



**THE GEORGE
WASHINGTON
UNIVERSITY**

WASHINGTON, DC

The Institute of Brazilian Issues (IBI) / Center for Latin America Issues (CLAI)

School of Business

MINERVA PROGRAM

IMPROVING THE EVALUATION FUNCTION

SUGGESTIONS FOR THE RIO DE JANEIRO STATE

Elizabeth da Costa Mendes Oliveira de Menezes

Rio de Janeiro State Secretariat of Planning and Management

(SEPLAG-RJ)

Advisor: Kathryn Newcomer

Washington D.C. – Spring 2015

Contents

Summary.....	3
1 Introduction	4
2 Evaluation in the Public Policy Cycle	5
3 Program Evaluation	8
3.1 Important Features	11
3.2 Why Evaluate	13
3.3 How to Select a Program to Evaluate.....	14
3.4 Evaluation Categories.....	15
3.5 Evaluation Methods.....	19
4 The Importance of the Logic Model.....	20
5 Program Evaluation in United States	21
6 The Pluriannual Plan of the Rio de Janeiro State	23
6.1 Legal Aspects	24
6.2 Program Analysis by the Rio de Janeiro State	25
7 Suggestions to Improve the Evaluation Function.....	26
7.1 To the Process of Implementation.....	27
7.2 To the Design	29
8 Conclusion.....	32
Bibliography	34

Summary

The focus of this paper is to improve the evaluation function in the Rio de Janeiro State with the main objective to promote its sustainable economic development. It starts with a context relating the economic growth and the capacity of planning in Brazil. It then highlights that nowadays, the climate of fiscal austerity requires an efficient, effective, and transparent action of government and describes the importance of evaluating throughout the public policy cycle, because all of the stages affect each other, up to the point of preventing the evaluation. At this point, the paper discusses the experience of the State with some unresolved issues. After that, it provides a description of categories and methods and the importance of the logic model, followed by some experiences in United States with the subject matter. At the end, suggestions are made to improve the evaluation function in the State, to disseminate a culture of evaluation, and in therefore, to pursue the main objective, which is to promote the sustainable economic development of Rio de Janeiro State.

Keywords: Evaluation function, public policy cycle, sustainable economic development.

1 Introduction

In the 2000s, Brazil experienced a period of low inflation rates and economic growth, when compared with the 1980s, which is considered “the lost decade,” due to the severe economic, political, and social crisis. During the 1980s, planning public policies was difficult because all efforts were concentrated to combat inflation and to promote economic stability. Just in 1994, with the Real Plan, the country gave its first paces for economic stability and, therefore, for planning capacity. Despite the fact that nowadays the economic scene allows planning, the government has many problems in all phases of the public policy cycle. The way that a policy enters the agenda, its drafting, implementation, monitoring, and evaluation have complex methods and political influences that may induce mistakes.

Furthermore, after the world economic crisis started in 2008, countries increased their spending and by consequence their debts, which in turn increased the demand for systematic data on performance of public policies around the world. According to Stack (2014): “In climate of fiscal austerity, it is far better to cut programs with minimal impact and improve existing programs, based on evidence from high-quality program evaluations.”¹

Normally, governments search for economic growth, but growth alone does not guarantee equity. The population welfare included in the concept of sustainable economic development must be the ultimate goal, and with the scarcity of resources and unlimited needs, government expenditures increasingly need to be grounded.

¹ STACK, K. Advisor for Evidence-Based Innovation at the U.S. Office of Management and Budget. Available on: businessofgovernment.org. Accessed in: 02/09/2015.

In this context, this paper focus on improving the evaluation function, drawing attention to its importance throughout the public police cycle, describing its categories and methods, and discussing some United States practices. Following that, the paper analyzes the legal instrument Pluriannual Plan, PPA, from the Rio de Janeiro State and, at the end, makes some suggestions in order to improve the information quality provided to society, to Audit Courts, and to government with an ultimate goal: to improve the sustainable economic development in the State.

2 Evaluation in the Public Policy Cycle

According to Souza (2006), there is an increasing importance of three factors affecting public policies:

- Restrictive policy expenditures that dominates the agenda of the countries in particular developing countries;
- New insights into the role of governments replaced by the Keynesian policies of the post-war by restrictive policies expenditure; and
- Most of the developing countries, especially Latin American ones, have not yet managed to form political coalitions capable of solving minimally the question of how to design public policies to boost economic development and promote the social inclusion of much of the countries` population.²

Lynn (1980) defines public policy as a set of government actions that will produce specific effects, and Dye (1984) summarizes the definition as "what the government

² SOUZA, C. *Políticas Públicas: uma revisão da literatura*. Sociologias, Porto Alegre, ano 8, jul/dez 2006, p.20/45.

chooses to do or not to do." ^{3,4} That said, there is no consensus about the definitions of public policy, but the best known explanation remains Lasswell's (1936). He claims that decisions and analyses of public policy involve answering the following questions: who gets what, why, and what difference it makes.⁵ These questions are comprehensive and view the cycle of public policy as a model.

According to Souza (2006):

"The cycle of public policy sees public policy as a deliberative cycle, consisting of various stages and constituting a dynamic learning process. The cycle of public policy consists of the following stages: agenda-setting, identification of alternatives, evaluation of options, selection of options, implementation, and evaluation."⁶ (Author's translation).

The Rio de Janeiro State Pluriannual Plan Manual describes the policy cycle as involving five phases: planning, implementation, monitoring, evaluation and review.⁷ It is noteworthy that the process begins with a demand or an opportunity proposal from the stakeholders who pressure for it to enter in the public agenda setting. After that, a legal instrument called PPA formalizes the first phase: planning. This instrument contains the public choices for all government, including the policies of all agencies of the three powers (executive, legislative and judicial). Planning is essential to the policy cycle

³ LYNN, L. E. *Designing Public Policy: A Casebook on the Role of Policy Analysis*. Santa Monica, Calif.: Goodyear. 1980.

⁴ DYE, T. D. *Understanding Public Policy*. Englewood Cliffs, N.J.: PrenticeHall. 1984.

⁵ LASSWELL, H. D. *Politics: Who Gets What, When and How*. The American Political Science Review. American Political Science Association, 1936.

⁶ SOUZA, C. *Políticas Públicas: uma revisão da literatura*. Sociologias, Porto Alegre, ano 8, jul/dez 2006, 20/45 p.

⁷ Manual de Elaboração do Plano Plurianual 2012 – 2015. RJ: SEPLAG/SUBPL, 2011. Available on: <http://download.rj.gov.br/documentos/10112/780695/DLFE-46913.pdf/ManualdeElaboracaoPPA20122015.pdf>. Accessed in: 02/12/2015.

because any failure at this point, even with adjustments during the process, may compromise the later phases and the results.

Implementation is the second phase during which adopted policies are put into effect. In Brazil it is not imposing, just authorizing. For a long time, with every contingency of budget, this issue is discussed. The 1988 Brazilian Federal Constitution gives to the executive power the possibility to give up certain policies, but most of the budget is mandatory. Discretion only focuses on 4.7% of the budget, according to “Orçamento Cidadão 2015.”⁸ Besides the revenue frustration, the inefficiency of the Public Administration is also a factor that contributes to a low level of implementation.

According to Gertler et al (2007), the third phase, monitoring, is a continuous process to inform program implementation and day-to-day management and decisions. During this phase, the focus must be on inputs, activities, outputs, and occasionally outcomes.⁹ The Rio de Janeiro State monitors the instrument PPA at three quadrennials reports that consider only the financial and physical relation of goods and services delivered to society and to the State`s government. This relationship does not express totally the results of the programs` implementation, but is the only analysis criterion possible to be realized by the Planning Undersecretary. The Pluriannual Plan is not prepared to receive information about indicators to allow other kinds of evaluation.

Still according to Gertler et al (2007) the phase, evaluation, is:

⁸ Orçamento Cidadão 2015. Available on: http://www.planejamento.gov.br/secretarias/upload/Arquivos/sof/PLOA2015/Orcamento_Cidadao_%202015.pdf. Accessed in: 02/09/2015.

⁹ GERTLER, P. J.; MARTINEZ S.; PREMAND P.; RAWLINGS, L. B.; VERMEERSCH C. M. J. Impact Evaluation in Practice. Interactive textbook, 2007. Available on: <http://www.worldbank.org/pdt>.

“[...] periodic, objective assessments of a planned, ongoing, or completed project, program, or policy. Evaluations are used to answer specific questions related to design, implementation, and results.”¹⁰

As noted, this phase cannot be implemented yet in the Rio de Janeiro State with criteria because it depends on the other phases, with planning according to a policy-planning model.

The 1988 Brazilian Federal Constitution explains the last phase “review” because it provides the Plano Pluriannual, PPA, for 4 years. As planning is dynamic, it is possible to review the programs every year. This phase is similar as the first one, but the changes are minor.

The evaluation function is more advanced at the Federal level than at the State level, but the Rio de Janeiro State, specifically the Planning and Management Secretariat, is expending efforts in this direction by creating a network of discussion and training, and developing a system capable of accumulating data. Some state agencies like health, education, environmental and security have their own methods to evaluate relevant policies, but the Planning and Management Secretariat is not involved in this process, and thus is unable to integrate the efforts and to release the budget according to a criterion based on results.

3 Program Evaluation

To evaluate means to assign value, but the word has been used with multiples meanings. Ala-Harja and Helgason (2000) explain that there is no consensus about the meaning because the term “public policy” pervades many disciplines, institutions, and

¹⁰ GERTLER, P. J.; MARTINEZ S.; PREMAND P.; RAWLINGS, L. B.; VERMEERSCH C. M. J. Impact Evaluation in Practice. Interactive textbook, 2007. Available on: <http://www.worldbank.org/pdt>.

performers, covering a number of issues, needs, and people. They define evaluation in simple terms: “the term includes assessing the result of a program in relation to the proposed objectives.”¹¹ According to the American Evaluation Association, AEA, evaluation means “assessing the strengths and weaknesses of programs, policies, personnel, products, and organizations to improve their effectiveness.”¹² The Organization for Economic Co-operation and Development, OECD, defines evaluation as “the systematic and objective assessment of an on-going or completed project, program, or policy, its design, implementation, and results in relation to specified evaluation criteria”.¹³ The Unit Budgetary Overview and Evaluation in the Directorate-General for Budgets of the European Commission considers evaluation “an essential part of modern public sector of management practice,” but considers it impossible to arrive a single definition. However, they identify some features, evaluation:

- Should be analytical - based on recognized research techniques;
- Should be systematic – they require careful planning and consistent use of the chosen techniques;
- Should be reliable - the findings of an evaluation should be reproducible by a different evaluator with access to the same data and using the same methods of data analysis;
- Should be issue-oriented - evaluations should seek to address important issues relating to the program, including its relevance, efficiency and effectiveness;

¹¹ ALA-HARJA, M.; HELGASON, S. Em direção às melhores práticas de avaliação. Brasília: ENAP/Revista do Serviço Público, 2000. n.4; 5 p.

¹² An Evaluation Roadmap for a More Effective Government. American Evaluation Association, 2010. Available on: <http://www.evaluationcenter.net/Documents/aea09.eptf.eval.roadmap.pdf>. Accessed in: 09/02/2015.

¹³ *Evaluating Development Co-operation*: summary of key, norms and standers. Available on: <http://www.oecd.org/development/evaluation/dcdndep/41612905.pdf>. Accessed in: 09/02/2015.

- Should be user-driven - this just means that successful evaluations should be designed and implemented in ways that provide useful information to decision-makers, given the political circumstances, program constraints and available resources.¹⁴

According to Wholey, Hatry and Newcomer (2010) the concept of program evaluation is:

- The application of systematic methods to address questions about program operations and results;
- The methods used are based on social science research methodologies and professional standards; and
- The field of program evaluation provides processes and tools that [can be used]...to obtain valid, reliable, and credible data to address a variety of questions about the performance of public and nonprofit programs.¹⁵

As seen, there are many concepts and categories of evaluation, but to unify the understanding this paper will consider the concepts in accordance with Newcomer's suggestion: "... situating performance measurement and data analysis within the broader field of evaluation would be theoretically parsimonious and fruitful."¹⁶

This field of study is complex and can be explained when one compares the goals of the private and the public sector. In the private sector, usually the ultimate goal is

¹⁴ *Evaluating EU Expenditure Programs: A Guide*. Ex post and Intermediate Evaluation. First edition, January 1997. Available on: <http://www2.stat.unibo.it/cazzola/didattica/valutazione/Evaluating%20EU%20Guide.pdf>. Accessed in: 02/09/2015.

¹⁵ NEWCOMER, K. E.; HATRY, H. P.; WHOLEY, J. S. *Planning and Designing Useful Evaluations*. Handbook of Practical Program Evaluation. Jossey-Bass Publishers, 3rd ed., 2010.

¹⁶ NEWCOMER, K. E. and BRASS, C. T. *Forging a Strategic and Comprehensive Approach to Evaluation within Public and Nonprofit Organizations: Integrating Measurement and Analytics within Evaluation*. 2014.

making a profit, and it is a simple value to verify. However, in the public sector, the ultimate goal is people's well-being. Therefore, the question is "how can we know that the goal was achieved?" Program Evaluation offers many categories and methods to help evaluators work and address some essential features that cannot be neglected.

3.1 Important Features

The literature considers some essential features to incorporate in a program evaluation. They are reliability, validity and ethics. Reliability refers to how consistent the evaluation is. When testing this feature an evaluator must ask himself "are the results to which this evaluation tool producing steady and consistent?" Furthermore, according to Newcomer (2011):

"Measurement reliability is the extent to which a measurement can be expected to produce similar results on repeated observations of the same condition or event. Is a question asked in the same way? Is information collected in the same way from one item to the next? Would anyone get the same answer if they repeated the question or data collection task? If not, your evidence may not be reliable or, therefore, competent."¹⁷

Validity refers to the adequacy or accuracy of the measurement, in others words, it answer the question an evaluator may ask, "are we measuring what we need to measure? However, Newcomer (2011) also draws attention to validity as a multidimensional concept:

"Validity generally refers to the accuracy and representativeness of a study or data, and the term is sometimes casually used to characterize whole studies as valid or "scientifically valid." But validity is a multidimensional concept, and since its many dimensions are never attained perfectly, the question is not "Is the study or are the data valid"

¹⁷ NEWCOMER, K. E. Strategies to Help Strengthen Validity and Reliability of Data. Copyright. 2011.

but “How valid are they in terms of measurement, internal, external, and statistical conclusion validity?”¹⁸

Newcomer (2011) explains that limitations in one of these dimensions will affect the internal validity because the dimensions are considered like steps of a pyramid that has at the bottom measurement validity and measurement reliability, in the middle external validity and statistical conclusion validity and at the top internal validity. She also defines the dimensions making some questions:

- External validity – may the results be generalized?
- Internal validity – is there a causal relationship between a specified cause and potential effect?
- Statistical conclusion validity: do the numbers generated accurately detect the presence of a factor, relationship, or effect of a specific or reasonable magnitude?¹⁹

Ethics questions are present in evaluation and cannot be confounded with morality. Ethics is broader than the sphere of morality because ethics is associated with values that guide human behavior in society and morality are the customs and rules concerned about what is right or wrong by an individual unit.

The American Evaluation Association provides some principles for evaluators:

- Systematic inquiry - evaluators conduct systematic, databased inquiries about whatever is being evaluated.
- Competence - evaluators provide competent performance to stakeholders.
- Integrity/honesty - evaluators ensure the honesty and integrity of the entire evaluation process.

¹⁸ NEWCOMER, K. E. Strategies to Help Strengthen Validity and Reliability of Data. Copyright. 2011.

¹⁹ Ibid.

- Respect for people - evaluators respect the security, dignity and self-worth of the respondents, program participants, clients, and other stakeholders with whom they interact.
- Responsibilities for general and public welfare - evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare.²⁰

According to Kimmel (1988): “social research evaluators are often confronted with ethical problems because they assume the responsibility for determining what is beneficial in social programs or policies.”²¹ Furthermore, in some situations, there may exist a fine line between what is ethical or not and thus, the issues can be controversial.

3.2 Why Evaluate

Programs and policies are designed to produce desired outcomes, but policy makers, commonly focus on controlling and measuring the inputs and immediate outputs rather than the outcomes and impacts. If they achieve their medium and ultimate goals, it is not known. The resources need to be spent more effectively to improve well-being, in other words, to reduce poverty and to improve people’s lives. Despite this, policy makers may see evaluation as a problem because it exposes the absence of an efficient and structured method for planning and for implementation. Furthermore, sometimes the high costs to evaluate a program are another factor that inhibits it. The disclosure of the benefits is one way to improve program evaluation and to justify its high costs, because

²⁰ Guiding Principles for Evaluators - 2004. American Evaluation Association. Available on: <http://www.eval.org/p/cm/ld/fid=1>. Accessed in: 09/02/2015.

²¹ KIMMEL, A. J. Ethics and Values in an Applied Social Research. Applied Social Researches Methods Series vol. 12. Sage Publications, Inc, 1988.

commonly an improved and transparent welfare program is worth than the costs of its evaluations.

In this context, evaluators intend to make a performance report that all stakeholders may use to: ask for more budget, convince the public that the policy was a right choice, abandon ineffective policies, determine which results were below what was expected, and to report for accountability²².

3.3 How to Select a Program to Evaluate

The selection process must consider some basic criteria that, according to Wholey, Hatry and Newcomer (2010) are:

- Utilization–focus - can the results of the evaluation influence decisions about the program?
- Timing - can the evaluation be done in time to be useful?
- Significance - is the program significant enough to merit evaluation?
- Perception of problems - is program performance viewed as problematic?
- Life cycle - where is the program in its development?²³

Clear definitions of “significance” and “problematic” are important to diminish the process selection subjectivity.

²² GERTLER, P. J.; MARTINEZ S.; PREMAND P.; RAWLINGS, L. B.; VERMEERSCH C. M. J. Impact Evaluation in Practice. Interactive textbook, 2007. Available on: <http://www.worldbank.org/pdt>.

²³ NEWCOMER, K. E.; HATRY, H. P.; WHOLEY, J. S. Planning and Designing Useful Evaluations. Handbook of Practical Program Evaluation. Jossey-Bass Publishers, 3rd ed., 2010.

3.4 Evaluation Categories

There are different categories of evaluation, but to understand the context of this paper it is important to clarify the concepts. It is also important to perceive that the division is made to facilitate understanding and that one category does not exclude the other. On the contrary, one category should complement the other to improve the evaluation function. Evaluations are split up in following categories.

- *According to the evaluator agent*

External – when experts from outside of the institution conduct the evaluation. There are advantages and disadvantages of this type. The advantage of an external evaluation is the unbiased evaluation conducted by outside experts. The disadvantage is that it may be more difficult for outside experts to access program data, and program managers may be less cooperative.

Internal – when the evaluation is conducted by the program managers themselves or by a group within the institution.

Mixed – conducted by both, in and outside the institution, with the objective to maximize the advantages and minimize the disadvantages. The internal and the external groups work together.²⁴

Participatory – the program beneficiaries are included in the process and are surveyed. Their responses and perceptions of the program support the evaluation of the

²⁴ CUNHA, C. *Avaliação de políticas públicas e programas governamentais: tendências recentes e experiências no Brasil*. Elaborated during the course: “The Theory and Operation of a Modern and National Economy”. Minerva Program – George Washington University – 2006.

program, according to Spink (2001), Abma, (2004), Kayano e Caldas (2002), apud Ramos and Schabbach (2012)²⁵.

- *According to the focus*

Formative – this category of evaluation examines the formulation of the program and the results can enable some adjustment during the process of implementation of the program.

Summative – this category of evaluation is made at the end of the program to verify if it achieved the expected results.

- *According to the period of evaluation*

Before and after – this category of evaluation occurs on specific dates, ex ante, and ex post the implementation program. Ex ante evaluation is used to support the decision process before the beginning of the program. It uses a diagnostic based on the logic of the program. Ex post evaluation occurs after program implementation, in part or in whole, and intends to prove that the program achieves its goals, or if another one is necessary.

Ongoing or monitoring – this category of evaluation occurs during entire period of program implementation and can inform the decisions on whether the program should continue or not, and if the program needs some amendment during the process of implementation.

- *According to what is to be measured*

²⁵ RAMOS, M. P. and 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 46 (5): 1271-294, set/out 2012.

This category focus on what to measure about the process, results and impacts. It is split up in three categories: performance measurement, outcome or impact evaluation and cost, but the definition of input, output, outcome and impact are important to understand. Inputs are the resources used to produce outputs, outcomes and impacts. Outputs are the goods and the services delivered by the activities, and outcomes, and impact are the initial, intermediate and final results of an activity with different approaches. The categories are:

Performance Measurement – as well as process analysis, is related to the management dimension with a focus on activities, outputs and, sometimes outcomes. The difference is that the nomination “process analysis” is used just when the performance measurement is made with the objective of detecting the difficulties that happen during the process, and to enable some corrections. Performance measurement allows understanding if the program is being conducted as planned, and if it seems to be achieving its goals and target audience. It may be used to improve efficiency and effectiveness. According to Burt and Nightingale (2010, 210 p.) process analysis is:

“Process analysis focuses on the ways that programs work once they have reached a steady state of mature functioning. Its key questions are, how are people getting into the program, how are services being delivered, how are people exiting the program, and are these exits appropriate?”²⁶

The United States Government Accountability Office, GAO, describes performance measurement as:

“Performance measurement is the ongoing monitoring and reporting of program accomplishments, particularly progress towards pre-established goals. Performance measures may address the type or level of program activities conducted (process), the direct products and services delivered

²⁶ BURT, M. and NIGHTINGALE, D. Repairing the US social secure net. The Urban Institute Press, 2010.

by a program (outputs), and/or the results of those products and services (outcomes).”²⁷

Outcome and Impact Evaluation – the focus of some program evaluations is on outcomes, or on impacts that may be logically linked to the program and not to other contextual factors. For impact evaluation a granting cause and effect must be present. If this impact is demonstrated in at least one rigorous study, it is called evidence-based. The difference between performance measurement and impact evaluation is based on its goals. Performance measurement focuses on the ongoing performance and impact focuses on the impact of the same intervention in a wider context - in other words, if the program reaches the target audience and if it changes the reality that it was intended to change.

Cost Evaluation – there are two types of cost evaluation: cost-benefit and cost-effectiveness. Cost-benefit seeks to place dollar values on the total costs and total benefits expected, and verify whether the benefits outweigh the costs. It can be used on the planning phase of the program to evaluate its feasibility. Cost-effectiveness analysis compares the relative performance of two or more programs or program alternatives in reaching a common outcome, to verify each one will achieve in a better way the expected goal at a minimum cost. The cost information is important to complement the outcome and impact evaluations.

Although, according to Burt and Nightingale (2010, 209 p.):

“There is no single "right" research design for every policy question. Gathering and analyzing information to inform policy decisions cost money, so the choice of a research design should be approached sensibly, by asking what level and type of information are really needed. In each

²⁷ United States General Accounting Office, GAO. *Performance Measurement and Evaluation: definitions and relationship*. May 2005.

specific case. The question of what is possible also arises, and sometimes the question of what is ethical.”²⁸

This passage summarizes the difficulty of evaluation design and draws attention to the necessity of more discussion with the actors involved in the process, with the objectives of disclosing the subject matter, and enabling the growth of the field.

3.5 Evaluation Methods

For any kind of evaluation, the following are the most appropriate methods. Implementation analysis uses methods like fieldwork, observation, surveys, and administrative data. Performance analysis collects data on outputs and sometimes outcomes like program outcome data and performance measures. Outcomes and impact evaluation use an experimental, a quasi-experimental, and a non-experimental design to collect data.

Some of these methods require statistical knowledge and some of their characteristics are worth mentioning:

- Necessity of establishing a cause and effect relation;
- Necessity of collecting information before and after the implementation of the program;
- Necessity of comparing different but similar groups, before and after the implementation, because other factors may affect achievement of results.
- Necessity of using a comparison or a control group and a treatment group.

The relation between cause and effect must be tested in the design of the program, thus it may be done during the implementation to correct the course of the program to

²⁸ BURT, M. and NIGHTINGALE, D. Repairing the US social secure net. The Urban Institute Press, 2010.

achieve its objectives. To prepare the design, a Logic Model methodology that helps the policy makers and stakeholders to follow a line of reasoning between the problem and its goals that will enable its posterior evaluation during and after the implementation.

4 The Importance of the Logic Model

According to the “Instituto de Pesquisa Aplicada, IPEA:

“The construction of the logic model helps to organize the component stocks of a program in connection with the expected results, while setting the assumptions and ideas that give meaning to the intervention. Considered a tool to explain the program theory, the application of logic model results in a process that facilitates planning and communicating what is intended with the program and its expected operation. In particular, it can be used as a tool to carry out ex-ante evaluation of programs to improve the consistency of its original design.”²⁹
(Author`s translation)

With this intention, using a logic model has three phases:

- Identification of the problem to solve, through a broad diagnosis of compiling its causes;
- Construction of a systemic model that expresses the logical conditions that must be met to resolve the problem; and
- Construction of a management tool that records the solution strategy, in the form of an array of sequential goals to be achieved. At this stage, the indicators are incorporated into the program to monitor the management of the proposed solution.

²⁹ CASSIOLATO, M. and GUERESI S. *Como elaborar Modelo Lógico: roteiro para formular programas e organizar avaliação*. IPEA. Brasília, 2010.

Therefore, briefly, the problem and its causes and consequences (problem tree) are the first step to use this methodology. When you know the causes, you can suggest actions to combat these causes (solution tree). These actions must have indicators to enable post hoc evaluation. This framework helps to display the logic between the problem and the results. The filling of this model must be completed backwards and helps summarize the logic of the program and the inputs needed.

Context	Problem	Inputs	Process	Outputs	Outcomes		
					Short Term	Medium Term	Long Term
History, economy, politic, demography, law, regulation etc.	Situation to be changed	Money, building, equipament, teachers, staff etc.	number	number	%	%	%

5 Program Evaluation in United States

“If we believe the government can make a difference in people’s lives, we have the obligation to prove that it works – by making government smarter, and leaner and more effective.”³⁰

With this thought, on January 2011, President Obama signed the Government Performance and Results Modernization Act, GPRMA 2010, bill into Law. It is part of a series of laws passed since the 1990s designed to improve federal government management. This law requires the agencies to develop: strategic plans with long-term goals, performance plans with annual goals and measures and performance reports on prior year performance (quarterly and annual). It also requires a government-wide performance plan. The GPRMA of 2010 also requires:

“... a central inventory of all Federal programs. The Federal Program Inventory has the potential to facilitate coordination across programs by making it easier to find programs that can contribute to a shared goal, as well as improve public

³⁰ President Barack Obama. April 13, 2011. Available on: <http://archive.performance.gov/>. Accessed in: 03/05/2015.

understanding about what Federal programs do and how programs link to budget, performance and other information.”³¹

This inventory identifies programs by mission and indicates the strategic goal, the strategic objectives each program supports, and its budget. In the past two decades, the original Government Performance and Results Act of 1993 helped to set up a solid foundation of results-oriented planning measurement and reporting in the federal government, but little information was used for decision-making. To encourage greater use, GPRMA called for agencies to: demonstrate leadership commitment, align individual, program, and agency goals, improve the usefulness of performance information, build analytic capacity to analyze and use performance information, and communicate performance information frequently and effectively.

The agencies` Strategic Plans state the mission, the vision, the values, and a strategic plan framework with long-term goals, outcomes, objectives, and performance indicators with prior information and targets. These goals refer to a federal strategic plan.

Agencies also need to develop Performance Plans split by goal with a description of what will be or is being done and the programs linked, with performance indicators with prior information and targets. These goals are the same as stated in the Strategic Plan.

Annually each agency makes a Performance Report with the focus on the achievement of the goals and the relationship between the program and its mission. It is important to report the risk of non-achievement and action strategies. Quarterly each agency makes a report with a priority progress review and use of performance information for the current year. It is important to note that there is a manager for each performance goal and that all information is available on the web to guarantee

³¹ Available on: <http://www.state.gov/s/d/rm/rls/fs/2013/209930.htm>. Accessed in: 03/05/2015.

transparency. The GPRMA is a way the American government uses to improve the information to Congress and society with the final objective of improving supply of goods and services to society.³²

Another important action taken by the U.S. government is the evidence-based initiative to social policy as described on the site of the Office of Management and Budget - OMB. It says that not all the efforts intend to dictate a particular program design to everyone, “we only insist that most of the money go toward the programs backed by the best available evidence, and the rest to programs that are promising and willing to test their mettle.”³³ Agencies must provide credible evaluation results that show promise and OMB also talks about learning culture:

“By instilling a culture of learning into federal programs, we can build knowledge so that spending decisions are based not only on good intentions, but also on strong evidence that carefully targeted investments will produce results.”³⁴

6 The Pluriannual Plan of the Rio de Janeiro State

The Pluriannual Plan is a legal instrument that contains the programs, actions, products, and physical and financial targets for a quadrennium. It has the government strategies for the medium term and is able to be reviewed each year to suit the dynamic of planning. The Planning and Management Secretariat of Rio de Janeiro State – Seplag – according with its internal statute, is responsible for the Pluriannual Plan and its policy

³² LICHT, B. The GPRMA Modernization Act. Public Law 111-352. GAO. March, 2012. Available on: <http://www.slideshare.net/JMKamensky/gao-gpra-modernization-act-overview>. Accessed in: 03/05/2015.

³³ Building Rigorous Evidence to Drive Policy. Jun, 2009. Available on: <http://www.whitehouse.gov/omb/blog/09/06/08/BuildingRigorousEvidencetoDrivePolicy/>. Accessed in: 03/10/2015.

³⁴ Ibid.

cycle, supporting with methodology and system the agencies. The manuals are available on its site and they contain important concepts.

The PPA Elaboration Manual of Rio de Janeiro State defines program and action as follows:

“A program is the instrument of governmental organization action. It is an articulated group of actions (budgetary and non-budgetary), that seeks the achievement of a common target (...). The program is implemented through the execution of its actions (projects, activities, special operations and non-budgetary actions) that should be sufficient to achieve the program objective. (...).”(Rio de Janeiro, 2011. Author`s translation)

“Action is an operation or a set of operations which beget a product (good or service) with the intention of contributing to the program objectives.”³⁵ (Rio de Janeiro, 2011. Author`s translation)

In this context, the design of a program must consider its objectives and some actions that contribute to the achievement of these objectives. It brings targets to these actions for each year of the quadriennium. Unfortunately, the state does not have enough qualified staff to draw the programs in logic models, and to incorporate indicators linked to them. This instrument covers the policies of all agencies of the three powers (executive, legislative, and judicial), and is the basis for the elaboration of the annual budget.

6.1 Legal Aspects

Brazilian Federal Constitution (1988), art.165 and Rio de Janeiro State Constitution (1989), art. 209 dispose the central role of the PPA in the process of institutional planning

³⁵ RIO DE JANEIRO. Manual de Elaboração do Plano Plurianual 2012 – 2015. RJ: SEPLAG/SUBPL, 2011. Available on:

<http://download.rj.gov.br/documentos/10112/780695/DLFE-46913.pdf/ManualdeElaboracaoPPA20122015.pdf>. Accessed in: 03/05/2015.

of the federal states. In art. 167, the Constitution provides that no investment whose execution exceeds one fiscal year may be implemented without prior inclusion in the PPA, or by law to authorize such inclusion, subject to criminal liability. The Federal Constitution also establishes that the Pluriannual Plan, the Budgetary Guideline Law, and the Annual Budget Law are components of an integrated system of planning and budget, and that the Budgetary Guideline Law should guide the allocation of resources in the Annual Budget Law.

The Bidding and Contracts Rules to Public Administration Law - Law N^o. 8666/93, art. 7 provides that the works and services may only be offered when the product does not adhere to the goals established in the Pluriannual Plan, if applicable.

The Fiscal Liability Law – Law N^o 101/00 provides that the creation, expansion, or improvement of governmental action that increases spending, must be accompanied by the declaration that it is compatible with the Budget Law, the Pluriannual Plan, and the Guideline Budget law.

In other words, the importance of the Pluriannual Plan is recognized by law, but it is not enough to provide a qualified product to society and to government. Evaluation is needed to help achievement of objectives, and thus implementation, in order to secure expected results.

6.2 Program Analysis by the Rio de Janeiro State

According to the internal statute of the Planning and Management Secretariat of Rio de Janeiro State - Seplag, the agency must “coordinate and evaluate the formulation and implementation of public policies, as well as propose corrective alternatives and resizing of governmental actions, when necessary” (Resolução Seplag N^o 894 art. 2^o IV; 2013. Author`s translation). Thus, responsibility for coordinating the public policy

process is in the hands of the Undersecretary of Planning. However, the agencies are responsible for elaborating and implementing their programs, and answering for accountability. The Undersecretary supports the agencies by providing the methodology and the system, and is responsible for begetting two types of analysis that are published in the State Official Journal: a quarterly and an annual analysis.

In 2015, the first annual analysis will relate the fulfillments of physical and financial targets, and the second will bring the same content, but the latter will include accumulated data and a report explaining whether or not the targets are achieved. It is noteworthy that government efforts are being made to make a previous program inventory to understand which one are necessary to continue and to begin to understand their logic. Although the design of the programs with logic models, and related with indicators will be difficult to achieve because the servants were not trained in this field and time is scarce.

With the objective to improve the public policy process, in 2013, Seplag decided to create a Planning Network divided in groups for each agency whose components belong to both agencies (Planning and the Sector). This network is still underused, but should be the way to disseminate the culture of evaluation and to make it a reality.

7 Suggestions to Improve the Evaluation Function

Within the objective to improve evaluation function in Rio de Janeiro State, some suggestions to the process and design are provided here. However, the intention is to accelerate a discussion that may assist the state in some of its assignments. The Planning and Management Secretariat have already a robust instrument to start the process, the Planning Network created in 2013. This network is the bridge between the central agency

and sectors agencies, and can allow it to work together with the responsible stakeholders, identifying training needs, among other issues.

7.1 To the Process of Implementation

Grob (2008) has described the fundamental pillars for a national evaluation system. These are:

- Understanding evaluation types and methods and how policy makers and managers use them;
- Appreciating the relationships among evaluation and related professionals;
- Establishing government agencies that evaluate public programs;
- Providing freedom and encouragement for growth of non-government evaluation organizations;
- Providing evaluation education and training programs;
- Using professional standards.³⁶

As the evaluation importance must be disseminated, the beginning of the process should be to establish evaluation, bringing the final products (PPA, LDO, LOA and Quarterly and Annual Reports) with its processes, responsibilities and contents integrated with the evaluation function on all phases of the public policy cycle (planning, implementation, monitoring, evaluation and review). It is important to require the agencies to develop legal instruments with some additional content that allows evaluation, and designates a responsible party for each strategic goal, program and action. Because the responsibility to improve programs is not only on the Planning and

³⁶ GROB, G. F. Teaching Evaluation in South Asia – 2008. Available on: <https://drive.google.com/file/d/0BzSNHlTfgDTHSllqN21ldzdGblk/edit>. Accessed in: 03/08/2015.

Management Secretariat, but also on agencies, sectors should learn evaluation to show a commitment by all.

It is important to focus the training for the different purposes and target audiences, for example, for leaders who commit resources, for the ones who will coordinate the process, for the public servants who will replicate the information, and for the evaluators who will actually evaluate the programs. The features cited before (reliability, validity and ethics) should be part of this training. This great need to train people is caused by the existence of few specialized courses in universities in Brazil, because the field of evaluation is new in Brazil, and because public servants do not have this expertise. Encouraging the relationship with non-governmental organizations and academics, with experience in evaluation, is a significant action to promote the evaluation field and to help the State on its tasks.

It is important to draft instruments to be used by the State. This draft should contain the instruments and suggestions regarding comprehensiveness of evaluation, and the use of evaluation. For example, the following items must be included:

- State Strategic Plan for 10 years to be reviewed each 2 years with long-term goals per agency (Currently there is no rule about the content, the period and the link between the other instruments);
- Sectorial Strategic Plans aligned with State Strategic Plan for 10 years to be reviewed each 2 years with long-term goals (Today there is no rule about the content, the period and the link between the other instruments);
- State Sectorial Development Plans aligned with Federal Sectorial Development Plans;
- Pluriannual Plan – PPA for 4 years to be reviewed each year with short and medium term goals per agency aligned with Sectorial Strategic Plans and State Sectorial

Development Plans (Currently the goals are just physical and financial outputs for each four years).

- Budget Guideline Law per year linking the PPA priorities with resource allocation on Budget Law; and
- Budget Law per year.

Despite the fact that the Government Plan (elections commitments) is not a state instrument, its commitments must be considered in agenda setting for the PPA. Another important matter is the link between the agency mission described on its Strategic Plan and the programs the agency is planning to implement. Some programs are multisectoral, in other words, they need actions by more than one agency, but this relation between the mission and the products must be clarified.

The establishment of government agency that evaluates public programs, as quoted, may be too expensive, but a group of servants to coordinate the improvement of program evaluation shall be necessary. Finally, yet importantly, an improvement on the existent data systems is needed to support the information to be collected, and to support the evaluations.

7.2 To the Design

As evaluation design should be tailored to context, this paper will focus on the PPA and LDO whose responsibility is the Undersecretary of Planning. According to Wholey, Hatry and Newcomer (2004) evaluation design is:

“a plan for conducting an evaluation that specifies (1) a set of evaluation questions, (2) the data that will be collected and analyses that will be undertaken to answer the evaluation questions, (3) the estimated costs and time schedule for the

evaluation study, and (4) how the evaluation information will be used.”³⁷

To PPA, we must consider a sequence of items for the design that according to Teaching Evaluation in South Asia – TESA, are:

- Describe the program/project: why it is necessary and who are the stakeholders;
- Determine what will be achieved and how;
- Identify the issues and questions of relevance to the program/project;
- Decide on methods and process;
- Develop an implementation plan with clear check points;
- Collect information relevant;
- Analyze, evaluate and interpret information and draw conclusions based on evidence;
- Develop recommendations; and
- Disseminate results.³⁸

The process begins with the description of the program with its actions and expected outcomes. This paper suggests that a program should come to the agenda (PPA) if it includes an Explanatory Program Report that contains the context, the relationship with the other instruments, and a clear objective. The logic model, the desired actions and outcomes should also be included. In addition, indicators to enable performance measurement (inputs, process and outputs), results and impact (if possible) and goals to achieve in a specific time should be identified. The methods used to evaluate are

³⁷ NEWCOMER, K. E.; HATRY, H. P.; WHOLEY, J. S. Handbook of Practical Program Evaluation. Jossey-Bass Publishers, 3rd ed., 2010.

³⁸ GROB, B. F. Teaching Evaluation in South Asia – 2008. Available on: <http://www.teachingevaluationinsouthasia.org/index.php/getting-started/presentations>. Accessed in: 03/08/2015.

important at this moment to enable collecting data before program implementation, and depending on the type of evaluation, strategy treatment and comparison or control groups. The importance of the goals are cited by Metzenbaum (2006):

“Goals and measures are among the most powerful performance-improving and accountability-enhancing tools government has at its disposal. Even without an explicit link to incentives, goals and measures drive behavioral change both in individuals and in organizations. Goals do this by serving as a focusing point and by influencing attitudes, effort, and creativity.”³⁹

The quarterly and annual reports will support the exercise of accountability. Policy makers and managers must demonstrate, by agency and by goal, the relationship between each program and its mission, and the importance to the agency of the achievement of the goals. In addition, they should clarify the development of the program and the risk of not achieving the goals, with strategies to combat lack of success.

The LDO, that lists the PPA’s priorities to allocate resources, must link the programs and actions with the long-term goals. The agencies must justify priorities, and the Budget Law must respect these priorities in resource allocation. With the growth of the evaluation function, the selection of the priorities should be based on programs’ achieved results.

As seen, before making these reports imperative, training must have been offered to Planning and Management Secretariat and sectorial agencies responsible for their policies. Finally, the involvement of leadership and regulations are the key to making evaluation possible and successful.

³⁹ METZENBAUM, S. H. *Performance Accountability: The Five Building Blocks and Six Essential Practices*. IBM Center for the Business of Government. 2006.

8 Conclusion

The main objective of this paper was to make suggestions for the Rio de Janeiro State to improve its public policies, and as results its sustainable economic development by promoting the field of evaluation. The suggestions are intended to open discussion, and to accelerate the process by drawing attention to the importance of the evaluation. The field of evaluation is growing around the world and the State is not immune to this movement, but the actions that are needed, demand a great effort and a commitment from public servants and leadership to be implemented. Some issues have been highlighted here. The first one is the need for commitment of the leadership shown by including the content into legislation. Without legislation it will be difficult, especially considering the great effort to be undertaken and the need for standardization of work.

The second need is for training in evaluation. A learning culture, especially in this case, accepting evaluation field, its benefits, categories and methods, must be disseminated to public servants who work with public policies, and especially to those who will coordinate the process. Partnerships with academics may be a good solution to support evaluation training.

The third need is to development a system to keep data to organize the large amount of data that will be collected.

Fourth, an Explanatory Program Report should be required. This report should clarify evaluation and its relation with other instruments, clear objectives, its logic model, and consequent actions with goals to achieve. As the State has nearly 200 programs, it is impossible to work with all of them at the same time. A schedule that considers their degree of relevance should make the evaluation requirements possible to implement. The

concept of relevance should enable objectivity to be used in selection of programs to evaluate.

The public policy priorities extracted from LDO, as reflected in the Federal and State Constitution, should guide the allocation of resources. The choice of priorities in a first time should be made by sectorial agencies, based on the relevance of the programs. The concept of relevance to be a priority needs to reflect objectivity in selection. With the growth of the use of evaluation, the results of the programs can be used to base these selections, beyond on the relevance already mentioned.

Newcomer and Brass (2014) offered pertinent recommendations as they say:

“at a minimum, we suggest that public and nonprofit leaders need to think strategically about evaluation, to: connect and plan out the evaluation efforts undertaken to benefit from the synergies among them; enhance organizational learning from all evaluation efforts, including performance measurement and analytics; train managers and executives to enhance their evaluation competencies; envision evaluation and learning as a shared responsibility of all.”⁴⁰

It is worth noting that the network created by Seplag/RJ in 2013 may be one way to disseminate the evaluation culture to all State agencies. Ultimately, despite knowing that the effort is great and that the reality is not in accordance with what we hope, the State can engage in a step-by-step process and find partnerships to help embrace evaluation in the agencies.

⁴⁰ NEWCOMER, K. E. *Forging a Strategic and Comprehensive Approach to Evaluation within Public and Nonprofit Organizations: Integrating Measurement and Analytics within Evaluation*. 2014.

Bibliography

BRASIL. Constituição da República Federativa do Brasil. Brasília, DF: Senado, 1988.

BRASIL. Orçamento Cidadão 2015. Available on:

http://www.planejamento.gov.br/secretarias/upload/Arquivos/sof/PLOA2015/Orcamento_Cidadao_%202015.pdf. Accessed in: 02/09/2015.

RIO DE JANEIRO. Constituição do Estado do Rio de Janeiro. RJ: ALERJ, 1989.

RIO DE JANEIRO. Resolução 894/2013. RJ: SEPLAG, 2013.

Available in: <http://download.rj.gov.br/documentos/10112/1587843/DLFE-61786.pdf/RES.SEPLAG894REGIMENTOINTERNOVERSAOFINAL.pdf>.

RIO DE JANEIRO. Manual de Elaboração do Plano Plurianual 2012 – 2015. RJ: SEPLAG/SUBPL, 2011.

Available on: <http://download.rj.gov.br/documentos/10112/780695/DLFE-46913.pdf/ManualdeElaboracaoPPA20122015.pdf>. Accessed in: 02/12/2015.

ALA-HARJA, M.; HELGASON, S. *Em direção às melhores práticas de avaliação*. Brasília: ENAP/Revista do Serviço Público, 2000. n.4; 5 p.

AMERICAN EVALUATION ASSOCIATION. *An Evaluation Roadmap for a More Effective Government*. 2010.

Available on: <http://www.evaluationcenter.net/Documents/aea09.eptf.eval.roadmap.pdf>.
Accessed in: 09/02/2015.

BURT, M.; NIGHTINGALE, D. *Reparing the US social secure net*. The Urban Institute Press, 2010.

CASSIOLATO, M.; GUERESI S. *Como elaborar Modelo Lógico: roteiro para formular programas e organizar avaliação*. IPEA. Brasília, 2010.

CUNHA, C. *Avaliação de políticas públicas e programas governamentais: tendências recentes e experiências no Brasil*. Elaborated during the course: “The Theory and Operation of a Modern and National Economy”. Minerva Program – George Washington University, 2006.

DYE, T. D. *Understanding Public Policy*. Englewood Cliffs, N.J.: Prentice Hall, 1984.

GERTLER, P. J.; MARTINEZ S.; PREMAND P.; RAWLINGS, L. B.; VERMEERSCH C. M. J. *Impact Evaluation in Practice*. Interactive textbook, 2007. Available on: <http://www.worldbank.org/pdt>.

GROB, B. F. *Teaching Evaluation in South Asia*. 2008.

LASSWELL, H. D. *Politics: Who Gets What, When and How*. The American Political Science Review. American Political Science Association, 1936.

LICHT, B. *The GPRMA Modernization Act*. Public Law 111-352. GAO. March, 2012.

LYNN, L. E. *Designing Public Policy: A Casebook on the Role of Policy Analysis*. Santa Monica, Calif.: Goodyear, 1980.

METZENBAUM, S. H. *Performance Accountability: The Five Building Blocks and Six Essential Practices*. IBM Center for the Business of Government, 2006.

NEWCOMER, K. E. *Strategies to Help Strengthen Validity and Reliability of Data*. Copyright. 2011.

NEWCOMER, K. E.; HATRY, H. P.; WHOLEY, J. S. *Handbook of Practical Program Evaluation*. Jossey-Bass Publishers, 3rd ed., 2010.

NEWCOMER, K. E. *Forging a Strategic and Comprehensive Approach to Evaluation within Public and Nonprofit Organizations: Integrating Measurement and Analytics within Evaluation*. 2014.

OCDE. *Evaluating Development Co-operation: summary of key, norms and standers*. Available on: <http://www.oecd.org/development/evaluation/dcdndep/41612905.pdf>. Accessed in: 09/02/2015.

POISTER, T. H. Performance Measurement in WHOLEY, J. S.; HATRY, H. P.; NEWCOMER, K. E. *Handbook of Practical Program Evaluation*. Jossey-Bass Publishers, 3rd ed., 2010, 100-124 p.

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* 46 (5): 1271-294, set/out 2012.

SOUZA, C. *Políticas Públicas: uma revisão da literatura*. *Sociologias*, ano 8, n. 16, 020-45, jul/dez, 2006.

STACK, K. Advisor for Evidence-Based Innovation at the U.S. Office of Management and Budget. Available on: <http://www.businessofgovernment.org>. Accessed in: 09/02/2015.

United States General Accounting Office, GAO. *Performance Measurement and Evaluation: definitions and relationship*. April 1988.

KIMMEL, A. J. *Ethics and Values in an Applied Social Research*. *Applied Social Researches Methods Series* vol. 12. Sage Publications, Inc, 1988.