above, it is observed that developed countries have already incorporated in their political agenda the evaluation as a process, different from what occurs in Brazil and in other developing countries, where the evaluation of public policies and programs still suffers from discontinuity and unavailability of reliable and complete data, as well as integrated information systems that facilitate access to information, which would contribute greatly to decision making.

## FINAL CONSIDERATIONS

In consonance with the search for articulated models for the development of governmental strategies to combat poverty and promote citizenship in Brazil, several international organizations have been developing integrative and intersectoral proposals aimed at improving the quality of life of the world's population. Actions such as the establishment of UNDP SDGs and comparative and shared studies among countries belonging to the OECD are examples of such efforts.

In Latin America, there is a growing search for evaluative actions to strengthen and improve the public policies adopted in the region, mainly related to addressing social vulnerabilities. Nevertheless, the scenario demands the development of the professional capacities of the governments and increase the performance of civil society in monitoring and control. The evolution and internalization of evaluation processes contribute to the development of robust and comprehensive methodologies. From the 1980 decade, there is a growing search for models capable of combining quantitative and qualitative information (mixed methods), aiming at improving and making the evaluation more realistic regarding the complexity of the analyzed policies. The search for new robust and comprehensive evaluation methodologies is essential for the development and evolution of evaluation activities, providing administrators with concrete, real, and applicable data. Therefore, the aim is to use innovative methodologies, providing professional and scientific training, and supporting decision-makers in relation to public social policies, especially health-related ones.

However, it is worth mentioning that, in the Brazilian case, the available public databases present problems of reliability, unavailable data and discontinuity in the production of indicators; such factors may contribute to the low adherence to quantitative evaluative studies. Thus, evaluations in Brazil would have more qualitative evaluations of local character and analysis of specific policies, mainly focused on accountability and performance measurement and in search of the quality of services provided. There are also experiments that promote interdisciplinary research and projects based on the processing, integration, and analysis of large data quantities (*big data*).

Fortunately, we should note that the data demonstrate the strengthening of the study of evaluation in Brazil, in order to seek best practices and theories, enormously contributing to the creation of an evaluation "culture" capable of embracing the complexity and specificities of each policy or program in question. However, it is also necessary to emphasize the importance of continuous evaluation actions, especially in scenarios with scarce financial, human, and material resources. Such evaluations may contribute to the adequate allocation of resources and the strengthening of control actions by the responsible agencies.

## REFERENCES

ALA-HARJA, M.; HELGASON, S. Em direção às melhores práticas de avaliação. **Revista do Serviço Público**, v. 51, n. 4, p. 5-60, 2000.

ALVES, T.; PASSADOR, C. S. **Educação pública no Brasil**: condições de oferta, nível socioeconômico dos alunos e avaliação. São Paulo: Annablume, 2011.

ARRETCHE, M. T. S. Tendências no estudo sobre avaliação. In: RICO, E. M. (Org.). **Avaliação de políticas sociais**: uma questão de debate. São Paulo: Cortez, 2001.

BILELLA, P. D. R. et al. **Diretrizes para avaliação para a América Latina e o Caribe**. Buenos Aires: Akian, 2016.

BLOCH, C. et al. Developing a methodology to assess the impact of research grant funding: A mixed methods approach. **Evaluation and Program Planning**, v. 43, p. 105-117, 2014.

BOBBIO, N.; MATTEUCCI, N.; PASQUINO, G. **Dicionário de política**. 1. ed. Brasília: Editora da UnB, 1998.

CAMELO JUNIOR, J. S. et al. Avaliação econômica em saúde: triagem neonatal da galactosemia. **Cadernos de Saúde Pública**, v. 27, n. 4, p. 666-676, 2011.

CARVALHO, M. C. B. Avaliação participativa: uma escolha metodológica. In: RICO, E. M. (Org.). **Avaliação de políticas sociais**: uma questão de debate. 2. ed. São Paulo: Cortez, 1999. p. 87-94.

CENEVIVA, R.; FARAH, M. F. S. Avaliação, informação e responsabilização no setor público. **Revista de Administração Pública**, Rio de Janeiro, v. 46, n. 4, p. 993-1016, 2012.

CENTRO DE INTEGRAÇÃO DE DADOS E CONHECIMENTOS PARA SAÚDE – CIDACS. **Quem somos**. 2018. Available at: <https://cidacs.bahia.fiocruz.br/sobre/quem-somos/>. Accessed on: July 18, 2018.

COHEN, E.; FRANCO, R. **Avaliação de projetos sociais**. 8. ed. Rio de Janeiro: Vozes, 2008.

COSTA, A. F. et al. Mixed-methods evaluation in complex programmes: the national reading plan in Portugal. **Evaluation and Program Planning**, v. 39, p. 1-9, 2013.

COSTA, F. L.; CASTANHAR, J. C. Avaliação de programas públicos: desafios conceituais e metodológicos. **Revista de Administração Pública**, Rio de Janeiro, v. 37, n. 5, p. 969-992, 2003.

COTTA, T. C. Metodologias de avaliação de programas e projetos sociais: análise de resultados e de impacto. **Revista do Serviço Público**, v. 49, n. 2, p. 103-124, 1998.

COTTA, T. C. Avaliação educacional e políticas públicas: a experiência do Sistema Nacional de Avaliação da Educação Básica (Saeb). **Revista do Serviço Público**, v. 52, n. 4, p. 89-111, 2001.

CRUMPTON, C. D. et al. Avaliação de políticas públicas no Brasil e nos Estados Unidos: análise da pesquisa nos últimos 10 anos. **Revista de Administração Pública**, Rio de Janeiro, v. 50, n. 6, p. 981-1001, 2016.

CRUZ, M. M. Avaliação de políticas e programas de saúde: contribuições para o debate. In: MATTOS, R. A.; BAPTISTA, T. W. F. (Org.). **Caminhos para análise das políticas de saúde**. Porto Alegre: Rede Unida, 2015. p. 285-317.

DERLIEN, H.U. Una comparación internacional en la evaluación de las políticas públicas. **Revista do Serviço Público**, v. 52, n. 1, p. 105-124, 2001.

DRAIBE, S. M. Avaliação de implementação: esboço de uma metodologia de trabalho em políticas públicas. In: BARREIRA, M. C. R. N.; CARVALHO, M. C. B. (Org.). **Tendências e perspectivas na avaliação de políticas e programas sociais**. São Paulo: IEE/PUC-SP, 2001. p. 13-42.

EVALPARTNERS. **Global Evaluation Agenda 2016-2020**. 2015. Available at: <http://www.evalpartners.org/sites/default/files/documents/EvalAgenda2020.pdf>. Accessed on: July 15, 2016.

EVALPARTNERS. **Global evaluation agenda**. [2016]. Available at: <http://www.evalpartners.org/global-evaluation-agenda>. Accessed on: July 27, 2016.

FARIA, C. A. P. A política da avaliação de políticas públicas. **Revista Brasileira de Ciências Sociais**, v. 20, n. 59, p. 97-109, 2005.

FGV CLEAR. **Centro de Aprendizagem em Avaliação e Resultados para o Brasil e a África Lusófona (FGV EAESP Clear)**. 2018. Available at: <http://fgvclear.org/pt/sobre-o-fgv-clear/>. Accessed on: July 17, 2018.

FREY, K. Políticas públicas: um debate conceitual e reflexões referentes à prática da análise de políticas públicas no Brasil. **Planejamento e Políticas Públicas**, n. 21, [n.p.], 2000.

GARCIA, R. C. Subsídios para organizar avaliações da ação governamental. **Planejamento e Políticas Públicas**, n. 23, p. 7-70, 2001.

HOWLETT, M.; RAMESH, M.; PERL, A. **Política pública – seus ciclos e subsistemas**: uma abordagem integradora. Rio de Janeiro: Elsevier, 2013.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. **Indicadores ODS**. 2018a. Available at: <https://ods.ibge.gov.br/>. Accessed on: Apr. 25, 2018.

MOKATE, K. M. Convirtiendo el "monstruo" en aliado: la evaluación como herramienta de la gerencia social. **Revista do Serviço Público**, v. 53, n. 1, p. 89-134, 2002.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT – OECD. **Glossary of key terms in evaluation and results-based management**. Paris: OECD, 2002.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS – ONU. **Transformando nosso mundo: a agenda 2030 para o desenvolvimento sustentável**. 2015. Available at: <https://nacoesunidas.org/pos2015/agenda2030/>. Accessed on: July 27, 2016.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS – ONU. **Global Evaluation Agenda**: Eval2016-2020. [s.l]: EvalPartners, 2016.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT – OECD. **Measuring distance to the SDGs targets**: a pilot assessment of where OECD countries stand. [2016]. Available at: <http://www.oecd.org/std/oecd-measuring-distance-to-the%20sdgs-target-pilot-study-web.pdf>. Accessed on: July 27, 2016.

PATTON, M. Q. **Utilization-focused evaluation in Africa**. 1999. Available at: <http://gametlibrary.worldbank.org/FILES/284_Utilisation%20Focused%20evaluation%20in%20Africa.pdf>. Accessed on: July 25, 2016.

PAULA, A. P. P. **Por uma nova gestão pública**: limites e potencialidades da experiência contemporânea. Rio de Janeiro: FGV, 2005.

PICCIOTTO, R. Democratic evaluation for the 21st century. **Evaluation**, v. 21, n. 2, p. 150-166, 2015.

PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO – PNUD. **Objetivos de desenvolvimento sustentável**. [2016]. Available at: <http://www.pnud.org.br/ods.aspx>. Accessed on: July 27, 2016.

RAMOS, M. P.; SCHABBACH, L. M. O estado da arte da avaliação de políticas públicas: conceituação e exemplos de avaliação no Brasil. **Revista de Administração Pública**, Rio de Janeiro, v. 46, n. 5, p. 1271-1294, 2012.

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

SCRIVEN, M. **Evaluation thesaurus**. 4. ed. Newbury Park: Sage, 1991.

SCRIVEN, M. Types of evaluation and types of evaluator. **Evaluation Practice**, v. 17, n. 2, p. 151-162, 1996.

SECCHI, L. **Políticas públicas**: conceitos, esquemas de análise, casos práticos. 2. ed. São Paulo: Cengage Learning, 2014.

SILVEIRA, S. F. R. et al. Políticas públicas: monitorar e avaliar para quê? In: FERREIRA, M. A. M.; ABRANTES, L. A. (Org.). **Políticas públicas, gestão e sociedade**. Viçosa: Triunfal, 2013. p. 301-327.

SOPHIE PROJECT. **Conclusions**. [2015]. Available at: <http://www.sophie-project.eu/pdf/conclusions.pdf>. Accessed on: Aug. 08, 2016.

SOPHIE PROJECT. **Evaluating the impact of structural policies on health inequalities and their social determinants, and fostering change (SOPHIE)**. [2012]. Available at: <http://www.sophie-project.eu/project.htm>. Accessed on: Sept. 13, 2016.

SOUZA, C. Políticas públicas: uma revisão da literatura. **Sociologias**, v. 8, n. 16, p. 20-45, 2006.

SPINK, P. Avaliação democrática: propostas e prática. **Fundamentos de Avaliação**, n. 3, p. 7-25, 2001.

THOENIG, J. C. A avaliação como conhecimento utilizável para reformas de gestão pública. **Revista do Serviço Público**, v. 51, n. 2, p. 54-71, 2000.

TREVISAN, A. P.; BELLEN, H. M. V. Avaliação de políticas públicas: uma revisão teórica de um campo em construção. **Revista de Administração Pública**, Rio de Janeiro, v. 42, n. 3, p. 529-550, 2008.

VEDUNG, E. **Public policy and program evaluation**. New Brunswick: Transaction, 2009.

VEDUNG, E. Four waves of evaluation diffusion. **Evaluation**, v. 16, n. 3, p. 263-277, 2010.

WEISS, C. H. Have we learned anything new about the use of evaluation? **American Journal of Evaluation**, v. 19, n. 1, p. 21-33, 1998.

WEISS, C. H. The interface between evaluation and public policy. **Evaluation**, v. 5, n. 4, p. 468-486, 1999.

Lilian Ribeiro de Oliveira

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

Doctor in Administration of Organizations at the Faculty of Economics Administration and Accounting of Ribeirão Preto at the University of São Paulo (USP); Member of the Center for Studies in Management and Contemporary Public Policies of the University of São Paulo (GPublic); Professor of Business Administration at the Educational Foundation of Ituverava (FEI/FFCL) and at Universitary Center Barão de Mauá (CBM), Ribeirão Preto – SP, Brazil. E-mail: lilianroliveira@fearp.usp.br

Claudia Souza Passador

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

Doctor in Education by the São Paulo University; Docent in Public Administration from the Faculty of Economics, Administration and Accounting of Ribeirão Preto, São Paulo University (USP) (2012); Associate Professor, Department of Administration, Faculty of Economics, Administration and Accounting, Ribeirão Preto, São Paulo University (USP); Coordinator of the Center for Studies in Management and Contemporary Public Policies at the São Paulo University (GPUBLIC/USP); Researcher at the Center for Research Support Center for Integrated Systemic Biology (NAP/CISBi - FMRP/USP) and at the Institute for Advanced Studies at the São Paulo University (IEA/USP), Ribeirão Preto – SP, Brazil. E-mail: cspassador@usp.br