

## **PROGRAM EVALUATION**

**Author: Raluca Antonie**

**Scientific Referents: Dr. Roger Hamlin (Michigan State University),  
Dr. Calin Hințea (Babeș-Bolyai University)**

## CONTENTS

CHAPTER I. PROGRAM EVALUATION – THE MAIN CONCEPTS .....	3
CHAPTER II. EVALUATION TYPES AND SYSTEMS.....	16
CHAPTER III. PROGRAM EVALUATION MODELS .....	25
CHAPTER IV. NEW EVALUATION MODEL: PAEM (PUBLIC ADMINISTRATION EVALUATION MODEL) .....	35
CHAPTER V. THE COUNTERFACTUAL EVALUATION MODEL (CEM).....	49
CHAPTER VI. INTRODUCING A MODEL FOR SOCIAL IMPACT ASSESSMENT OF PUBLIC ADMINISTRATION REFORM IN ROMANIA .....	62
CHAPTER VII. CONCEPTS OF RESEARCH METHODS AND STATISTICS USED IN PROGRAM EVALUATION .....	72
CHAPTER VIII. THE ROLE OF PROGRAM EVALUATION IN THE DECISION- MAKING PROCESS .....	84
CHAPTER IX. EVALUATION CULTURE AND CAPACITY IN ROMANIAN PUBLIC INSTITUTIONS AT REGIONAL AND LOCAL LEVEL.....	94
CHAPTER X. LEVERS SUPPORTING PROGRAM EVALUATION CULTURE AND CAPACITY IN ROMANIAN PUBLIC ADMINISTRATION: LEADERSHIP .....	112

## **CHAPTER I. PROGRAM EVALUATION – THE MAIN CONCEPTS**

**Summary:** *Evaluation of projects and programs in public administration is a specific stage, extremely useful in planning and management, a research technique and a tool for public policy, successfully used by institutions responsible for management and organizations, the coordination of projects and programs supported from public or private funds. Senior civil servants, politicians, managers, institutions and organizations, program directors and coordinators of projects use various evaluation models to notify, in due time, the effects of interventions they implement or intend to implement. The aim is to seize and to counteract the unwanted effects for the groups of people, communities and society and to encourage the positive elements of programs and projects.*

**Key concepts:** *evaluation, program, project, monitoring, audit, relevance, efficiency, effectiveness, sustainability, impact*

## **Introduction**

Evaluation study is relatively new, especially in Romania. It started as a practice field and became a recognized discipline of scientific research since 1960. During its evolution as a scientific domain, there were a series of disputes over terminology, methodology and ethics of evaluation. In 2004, Fitzpatrick, Sanders and Worthen have identified nearly 60 models published between 1960 and 1990 (Fitzpatrick, Sanders, and Worthen, 2004). 1965 meant the start of massive U.S. programs to combat poverty (War on Poverty - the general name of the set of programs) (Sorin Dan Sandor, 2005: 73).

## **Great Society. Watts riots. War against poverty**

A package of reforms in education, health, social welfare and environmental protection were part of President Johnson's plan to rebuild America as a Great Society (The Great Society). Initially, it has the full support of public opinion. Great Society to represent the most comprehensive reform of the New Deal promoted by Ronald Reagan. In 1960, despite the climate of prosperity, almost a quarter of American families lived below the poverty line. Whole regions have not reached economic prosperity in the postwar period. Moreover, technological developments in the industry leave jobless people unprepared. In 1964, 44% of people over 65 had no health insurance. The poor state of health with age due to their automatic entry cause among the poor. In fact, more than one in three Americans under 65 lived below the poverty line. Therefore, President Johnson held that "the war against poverty" (War on Poverty) should occupy the center of concerns for building large companies. A significant impetus to initiate war against poverty have been the Watts' riots. Riots in Watts (Watts' Riots) were a series of large-scale civil conflict, which erupted in the Watts neighborhood of Los Angeles, in the south of the United States of America. They started in 11 August 1965 and lasted six days. In the end, 34 people were killed, 1100 injured, 4000 personnel were arrested, 600 buildings were burned and there were about 100 million dollars damage. Most of the damage were recorded for businesses that have caused resentment in the neighborhood due to perceived social injustices. Homes were not attacked, but took some heat because of proximity to other buildings on fire. A government commission was appointed to investigate the riots. The main cause has been established as unemployment, poor schools and other inferior living conditions.

Costs caused by the war in Vietnam, with the costs of reform programs have resulted in greater pressure on the U.S. economy. With the decline in popularity of the war in Vietnam, President Johnson lost his political capital necessary to continue reforms.

Critics claim that the Great Society reform programs had effects far beyond the expected, if not reverse, despite the large public investments. There were created only new bureaucratic systems that have swallowed money without producing results. The best example is circulated welfare reform that has produced a fragmented city, with the most harmful consequences. It was responsible for creating welfare dependency among the poor, to create vicious circles of poverty.

Large sums of money were invested to eradicate social problems, but desired results were still not present. It was concluded that there was not enough money to solve all problems. It was also noted that regarding the problems faced by communities, money is not the only problem. Evaluation of programs launched by assuming these two constraints (Michael Quinn Patton, 1997) by the budgeting process-Programming Planning Budgeting System (PPBS), which encourage monitoring and evaluation. Public opinion has become increasingly cautious, seeking evidence for the success of programs that would be supported by public money. In these circumstances, evaluation experienced a real boom, the number considered essential for the success of a program dramatically increased.

As anticipated (Chelimski, Shadish, 1997) evaluation methodology is continually diversifying, beyond national borders, and is extremely useful in a broad range of growing concerns.

### **Raw material for program evaluation**

The raw material of evaluation consists in projects, programs and public policies, activities and strategies, etc. Project level is the reference evaluation level. A project is conceived as an organized effort to implement an idea. Of course, we mean in this particular study the socio-economic development projects. Key elements of a project include: goals, objectives, actors (initiators, direct and indirect beneficiaries, donors, etc.), activities, timing, resources and multiplier effects. Projects may be initiated and implemented by various entities including:

government institutions, NGOs and even companies in the private sphere. Usually projects are implementing programs or specific objectives of programs. The program has the degree of generality higher than the project, but respects a similar structure. Implementation of a program is achieved through several projects, detailing and implementing one or more of the objectives of the program. Public policy is the drive with the highest degree of generality, corresponding to a strategic action in a particular field. As an example, building administrative capacity of public administration in Romania is a specific public policy. A program corresponding to this policy is the Modernization of Local Public Administration. One of the projects implementing this program is Program Evaluation in Public Administration. The project is applied at the level of municipalities and apply an objective of the program: increased accountability for spending public money. Programs and projects can be funded by institutions of central and local government, international organizations (EU, World Bank, etc.), nonprofit organizations and other entities. Usually, the sponsor is interested in project results, evaluation of proposed objectives. In many cases, public institutions are co-financing partners in development projects that affect groups of people, communities and the whole society. Furthermore, good management involves organizing public institutions work on a project basis, to be more easily managed, improved, monitored, evaluated and controlled. This trend is supported and encouraged throughout the European Union. Need assessment in the administrative system is growing. Public administration reform and administrative capacity development requires the development of evaluation capacity. This can be achieved by creating a legal framework, the institutional capacity and human resource necessary for evaluation issues.

### **Definitions of evaluation**

Throughout its evolution, evaluation has received many definitions. A summary of the different dictionary definitions for evaluation reveals some key terms: determining merit, value, estimation, appreciation, etc.. Evaluation of projects or programs is closely related to the meaning of those terms. The assessment definitions also appear as a constant number of elements of the methodology.

In his book on "Analysis and research in public administration", Dan Sandor describes synthetic evaluation programs that relate to a systematic analysis to see how well projects and programs were implemented as intended and achieved their objectives.

European Commission proposes five criteria relevant to the evaluation: relevance, efficiency, effectiveness, sustainability and impact. On this basis, we define evaluation as the process by which, using methods and instruments, we can measure the degree to which project objectives and results were relevant, economic resources are consumed to achieve the objectives, whether the project is likely to continue after funding ends, the extent to which activities reach their target group and whether their impact is felt long after the implementation process ends. The key to a correct understanding the difference between systematic and continuous in evaluation of the resulting difference between evaluation and (sequential- conducted systematic but in certain moments in the life of a program) and monitoring (continuous data collection during the implementation of a program). "The term evaluation refers to the collection, analysis and reporting on information that can be used to change attitudes and improve a project or program." (Allum, 2000: 3)

Summarizing, we emphasize the following key elements of evaluation/assessments:

- Assessment is a useful tool in management policies, programs and projects;
- It involves judgments based on criteria;
- Assessment is useful in any stages of developing a program: In the design stage, before implementing a program (ex-ante); During the implementation of a program (interim); After implementation (ex-post);
- Assessment is an explanatory process: based on some questions for which answers are found;
- It is more comprehensive than monitoring;
- Evaluation involves a systematic and scientific analysis (collecting data, analyzing them, comparing them based on certain criteria);
- Assessment is based on granting information and explanations about the program assessed for decision-makers that can lead to alteration of design or modification of the implementation process. Such decisions may concern the continuation, modification or end of the program.

According to one of the classical authors in the evaluation field, Michael Quinn Patton, assessment is the systematic collection of information about activities, characteristics and outcomes of programs in order to be used to reduce uncertainties, improve effectiveness and to make decisions appropriate in connection with the programs "(Michael Quinn Patton, 1997: 23). This definition, if not understood properly can cause confusion. The key to a correct

understanding is the correct understanding of the difference between systematic and continuous that has been explained before.

To understand what evaluation is, one needs to distinguish between evaluation or assessment and other terms sometimes incorrectly used as interchangeable. Different authorities or agencies give different meanings to the same terms. As we live within the general context of the European Union and we refer to the evaluation of programs in the public sphere, we adopt the terminology used by the European Commission in assessing the programs.

### **Monitoring**

Monitoring refers to quantifying the project or program implementation results in real time throughout the development of objectives, resource consumption, reaching the target group, systematic quantification of changes arising from implementation of the program or project. As a result of the monitoring process, one can permanently seize the input – output report, income, expenses, planned activities, activities conducted, proposed target group, reached target group, recording any discrepancies. Assessment explains why these discrepancies exist (if any). Monitoring is descriptive, while evaluation is explanatory. The link between monitoring and evaluation is very strong. Evaluation cannot be done (or can be performed extremely difficult), if there is no coherent monitoring system. This usually involves a set of indicators and a monitoring plan. Specific resources should be allocated to the monitoring process. Usually, human resources are mobilized for monitoring project implementation. But it is possible, especially for large projects, that the monitoring activity is performed by qualified personnel from outside the project team.

### **Audit**

The audit is a review of the financial provisions of a project and the extent to which these criteria are met in accordance with legal and technical requirements. European Commission auditors and agents assigned to the audit period have a much wider range of meanings: verify project need, the extent to which project or program activities and results justify the financial investment, the existence and visibility in place of a plus-value generated by the project or program (Hughes, 2000: 3).



## **Evaluation**

Evaluation is a process by which to measure program performance and identify solutions to existing problems. More specifically, the assessment may be, among others, to analyze the results of a program to compare its costs, to help authorities respond to the citizens for their actions to help the allocation of resources and help improve their programs.

Data derived from an assessment are valuable for improving program implementation, as well as the decision-making. The effectiveness of projects and programs is determined by the answers one gives to the following questions: What works?, For whom? and Under what circumstances? The assessment also supports the planning of future activities, the distribution of human and financial resources, etc.

## **Evaluation Criteria**

What does the evaluation team measure? The answer is simple: the evaluation criteria. The EU Commission approach (Tavistock Institute et al, The Evaluation of Socio-Economic Development: The GUIDE, 2003) is based on five main criteria:

- Relevance
- Effectiveness
- Efficiency
- Impact
- Sustainability

**Relevance** refers to the extent that the program meets the real needs of those concerned. It also takes into account any changes in context that may result in some changes in the type of needs that the program should address or change in their hierarchy. A program that is relevant throughout its implementation take into account these changes in context, has the flexibility needed to replicate as many times as necessary to meet the needs it targets. And a program becomes irrelevant when, during its existence, fails to meet the need that is proposing to cover or cover them incorrectly, reported in the original plan. One aspect of relevance is the need for certain programs to be supported by public money. What are these programs? What are the reasons why the state should be involved in its implementation? Private or nonprofit

sector has failed to meet the needs of stakeholders? Why? To find out if a program is relevant and to what extent, the evaluation team must consider all these elements.

**Effectiveness** considers the extent to which programs achieve their objectives. Also, the degree to which project results meet the needs identified in the design phase is a measure of program effectiveness.

**Efficiency** takes into account an additional, essential element in the existence projects and programs: the financial issue. Moreover, efficiency takes into account the following aspects:

- Could be obtained the same results, in the same circumstances, with fewer resources?
- Unit costs are too high?
- Even if goals are met, is the project / program too expensive to be continued?

The net **impact** is the effect produced exclusively by a program. Because of the numerous external variables distorting the impact of a project, is difficult to calculate accurately the net impacts. It is difficult to differentiate the effect derived exclusively from a program in an extremely complex socio-economic context. However, using appropriate methods it is possible to give the answer to the following questions, with an assumed margin of error:

- What changes have resulted from the program?
- Are there other benefits of the program, along with the expected ones?

Another sense of the impact considers the long-term effects of a program.

**Sustainability** refers to the continuity of the program after withdrawal of funding from the original source.

- The effects of programs continue after the end of the implementation?
- Can there be identified alternative sources of funding?

Sustainability gives, along with other criteria to measure performance of a project or program. Besides these criteria, others can be mentioned just as important for assessing program performance. First, we refer to equity and community involvement.

Equity refers to issues such as access to services provided by the project regardless of age, gender, social and material conditions. Usually projects are promoted to discourage discrimination of any kind.

Community participation is another criterion to be taken into account in evaluating certain projects.

All these criteria, along with others that we have not mentioned, but can be just as important, are used in evaluation process.

In addition to consideration of as many of the criteria listed, the evaluation should be analytical, systematic, reliable, focused on issues or users, depending on the model used for evaluation.

## References

1. Allum, N.C, Bauer, M.W., G. Gaskell, *Quality, Quantity and Knowledge interests, Avoiding confusion in Bauer and Gaskell, Qualitative research*, London, Sage, 2000.
2. Chelimski, E., Shadish, W.R. (Editors, *Evaluation for the 21<sup>st</sup> Century*, SAGE Publications Inc.), 1997.
3. Chen, H. T., *Theory-driven evaluation*. Thousand Oaks, CA: Sage Publications, 1990.
4. Chen, H. T., *Practical Program Evaluation: Assessing and Improving Planning, Implementation, and Effectiveness*, Sage Publications Inc, 2005.
5. Chiriac D., Humă, C., Impactul socio-economic al fenomenelor naturale dezastruoase în România - inundații, alunerări de teren, secetă (62 p.), Probleme Economice vol. 20-21/2002 CIDE, 2002.
6. Cook, T. D., Campbell, D. T., *Quasi-experimentation: Design and Analysis Issues for Field Settings* Chicago, Rand McNally, 1979.
7. Curley, Hilary, Eugen Perianu (2006): *Assessment of the Evaluation Culture in Romania*, [http://static.anaf.ro/static/10/Mfp/evaluare/Assessment\\_Romanian\\_Eval\\_CultureEN.pdf](http://static.anaf.ro/static/10/Mfp/evaluare/Assessment_Romanian_Eval_CultureEN.pdf) (accessed on February, the 07<sup>th</sup>, 2012).
8. Dale, A., T et all, *Social Assessment in Natural Resource Management Institutions*. CSIRO Publishing. PO Box 1139, Collingwood, Victoria, 3066 Australia., 2001.
9. Fitzpatrick, J.L., Sanders, J.R., și Worthen, B.R., *Program Evaluation*. (3<sup>rd</sup> Edition) Boston: Pearson, 2004.
10. Goss, J.D., Leinbach T.R., 1996, *Focus groups as alternative research practice*, Area 28 (2): 115-23.
11. Hansen, H. F., *Choosing Evaluation Models, A Discussion on Evaluation Design*, 2005, in Evaluation, Vol. 11, No. 4, p. 447-462, SAGE Publication.

12. Hughes, J., *Project Manager's Guide to Evaluation*, <http://www.evaluate-europe.net/handbook/handbookpdf>, 2000.
13. Ilie, S., *Formal and informal incomes of the Romanian Households* in Stanculescu, M, R. Neef (editori) *The Social Impact of Informal Economies in Eastern Europe*, Ed.Ashgate U.K., 2002.
14. Iluț, P., *Abordarea calitativă a socioumanului*, Polirom, Iași, 1997.
15. Innes de Neufville, J., *Social Indicators and Public Policy: Interactive Processes of Design and Application* (New Brunswick, NJ: Transaction), 1975.
16. Kubish, A., Schorr, L. et al, *Ne approaches to evaluating community initiatives*. Washington DC: The Aspen Institute, 1995.
17. Malan, J., *Benchmarking Evaluation Capacity in the New Member States*, Centre for Strategy and Evaluation Services, 2004.
18. Marginean, I., *Politica Socială. Studii 1990-2004*, Expert, 2004.
19. Patton, M. Q., *Utilization-Focused Evaluation: The New Century Text*, 3<sup>rd</sup> ed., Beverly Hills: Sage Publications, 1997.
20. Patton, M. Q., *Qualitative Research and Evaluation Methods*, SAGE, 2002.
21. Rossi, P., Freeman, H., Lypsey, M., *Evaluation, A Systematic Approach*, 6th ed., SAGE Publication, 1999.
22. Scriven, M., *The evaluation thesaurus* (4<sup>th</sup> ed.). Thousand Oaks, CA: Sage, 1991.
23. Scriven, M., Hard-won lessons in program evaluation. *New Directions for Program Evaluation*, 58. San Francisco, CA: Jossey-Bass,1993.
24. Scriven, M., Evaluation in the new millennium: The transdisciplinary vision, 2003, in S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the millennium* (pp. 19-42). Wahwah, NJ: Lawrence Erlbaum Associates.

25. Scriven, M., *The transdisciplinary model of evaluation: Radical implications*. Paper presented at the meeting of the American Evaluation Association/Canadian Evaluation Society, Toronto, Canada, 2005.
26. Shadish, W. R. Jr., Thomas, D., Cook, L., Leviton, C., *Fundamentele evaluării programelor: Teorii ale practicii*, FIMAN, 1999.
27. Stake, R., *Standards-Based and Responsive Evaluation*, SAGE, 2003.
28. Stanculescu, M, R. Neef, editors, *The Social Impact of Informal Economies in Eastern Europe*, Ed. Ashgate U.K. , 2002.
29. Stanculescu, M. (coauthor), *Toward Country - Led Development. A Multi-Partner Evaluation of the Comprehensive Development Framework. Findings from Six Country Case Studies: Bolivia, Burkina, Faso, Ghana, Romania, Uganda, Vietnam*, IBRD, The World Bank Washington USA, 2003.
30. Stanculescu, M. (coauthor), *Pre-Accession Impact Studies. EU Cohesion Policy and Romania's Regional Economic and Social Development (bilingv)*, Institutul European (ed.) , 2003.
31. Stufflebeam, D. L., *Evaluation models. New Directions for Evaluation*, 89. San Francisco, CA: Jossey-Bass, 2001.
32. Stufflebeam, D. L., *The 21<sup>st</sup> century CIPP model*. In M. C. Alkin (Ed.), *Evaluation Roots* (pp. 245-266). Thousand Oaks, CA: Sage, 2004.
33. Stufflebeam, D. L., *CIPP model (context, input, process, product)*. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 60-65). Thousand Oaks, CA: Sage, 2005.
34. Șandor, S. D., *Analiză și cercetare în administrația publică*, Ed. Accent, 2005.
35. Tavistock Institute, *The Evaluation of Socio-Economic Development: The GUIDE*
36. Vanclay, Frank, 2003, *The International Handbook of Social Impact Assessment: Conceptual and Methodological Advances*, Edward Elgar Publishing - ISBN: 1840649356, 2003.
37. Weiss, C. H., "Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families." , 1995 in James P. Connell, A. C., Kubish, L., Schorr, B. and Weiss, C. H. (eds.) *New Approaches to Evaluating*

*Community Initiatives. Vol.1: Concepts, Methods, and Contexts.* Washington, D.C.: The Aspen Institute

38. Wholey, J. S., Hatry, H. P. , Newcomer, K. E. (Editors), *Handbook of Practical Program Evaluation* 2nd Edition, 2005.

## CHAPTER II. EVALUATION TYPES AND SYSTEMS

**Summary:** *There are many classification criteria that can be used in order to present the types of evaluation: the objectives of the assessment, the time when the evaluation takes place during the life of a program, after the unit of analysis, the position of the evaluation team members, etc.*

**Key concepts:** *formative and sumative evaluation, ex-ante, interim and ex-post evaluation, reactive and pre-determined evaluation*



## **Introduction**

The main classification criteria that can be used in order to present the types of evaluation are: the objectives of the assessment, the time when the evaluation takes place during the life of a program, after the unit of analysis, the position of the evaluation team members.

### **From the perspective of the objectives of the assessment:**

**Formative evaluation** is usually done during the implementation in order to analyze the situation and to facilitate program improvement.

This terminology was first used by Michael Scriven (Scriven,1991), after being taken from all relevant sources in the field.

**Summative evaluation** is used for analyzing the results at the end of the program or at the end of a stage in the development of the program in order to determine program performance and to estimate its progress in reaching the objectives.

Robert Stake offers an extremely suggestive definition of summative and formative assessment: "When the cook tastes the soup, it is formative assessment, and when the guests taste the soup it is summative evaluation" (Stake, 2003: 52).

Robert Stake, evaluator of educational programs in the United States of America in the '60s and '70s, is one of the first advocates of the use of qualitative methods in evaluating social programs. Stake perceives evaluation as a service and as a reflection of values.

Shadish, Cook and Levinton consider that programs should be evaluated selectively, and defines evaluation as "an approach that sacrifices some accuracy in the hope that selection would increase the value of discovery for people outside and within the programs" (Shadish Jr., Cook, Leviton 1999: 278).

### **From the perspective of methodological flexibility: Pre-determined and reactive assessment**

Robert Stake differentiate between pre-determined assessment recognized by: focus on objectives, using objective tests, use of standards held by program managers in designing the research and evaluation reports and reactive type, characterized by the following elements:

- Put more emphasis on program activities rather than his intentions.
- Focus on the public needs.
- Perspectives on the value of actors are taken into account in reference to program failures and successes.

The term "reactive" (derived from the relationship stimulus - response) promotes the idea of a methodologically flexible evaluation process. One of the main disadvantages of pre-assessment is that researchers (traditional quantitative) focuses on variables that can not be controlled by management, thereby losing utility. Standardized indicators generated by initial evaluation design proved often less relevant than the indicators that are built spontaneously in the program (by contact with side players, or according to further discussions that were pursuing an activity, etc.).

Some key features are specific to responsive assessment: focus on observation and flexibility, preference for qualitative methods and focus on improving local practices. Reactive evaluation has several advantages and disadvantages. Benefits include: highlighting important variables of the program, change the role of those involved in a program to encourage increased local control. However, "pre-defined assessment should be preferred to reactive evaluation when it is important to know whether certain goals have been met, certain promises were kept and the hypotheses to be investigated have been tested. We can expect ... predetermined measurements are more objective and reliable." (Shadish Jr., Cook, Leviton 1999: 283) In conclusion, both types should be considered depending on the methodological flexibility needed.

Reactive evaluation has a number of common features with formative assessment, while pre-defined evaluation partly overlaps summative evaluation.

Characterized by some critics as being "superficially attractive" (Shadish Jr., Cook, Leviton 1999: 317), reactive-predefined distinction remains valid in theory because it highlights certain aspects of the evaluation, that remain otherwise in the shadow: the importance of methodological flexibility usefulness of qualitative methods, the emphasis should be placed on activities, not only on targets etc. Both are considered relevant and useful to the public.

### **3. According to the time of evaluation:**

- Ex-ante evaluation
- Interim evaluation
- Ex-post evaluation

### **Ex-ante evaluation**

Ex-ante evaluation (Tavistock Institute, 2003: i, ii) is an assessment made in the first stage of a program or project cycle (planning and design phase), before being implemented. Ex-ante requires a SWAT analysis, in which will be considered the defining characteristics of the locality, region, state that implements the project, a needs analysis and some simulations of socio-economic effects. This assessment ensures relevance and coherence of the program depending on context.

If it must be selected a program to be implemented first from several programs, which will be selected by tender, the ex-ante evaluation helps determine project selection criteria and the selection of projects to be financed. Also the new status of Romania as an EU member state should be in accordance to specific compliance regulations and EU standards in providing quality goods and services.

### **Interim evaluation**

Interim evaluation is performed in the second stage of a project or program cycle: during implementation. The purpose of this assessment is to improve the design and the implementation of a project or program. Interim evaluation has in common with formative assessment several elements: the progress of the objectives until the assessment, how well the schedule and the budget has been respected, utilization of other resources, etc. By comparison with the initial situation, in the interim evaluation can be highlighted certain relevant changes in the socio-economic context that can affect the program. Interim evaluation use information from monitoring and from the ex-ante assessments. Usually interim evaluation involves peer review of the interim results of the project, but it is recommended as well a detailed analyze of the likely impact that has not yet had time to manifest, but is inherent. Based on the findings, peer can improve both design and program management and any predictable negative impacts can be prevented.

## **Ex-post evaluation**

Ex-post evaluation is an analysis of the entire program considered primarily from the perspective of comparison to the initial results and in terms of impact. There are many common elements of ex-post evaluation and summative evaluation. In addition to results and impact analysis, the ex-post can use the following methods: benchmarking, cost-benefit and cost-effectiveness, process analysis and multi-attribute analysis.

The purpose of ex-post is multiple: quantifying the intentional or non-intentional results and effects of a program, qualitative and quantitative analysis of performance, learning lessons to improve future management programs, evaluate program performance in comparison with the performance of other similar programs, etc.

## **From the perspective of the position of the evaluation team members**

From the perspective of the position of the evaluation team members there are two basic types of evaluation:

- Internal evaluation
- External evaluation

**Internal evaluation** is an evaluation performed by the personnel implementing the program within the institution. Usually, the implementing institution has qualified staff and data necessary for the evaluation. However, in Romania, there are many institutions (if we consider the public sector) or organizations (if we take into account the non-profit) that do not have the internal evaluation capacity (people not specialized in evaluation), although there are experts in the various sectors relevant to the project under evaluation. These institutions will use, even for internal evaluation program evaluation services of outside experts, preferably independent (without any connection with the project team evaluated or with the project).

Purpose of internal evaluation is to provide an analysis from the perspective of implementation team who has access to data easier than any other actor involved (either donor or external evaluator). Therefore, this type of evaluation is extremely rich in data and the evaluation report is very clear and explanatory. Those who have implemented know the best

program theory, processes and results and why certain changes were implemented in the initial design, possible reasons for which the original schedule was not met, for the exceeded budget or, conversely, for not having used the resources that have been allocated, along with the distorting elements and what effects have these items on the program.

Internal assessment is very suitable for formative assessment, helping to control quality assessment and develop internal evaluation capacity. Disadvantages of internal evaluation are: lack of sectoral expertise and lack of independence.

**External assessment** is the evaluation performed by independent evaluators, usually outside the institution or organization implementing the program or who are among the actors participating in the program. The main advantages of this type of evaluation are independence and potential of a wide range of expertise. External evaluation is especially appropriate for summative assessments. The disadvantages of this type of evaluation are possible pressures that can limit independence, it does not help internal evaluation capacity development and the high costs involved.

### **Other types of assessment**

Besides the types of assessment mentioned above, the literature mentions other types of assessment as well, including participatory evaluation and assessment based on theory.

In **participatory evaluation**, the evaluator's perspective is on equal footing with the actors involved in the program perspective. The intention is to have an evaluation with conclusions and recommendations relevant and useful for future projects of the actors involved. This type fits with both the summative evaluation and the formative one. Involvement of all participants on equal footing usually brings a significant addition of information in the evaluation process. Participatory evaluation is often presented as a modern assessment and evaluation is presented as opposed to traditional (Sorin Dan Sandor, 2005: 81)

**Evaluation based on the theory** is applied by some researchers in the field of addictive substance abuse (Chen, 1990) or evaluating comprehensive community initiatives (Weiss, 1995). This type of evaluation is characterized by the lack of statistical analysis of data, mostly because of their diversity. Therefore, there are used mainly qualitative research methods. Usefulness of this type of evaluation is especially evident in the evaluation of

community initiatives whose effects can not be analyzed statistically, but can provide information about the efficacy of this type of initiative. Some researchers (Schorr, Kubisch, 1995) argue that, by combining data about the results of a program with information on the process of project implementation, we can obtain valuable information on the effects of the program and its impact.

Evaluation based on the theory assumes that any social intervention or program is based on a theory about how a particular process, under what conditions will achieve maximum efficiency and effectiveness, etc. This theory can be implicit or explicit. Theory of a program is roughly equivalent to its logical model that explains how it works. This evaluation provides many enlightening information on how to implement similar complex programs, indicating risks, key elements and lessons learned during implementation.

**Impact assessment.** In the analysis of impact there are measured on the one side the net effects of the intervention (net impact) and, on the other, program's or project's effects on medium and long term. The main problem of determining the net impact is differentiating the effects of the program and those due to other causes/variables. The net impact can be determined both before implementation (estimated impact), during implementation and after. Medium and long term impact can be anticipated prior to implementation and can be calculated during and after implementation. Determining whether the net impact or medium and long term effects, impact assessment can be extremely useful for improving current and future projects' design, to base decisions on continuing or stopping certain initiatives.

### **Evaluation systems: centralized assessment and evaluation of decentralized**

The question for evaluation systems is commonplace in government issues: what is the optimal level of centralization / decentralization in the evaluation of programs. As in the case of public services, there are advantages and disadvantages, whether there is an option for centralization or decentralization. While excessive centralization lacks flexibility, induces no administrative discretion, decentralization may lead to lack of coherence, use of poor methodology, etc. For this reason it is often chosen an intermediate solution: neither centralization nor excessive decentralization. Particular attention should be paid to the needs of each industry. Furthermore, although particular attention will be paid to evaluation EU funding programs (PHARE, SAPARD, structural funds, etc.) other programs funded from

domestic sources, but also from other external sources should not be ignored. The evaluation will thus gain in complexity.

The proposed model for Romania by specialists of Evaluation Central Unit for 2007-2013 is a combination of centralized and decentralized model.

## References

1. Chen, H. T., *Theory-driven evaluation*. Thousand Oaks, CA: Sage Publications, 1990.
2. Chen, H. T., *Practical Program Evaluation: Assessing and Improving Planning, Implementation, and Effectiveness*, Sage Publications Inc, 2005.
3. Cook, T. D., Campbell, D. T., *Quasi-experimentation: Design and Analysis Issues for Field Settings* Chicago, Rand McNally, 1979.
4. Scriven, M., *The evaluation thesaurus* (4<sup>th</sup> ed.). Thousand Oaks, CA: Sage, 1991.
5. Scriven, M., Hard-won lessons in program evaluation. *New Directions for Program Evaluation*, 58. San Francisco, CA: Jossey-Bass, 1993.
6. Scriven, M., Evaluation in the new millennium: The transdisciplinary vision, 2003, in S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the millennium* (pp. 19-42). Wahwah, NJ: Lawrence Erlbaum Associates.
7. Scriven, M., *The transdisciplinary model of evaluation: Radical implications*. Paper presented at the meeting of the American Evaluation Association/Canadian Evaluation Society, Toronto, Canada, 2005.
8. Shadish, W. R. Jr., Thomas, D., Cook, L., Leviton, C., *Fundamentele evaluării programelor: Teorii ale practicii*, FIMAN, 1999.
9. Stake, R., *Standards-Based and Responsive Evaluation*, SAGE, 2003.
10. Șandor, S. D., *Analiză și cercetare în administrația publică*, Ed. Accent, 2005.
11. Tavistock Institute, *The Evaluation of Socio-Economic Development: The GUIDE*



## CHAPTER III. PROGRAM EVALUATION MODELS

*Programs and project evaluation models can be extremely useful in project planning and management. The aim is to set the right questions as soon as possible in order to see in time and deal with the unwanted program effects, as well as to encourage the positive elements of the project impact. In short, different evaluation models are used in order to minimize losses and maximize the benefits of the interventions upon small or large social groups. This chapter introduces some of the most recently used evaluation models.*

*Key-concepts: evaluation models, evaluation focused on beneficiaries, on experts or on different stakeholders, realistic evaluation, theory-driven evaluation, utilization-focused evaluation*

## **The knowledge stage in the area at international level**

Presently, the knowledge stage in the area on international level is extremely advanced, in spite of the relatively short time passed since the first systematic approaches in the field have appeared.

The essential contributions come from the academic research, from the non-profit sector and from the international organizations, implementing a series of programs and projects in numerous states and communities (The European Union Commission, The European Bank for Reconstruction and Development (EBRD), The World Bank (WB), The International Monetary Found (IMF), The Organization for Economic Co-operation and Development (OECD), The United States Agency for International Development (USAID) etc.)

### **Academic research**

The research regarding programs and project evaluation models developed extensively in the last years. The design of the evaluation models and the selection of the right evaluation model for a certain project or program became the aim of numerous investigations. This is specific to the classical authors in the field: Michael Quinn Patton, Michel Scriven, Peter H. Rossi. They were joined by an increasing number of researcher most of them coming from the academic research field: Daniel Stufflebeam, Egon Guba, Yvonna Lincoln, Ernest R. House, Kenneth R. Howe, Hanne Foss Hansen etc.

The toolkit is vast: formative and summative evaluation, evaluation focused on beneficiaries, on experts or on different stakeholders, realistic evaluation, theory-driven evaluation, utilization-focused evaluation are just a few examples.

An evaluation model stipulates the question or the set of questions that a specific evaluation seeks to answer. It also involves a certain methodology to set up the criteria for assessment (Hansen, 2005). The literature on programme evaluation and that on organizational effectiveness offer several typologies of evaluation models. Hansen (2003, 2005) and Scriven (2003) propose some of the most recently appeared and comprehensive typologies. These mainly consists in six different categories of

models that are common at some point to other authors as well such as Birkmayer and Weiss (2003). The six categories are: results models, process models, system models, economic models, actor models and programme theory models.

The results models are interrelated with summative evaluation and they focus especially on the results of a programme. Among the subcategories of the results models there are: *goal-attainment model* and *the effects model*. The *goal-attainment model* results are measured according to the goals that have been set. The main question is: Have the set goals been achieved? The *effects model* focuses on consequences of the evaluated program. It is about both the desired and not desired consequences. In this case, the question is: What are the effects of the program? What are the positive and negative consequences of the program?

Process models focus on the processes involved by a program. This is an explanatory model. Process evaluation is done usually concomitant with the implementation phase of the program (in real time), or by historical analyses. The main questions are: How are the the activities implemented? Are there any delays? If so, How are they motivated?

The system model uses a system perspective. What matters here is: input, structure, process and outcome in terms of results. The evaluation consists in comparisons of planned and realized input, structure, process and results or in benchmarking. In this case, the main question is: How has the program functioned as a whole?

The economic models (Cost-efficiency model, Cost-effectiveness model, Cost-benefit model) are considered related to the system perspective. (Hansen, 2005) The question specific to *cost-efficiency model* is: How is productivity? Is it satisfactory or not? The question connected to the *cost-effectiveness model* is: How satisfactory is the effectiveness? The *cost-benefit model*, focus on utility: How satisfactory is economic utility?

Actor models (Client-oriented model, Stakeholder model, Peer review model), are based on the actors' perspective. The *client-oriented model* focuses on the clients' perspective. Are clients satisfied? *The stakeholder model* focuses on all the relevant stakeholders perspective. Are the different groups of stakeholders satisfied with the program? The *peer review model* focuses on the opinions of professionals. Does the program respond to professional standards?

The programme theory model focuses on assessing the validity of the programme theory on which the given intervention occurs. The target of the programme theory model is to continually improve program theory according to the changing context. The main questions are: What works as established? What exactly has changed as a consequence of the changing context?

In 1997 Vedung had already depicted evaluation models as being organized into three main classes, very similar to the above-mentioned categories (E. Vedung, 1997). Definitely, there is an increasing number of models proposed by the professional literature. Beside the models already presented, the next evaluation models are relevant due both to their actuality and their possibility of being applicable to the Romanian Public Administration.

The CIPP model (Stufflebeam, 2002) - context, input, process, product- focuses on effectiveness and sustainability, metaevaluation and synthesis. Therefore, the starting question: Is the program a success or a failure? Did the program reach the target group? What are the needs that have been satisfied by the program? etc. The evaluation criteria are derived from the aim and the objectives of the program.

The constructivist model promoted by Egon Guba and Yvonna Lincoln in 2001, (Egon Guba & Yvonna Lincoln, 2001), is based on three fundamental assumptions: ontological, epistemological and methodological. According to this model, the main questions are: Who is going to use the evaluation?, What is the perspective/What are the perspectives of the evaluation? etc. The evaluation criteria are established by agreement by the actors involved in the program.

The qualitative model presented by Michael Q. Patton in 2003, (Michael Q. Patton, 2003) focuses on the utilization of qualitative methods (observation, individual interview, focus-group, Delphi method) for evaluating programs, especially when it is about finding out more details about the specific program. The questions specific to this evaluation model are naturally of qualitative nature: What was the manner in which the objectives have been accomplished? Why certain objectives have been dropped?

Utilization-focused evaluation model (Michael Q. Patton, 2002) has many similar features with formative evaluation. According to this model, the evaluation process starts together with the design of the project, and ends after its implementation. The

focus is on the different utilities given to evaluation by the stakeholders. According to these, more specifically according to the aim and the objectives of the evaluation, the structure of the evaluation takes shape. The main questions are: Who are the users of the evaluation?, What are their objectives? What are the pieces of information they need? etc. The evaluation criteria are established by the users of the evaluation.

### **International Organizations**

The Commission of the European Union, The European Bank for Reconstruction and Development (EBRD), The World Bank (WB), The International Monetary Fund (IMF), The Organization for Economic Cooperation and Development (OECD), The United States Agency for International Development (USAID) etc. use evaluation to guide their investment and intervention projects in different geographical areas and fields of activity.

The World Bank uses evaluation at large scale. With an entire department engaged in evaluation, WB makes such studies for each of the projects it finances. WB experts assess the impact following rigorous methodologies, clearly stated. On the web site of the WB there are presently 134 documents that can be accessed and that contain elements of evaluation. There are made available to the public the methodology, the data collection and their sources, and, selectively, certain evaluation studies organized according to the type of impact and to the country where it has been studied. Numerous evaluation studies are catching the eye due to their diversity of methods and of the projects under investigation. A series of handbooks are being presented, that explain the methodology and processes involved by an evaluation. Among the most significant is Judy Bakers' *Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners, Directions in Development*, World Bank, Washington, D.C. edited in 2000. In *Sourcebook for Poverty Reduction Strategies*, World Bank, Washington D.C., appears during the same year *Monitoring and Evaluation* written by Prennushi, G., G. Rubio, and K. Subbarao . An impressive series of articles written by WB experts among who we can mention James J. Heckman, Jeffrey A. Smith, Nancy Clements, Christopher Taber Grossman, Jean Baldwin Karen Fulbright-Anderson, Anne C. Kubisch and James P. Connell and many others.

The distinctive feature consists in the fact that the vast majority of studies are made on

WB projects focused on the fight against poverty. Therefore, considerable stress is placed upon the economic dimension of evaluation. But, from the perspective of the evaluation models, the socio-economic profile of World Bank projects promotes the complementary use of three evaluation models: economic (which focus on the financial input), result models and qualitative model. It is quite natural, considering the mission of the WB is fighting the poverty. There is no major difference between the models use by the World Bank and those proposed and used by the academic community.

Otherwise, there is no major distinction between the tools used by WB and those built by the academic community.

Still, the wide geo-political and cultural area of action is obvious in the methodology used by the WB and especially in the diversity of variables and tools.

Recently, the EU Commission published a guiding catalogue of indicators that should be considered in evaluation. Among these, there are: Social Cohesion (social integration, poverty or extreme poverty dimensions, the risks of poverty or social exclusion, geographical social cohesion, long term unemployment, the accessibility of services of general interest), Employment Quality (occupational health and safety arrangements, the rights of the workers, labor market organization, the balance between personal and professional life, employment opportunities, integration through employment, etc) Social Protection and Social Services (levels of social protection, accessibility etc.), Consumer Interests, Education, Social Capital, Livable Communities, Fundamental Human Rights, etc.

As well on the site of the Commission, there is a Handbook for the implementation of evaluation: [www.evaled.info](http://www.evaled.info). This explains, step- by- step, the procedure that must be used for an evaluation study within the European Union. Still it must be particularized for each country and cannot be applied as such. The evaluation models promoted by the European Union Commission are mixed, at the intersection between economic model (which focus on the financial input), result models and qualitative model.

### **The Non-Governmental Organizations (NGO)**

The Non-Governmental Organizations, especially the grant makers are usually

interested in evaluation. The main donors developed their own toolkit for evaluation and use their own models. For instance, Ford Foundation, USAID, Rockefeller as well as others have made public their instruments for evaluation, accompanied by numerous case-studies. One of the goals is, for sure, the accountability of their actions. It is worth mentioning though, that their evaluation toolkit is in perfect agreement with the methodology specific to the academic research and to international organizations. Certainly, the evaluation models and the research methods and techniques for social sciences need not to be reinvented. But their application is in accordance to the interests of the financing entities and to the cultural background of the researchers who conceive the instrument and effectively realize the research.

### **The present knowledge stage in Romania**

Regarding the program evaluation models, the Romanian scientific literature is relatively scarce. There are studies regarding the evaluation models, but most often they refer to technical evaluation, strictly economic evaluation (such as the country risk indicators) or specific to other fields (constructions, environmental protection, software etc.) and not program evaluation, or program evaluation financed or co-financed from public money: country risk evaluation models, evaluation models for the safety of the buildings, evaluation models for the polluting agents dispersion into the atmosphere, evaluation models for the cost of software testing, etc.

### **Trends**

Worldwide, the tendency is to use more than one evaluation model at a time (Hansen, 2005:448). But they have to be carefully chosen and adjusted. Evaluation studies and reports have to answer to more and more questions regarding the process and the results of the projects; it is of interest at the same time economic effects and social impact effects. Especially in the case of the projects financed from public money the tendency is to use evaluation models that focus on results and models that focus on the beneficiaries' (citizens') perspective. To accomplish this, it is necessary the use of some evaluation models focused both on results and on the actors involved. These are complex models generated by the selection analysis and synthesis of simple evaluation models.

The evaluation models that are used in order to evaluate a certain program must fit the evaluation objectives, the project development stage. On the other hand, the evaluation methods and the research methods should be carefully chosen, in perfect agreement with the evaluation models and the evaluation objectives.



## References

1. Birckmayer, J. D., Hirschon Weiss C., „Theory-Based Evaluation in Practice”, 2003, *Evaluation Review* 24(4): 407–31.
2. Dale, A., Taylor, N., Lane M., „*Social Assessment in Natural Resource Management Institutions*”, CSIRO Publishing, 2001
3. Greene, J. C., „Understanding social programs through evaluation” in N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research*, 2nd. ed. (pp. 981- 999). Thousand Oaks, CA: Sage, 2000
4. Foss, H. H., „Choosing evaluation Models”, 2005, in *Evaluation*, 11(4), Sage Publications
5. Lincoln, Y. S., Guba, E. G., „The Qualitative Inquiry Reader”, CA: Sage, 2001
6. Patton, M. Q. „Practical Evaluation”, Newbury Park: Sage Publications, 1982
7. Patton, M. Q., „Qualitative research and evaluation methods” (3rd ed). Thousand Oaks, CA: Sage, 2002
8. Rossi, P. H., Lipsey, M. W. and Freeman, H. E., „Evaluation: A Systematic Approach” Thousand Oaks, CA: SAGE, 2004
9. Schwandt, T. A., „Dictionary of qualitative inquiry”, Thousand Oaks, CA: Sage, 2001
10. Scott, W. R., „Organizations: Rational, Natural and Open Systems” Englewood Cliffs, NJ: Prentice-Hall, 2003
11. Scriven, M., „Evaluation Theory and Metatheory”, in Thomas Kellaghan, Daniel L. Stufflebeam si Lori A. Wingate (editors) *International Handbook of Educational Evaluation*. Dordrecht: Kluwer Academic Publishers, 2003
12. Stufflebeam, D., L., „Fundational Models for 21<sup>st</sup> Century Program Evaluation”, in Daniel L. Stufflebeam, George F Mandaus si Thomas Kellaghan (eds.) „Evaluation Models: Viewpoints on Educational and Humain Services Evaluation”, pp. 33-84. Boston, MA: Kluwer Academic Publishers, 2000
13. Stufflebeam, D. L., „The metaevaluation imperative”, 2001, in *American Journal of Evaluation*, 22(2), 183-209
14. Stufflebeam, D.L., „CIPP Evaluation Model Checklist”, 2002, on <http://www.wmich.edu/evalctr/checklists/cippchecklist.pdf>

15. Vedung, E, „Public policy and program evaluation”, London, Transaction Publishers, 1997.

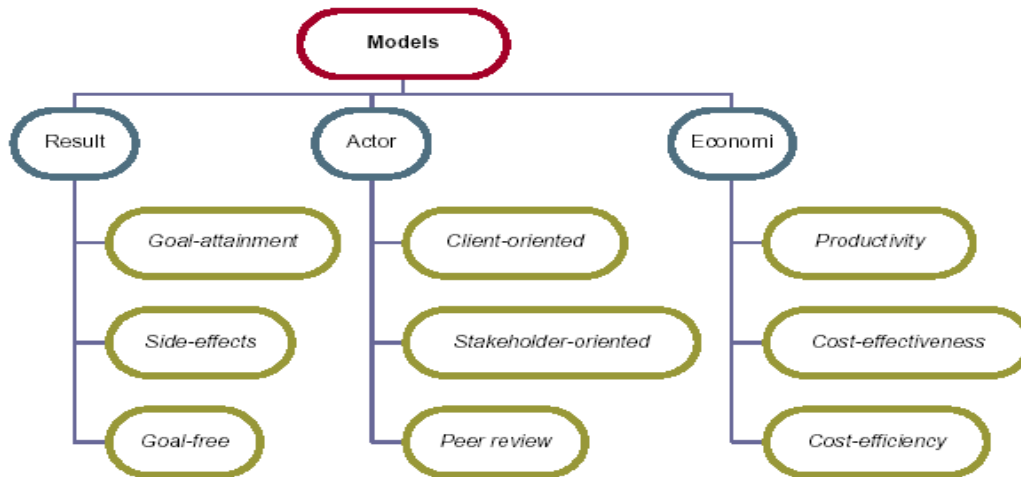
#### **CHAPTER IV. NEW EVALUATION MODEL: PAEM (PUBLIC ADMINISTRATION EVALUATION MODEL)**

*In order to be effective, evaluation should be based on solid monitoring systems. If the monitoring system is missing, what should be done? In this respect, the new model, called PAEM (Public Administration Evaluation Model), is a possibility. It generates structured programs and projects out of mere unstructured activities of public institutions. It helps building an indicator system that has a double functionality: it secures the future existence of a monitoring system and provides a functional evaluation toolkit.*

*Key-concepts: PAEM (Public Administration Evaluation Model), structured programs, indicator system, monitoring system.*

## The need for a new model, adjusted to Romanian realities

In 1997, Vedung had described evaluation models as being organized into three main classes, as shown in the figure below.



Source: Vedung (1997), "Public Policy and Program Evaluation", Transaction Publisher.

Figure 1. Evaluation Models according to Vedung (1997)

After studying the applicability of different existing evaluation models in the Romanian public institutional framework, we encountered several difficulties that made the applied models difficult to apply if not unpracticable.

The main cause is that most existent evaluation models are based on a monitoring system and capacity. In the case of Romania, it does not exist. Another cause is related to the fact that the existing evaluation models assume an evaluation culture and capacity, which is more than a reporting culture. As we have presented in another study (Gârboan, 2007), Romania's public administration has just passed the pre-culture stage of evaluation culture and it entered a developing stage of an evaluation culture.

Table 1. Evaluation culture. Distribution on frequencies

### Evaluation Culture

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2 preculture (11p-20p)	35	36.1	41.2	41.2
	3 weak culture (21p-30p)	46	47.4	54.1	95.3
	4 developed culture (31p-40p)	3	3.1	3.5	98.8
	5 extremely developed culture (41p-48p)	1	1.0	1.2	100.0
	Total	85	87.6	100.0	
Missing	System	12	12.4		
Total		97	100.0		

Therefore, there is a big proportion (41.2%) from Romanian investigated institutions that are situated on the pre-culture zone of evaluation. These still have to make progresses on data management systems, using previous experiences in improving general activities and even in organising their activities based on programs.

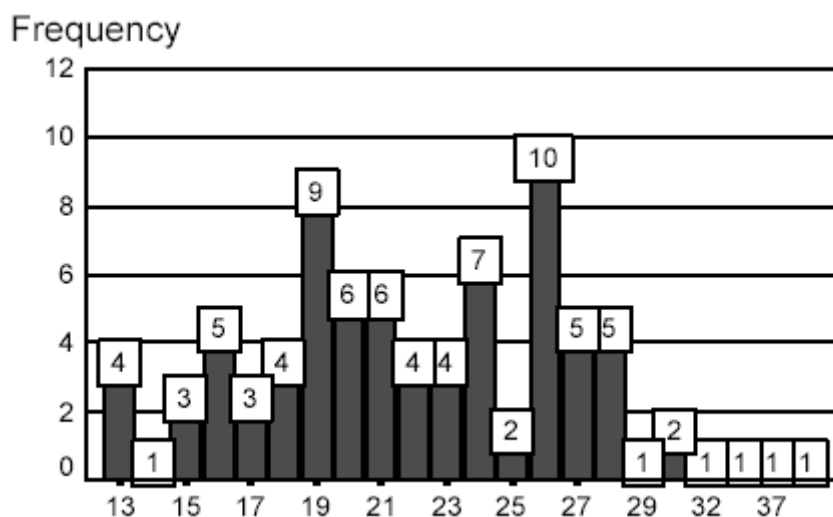


Figure 2. Evaluation culture

Figure 2 represents the repartition of investigated public institutions on the scale used to measure evaluation culture.

At his turn, when measured, the evaluation capacity is quite weak. 45.1% of the respondent institutions, have a very low capacity of evaluation, while 52.4% have a low capacity.

Tabelul 2. The capacity of evaluation. Frequency distribution.

capacity of evaluation\_grouped together values (alt\_4.7 + con\_4.13 + ext\_4.14 + cap\_4.15 + cultev)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2 very low capacity (16p-30p)	37	38.1	45.1	45.1
	3 law capacity (31p-45p)	43	44.3	52.4	97.6
	4 high capacity (46p-55p)	2	2.1	2.4	100.0
	Total	82	84.5	100.0	
Missing	System	15	15.5		
Total		97	100.0		

By watching the graphics of the absolute values which are not grouped together (Figure 3), we can clearly see that the mode value is 26, which is similar to the mode value discovered at the culture of evaluation.

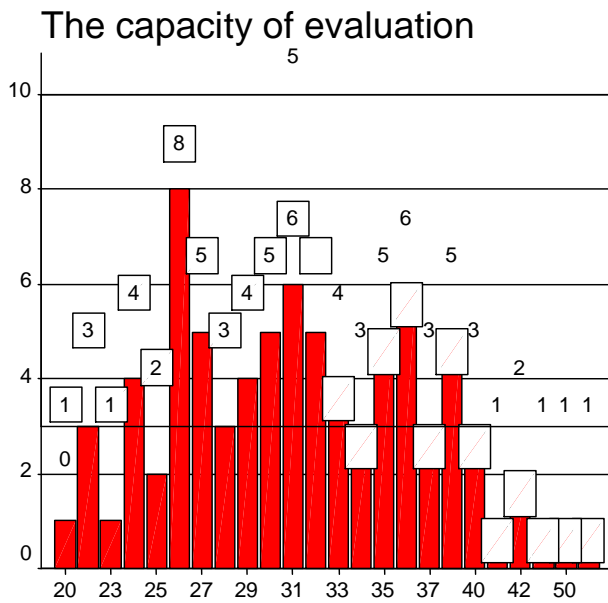


Figure 3. The capacity of evaluation (absolute values)

What is more, concerning the expertise in evaluation, the human resources who received training in this domain, or have practical experience in program evaluation, the situation is just as dramatic (see Figure 4 and Figure 5).

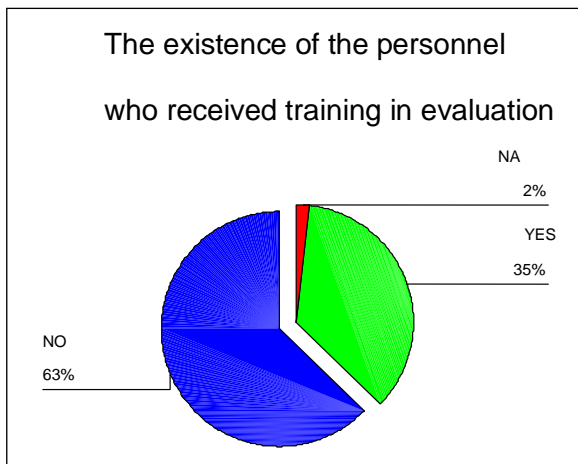


Figure 4. The existence of the personnel who received training

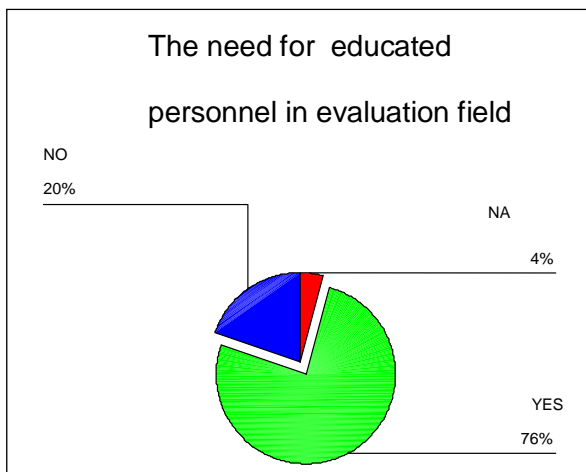


Figure 5. The need for educated personnel in evaluation field

63% of the public institutions which were questioned don't have in their structures specialized personnel in evaluation field (Figure 4) and 76% are aware of the existence of this need (Figure 5), indicates that 13%, even if they have specialists in evaluation, are conscious of the fact that the need for evaluation specialized personnel is even bigger. This fact shows the

tendency to development of the capacity of evaluation and the unquestionable existence of the evaluation culture in Romanian public institutions.

The evaluation capacity is also given by the estimation capacity of the necessary resources, which, as the present study indicates, is very low, only one third of the respondents being able to estimate the costs of an evaluation.

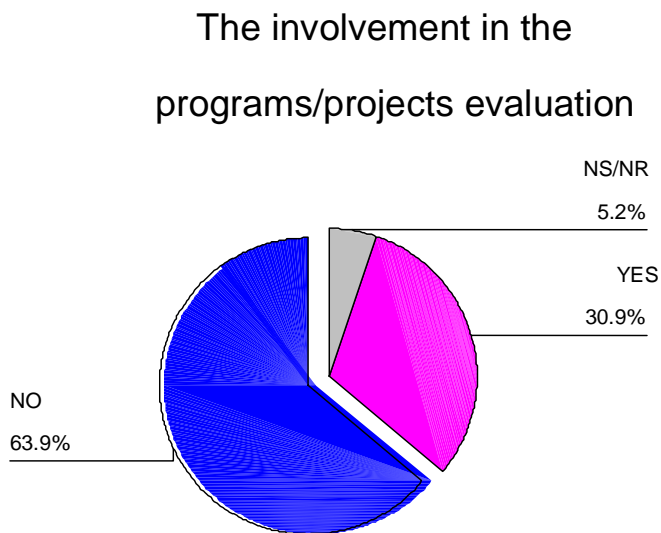


Figure 6. The involvement in the programs/projects valuation

The study's results also reveal (figure 6) that the institutions' involvement in evaluations of projects and programs is extremely low (only 30.9%), even though the most of the respondents have participated in such evaluations and have evaluated themselves when it came to projects which required European financing.



## The expertise gained through training/practical activities

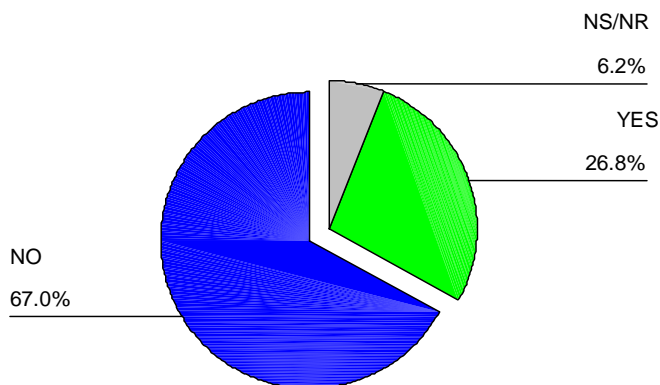


Figure 7. The expertise gained through training/practical activities

If it is about the expertise gained through training and/or practical activities, the situation is even more dramatic (Figure 7). Here from it comes the need for training in evaluation and research methods fields, and the emphatic need for research projects which aim at projects and programs evaluation with the implications of the public sector employers.

Generally, the results models, the process models, the system models, the economic models, the actor models and the programme theory models requires, for correct application a solid monitoring and evaluation culture and capacity. Romania is at its very beginning in this area.

Romanian public institutions are connected with program monitoring and evaluation mainly in the context of the programs and projects financed from pre-accession funds (PHARE, ISPA, SAPARD) which have had well-established monitoring and evaluation systems. This led to an initial development of the monitoring and evaluation culture in Romania. This has been an important step, even if it still is more about reporting than it is about evaluating.

At the half of the year 2005, the Management Authority<sup>1</sup> of Public Finance Ministry, began developing a National Evaluation Strategy based on a technical assistance contract financed by PHARE 2003. The National Evaluation Strategy, launched in November, 2006, is an important instrument that could gather evaluation's tendency of development in Romania. The interest for

---

<sup>1</sup> Which coordinates the evaluation and monitoring of PHARE and the Operational Programs funded from EU structural instruments

evaluation is growing but the precise request for evaluation is just at the beginning. There aren't any examples presenting evaluation as being built in lack of a list of programmes financed by European or external funds. The evaluation request depends on the existence of a legislation which provides a controlled evaluation and general acknowledged necessities and utility of implementation politics, strategic management and defined budget course.

Nowadays, the problem is the lack of laws blocks the development of evaluation system. It seems to be a malfunctioning mixture between supervision and evaluation functions.

As a consequence of studying the applicability of several evaluation methods in the context of Romanian public institutions, we developed a new model that can be used to evaluate programs especially when evaluators cannot benefit of a well established monitoring and evaluation system. This is an appropriate model not only for Romanian institutional realities, but also for other developing countries.

The new model, called **PAEM (Public Administration Evaluation Model)**, is composed of 10 different stages, of which, the first one deals with program design or standardizing and the others are about evaluation.

### **1. Program's standardization**

Romanian public administration has not passed to a program and project budgeting system, despite several attempts in this direction<sup>2</sup>. That is the reason why numerous activities of public institutions have not been conceived as programs and projects. Though, they have many functional particularities that make them feasible for monitoring and evaluation. That is why, these activities can be called unstandardized programs and projects. They do not have clearly specified objectives, activities, time framework and other program elements. In order to be able to evaluate them, standardization of an institution activities is required. By this process, most of the public institutions activities could be turned into programs and projects. The gain would be an easier and more effective monitoring and evaluation. In order to standardise, we could use a Program's or Project's Form.

### **The Program's/Project's Form**

Activities standardization is necessary whenever we need to turn them into programs and projects in order to monitorize and/or evaluate them. This could be achieved by simply filling in a Program's Form. This is an instrument that must contain several elements, such as the

---

<sup>2</sup> Law of public finances 1989/1998

context of the program, the organizational structure, the documents of the program, the actors involved, the specific activities, the time framework, the resources involved, the expected results as well as other significant elements according to each case.

**The context of the program** involves a short history of the activity that is being standardized (details on its origin and its initiator), the summary of the specific activities of the program and of the delivery methods for different services, information on similar activities and services as well as details concerning the unicity of the program. Another element of the Program's Form is **the organizational structure** which refers mainly to the institution's levels of functioning and control and to the distribution of the responsibilities. **The documents of the program** must refer to the scope and the objectives of the program, the methodology of the program meaning the strategies used in order to reach the goals, short-term, medium-term and long-term expected results, the system of performance indicators. The documents of the program does not have to lack the description of the specific activities, where it should be mentioned the name, the location, the time framework of each individual activity, the responsible person for each activity and the incentives and penalties related to their proper and improper accomplishment respectively. It also should be mentioned evaluation and monitoring methodology as well as other significant elements according to the particular situation. Any Program's Form should include a chapter of observations and additional notes. The Program's Form must be filled in based on repeated interviews conducted with the involved parties. The program initiators could give information on the context of the program. The organizational details may be delivered by those implementing the program. They could also give details on the documents of the programs.

After collecting all these pieces of information, we already have an overall image of the program and we can prepare the next step: the establishment of the monitoring and evaluation system of the program.

## **2. Setting the evaluation objectives**

During this stage the goals of the evaluation should be accurately presented. Even if not many evaluation questions can be answered because the lack of a monitoring system, several elements can be measured: the effects of the program as perceived by a group of actors involved in the program, the degree of their satisfaction, etc. The presentation of the evaluation perspective is crucial. This could be the perspective of the target group, that of all the actors involved, the

perspective of the financing entity or that of the implementation unit. The type of the evaluation and the evaluation methods should also be specified here in order to set the area of the investigation.

### **3. Conclusions of previous evaluations**

If any evaluation has been done previously, their findings should be presented. It could be useful to present especially strong and weak aspects that have been noticed in the past, as well as any other finding that could set the basis for present or future benchmark. In the category of previous evaluations may fall accreditation processes as well as any other evaluation or self-assessment.

### **4. The construction of an indicator system used for monitoring and evaluation**

The indicator system can be built starting from the concept operationalization procedure. This is an operation specific to social sciences research methodology. The first step would be to turn the concepts into variables. These variables represent the indicators of the future monitoring and evaluation indicator system. Next, several computations could be accomplished in order to obtain indices which can offer a synthetic image of different tendencies. During this stage, the type of indicators should also be established: performance indicators, impact indicators, efficiency indicators, etc. The indicator system that has been obtained can be used in a first instance both as a monitoring indicator system and as an evaluation indicator system. Then, progressively, the evaluation indicator system will develop specific traits.

### **5. The selection and use of research methods in program evaluation**

From the methodological toolkit of social sciences research, program evaluation uses both quantitative and qualitative methods. Therefore, it would be more appropriate to admit that program evaluation actually uses the multimethod approach. According to the type of the program and to the evaluation objectives diverse combinations of qualitative and quantitative methods should be used.

### **6. Data analysis and interpretation**

Data interpretation is done, according to the type of collected data, with the help of statistical (quantitative) analysis methods or with the help of qualitative analysis methods.

### **7. Filling in the indicator system**

During this stage, the system of indicators will be filled in with measurement data obtained data analysis and interpretation stage.

### **8. The development of a plan to fight the unwanted effects**

This stage is specific to the internal institutional management. The evaluation process can bring a significant in-put by rendering explanations conected to functional and not functional aspects that have a positive or a negative impact. Even if the development of a plan to fight the unwanted effects is not generally perceived as being part of an evaluator's work, some useful sugestions can be made at this stage.

### **9. Writing the evaluation report**

The evaluation report represents the synthesis of the evaluation studies. It is about a document based on the evaluation process that can have different destinations. The evaluation report can address to the management of the institutions, to the financing entities that are supposed to have the responsibility of decision-making. The evaluation could suggest through the evaluation report some actions or decisions. In the same time, an evaluation could address to the media or to the general public. According to the target group of the evaluation, the language used and the detail presented should be carefully considered.

### **10. Planning the integration of the evaluation results in the design of future activities and programs**

Every evaluation should end by integrating some of its findings in future actions. The planning of integration of the evaluation results in the design of future activities and programs represents just as the plan for fighting the unwanted effectsan additional but very usefull document of the evaluation.This is usefull mainly for increasing the performance of the future actions.

### **Trends**

As most sources indicate, the tendency in the field is to use more than one evaluation model at a time (Hansen, 2005:448). But they must be carefully chosen and adjusted to country and institutional realities. In the public field, evaluation studies and reports are put to answer more and more questions regarding the process and the results of the projects; it is of increasing interest at the same time the economic effects and social impact effects. Especially in the case of the projects financed from public money the tendancy is to use complex evaluation models that focus on more then one perspective.

In order to be effective, evaluation should be based solid monitoring systems. If the monitoring system is missing, evaluation models should be consequently adjusted. In this respect, the new model, called **PAEM (Public Administration Evaluation Model)**, is a possibility. It generates structured programs and projects out of mere, unstructured public institutions' activities, it helps at building an indicator system that has a double functionality: secures the future existence of a monitoring system and provides a functional evaluation toolkit.

## References

1. Birckmayer, Johanna D. and Carol Hirschon Weiss – „Theory-Based Evaluation in Practice”, 2003, *Evaluation Review* 24(4): 407–31.
2. Chiriac D., C. Humă- “Impactul socio-economic al fenomenelor naturale dezastruoase în România - inundații, alunări de teren, secetă “(62 p.), *Probleme Economice* , 2002, vol. 20-21/ CIDE
3. Dale, Alan, Taylor, Nick, Marcus, Lane, *Social Assessment in Natural Resource Management Institutions*, 2001, CSIRO Publishing.
4. Gârboan, Raluca , “Evaluation Culture and Capacity in Romanian Public Institutions at Regional and Local Level”, in *Transylvanian Review of Administrative Sciences*, 2007, 21 E, pp. 47-60
5. Marginean, Ioan , *Politica Socială. Studii 1990-2004*, Expert, 2004
6. Patton, M. Q., *Qualitative research and evaluation methods (3rd ed)*, 2002, Thousand Oaks, CA: Sage.
7. Rossi, Peter H., Mark W. Lipsey and Howard E. Freeman, *Evaluation: A Systematic Approach*, 2004 Thousand Oaks, CA: SAGE.
8. Schwandt, T. A., *Dictionary of qualitative inquiry*, 2001, Thousand Oaks, CA: Sage.
9. Scott, W. R., *Organizations: Rational, Natural and Open Systems*, 2003, Englewood Cliffs, NJ: Prentice-Hall.
10. Stanculescu, M., *Pre-Accession Impact Studies. EU Cohesion Policy and Romania's Regional Economic and Social Development*, 2003, Institutul European (ed.).
11. Stanculescu, M., *Toward Country - Led Development. A Multi-Partner Evaluation of the Comprehensive Development Framework. Findings from Six Country Case Studies: Bolivia, Burkina, Faso, Ghana, Romania, Uganda, Vietnam*, 2003, IBRD, The World Bank Washington USA.
12. Stufflebeam, Daniel, L. “Fundational Models for 21<sup>st</sup> Century Program Evaluation”, in Daniel L. Stufflebeam, George F Madaus si Thomas Kellaghan (eds.) *Evaluation Models: Viewpoints on Educational and Humain Services Evaluation*, 2000, Boston, MA: Kluver Academic Publishers. pp. 33-84.
13. Stufflebeam, D. L., „The metaevaluation imperative” in *American Journal of Evaluation*, 2001, 22(2), pp.183-209
14. Stufflebeam, D. L., „CIPP Evaluation Model Checklist” in

<http://www.wmich.edu/evalctr/checklists/cippchecklist.pdf>, 2002

15. Vedung , E., “Public policy and program evaluation”, 1997, London, Transaction Publishers.



## **CHAPTER V. THE COUNTERFACTUAL EVALUATION MODEL (CEM)**

*The present chapter aims at bringing closer to the public the logic behind Counterfactual Impact Evaluation (CIE). It starts with a semantic disclosure, continuing with asserting the the main counterfactual theories and their application in the Program Evaluation field and ends with the construction of a Counterfactual Evaluation Model. The chapter intends to present, in an introductory manner, some of the possible and probable uses of the CEM. The main question is: does it make any sense to go counter to the facts in Program Evaluation?*

*Key-concepts: counter factual evaluation, program evaluation, causal models*

## **Introduction**

The new public management theories encourage strategic abilities and functionalities. “Limiting public management to the execution function is a continuation of the old dichotomy (legal/managerial approach) and is not in line with the new public management approach seen in all western countries” (Mora, Țiclău, 2008: 96). This strategic perspective can be achieved in the presence of an accurate image of actions, interventions, programs and projects. We refer mainly to a clear image of their short-, medium- and long-time effects. This accuracy is possible if program evaluation tools are being used systematically. One of the most dynamic evaluation model is the counterfactual one. It is fit to the new public management paradigm from the perspective of its complexity and dynamism. “Even though management and leadership have a common basis and share key characteristics there are significant differences that make a managers and leaders job different”. (Hințea, Mora, Țiclău, 2009: 90) Part of the common basis is the use of program evaluation and of the counterfactual evaluation, particularly.

According to The American Heritage® Dictionary of the English Language, *counterfactual* is an adjective meaning “Running contrary to the facts”. More explicitly, Collins English Dictionary – Complete and Unabridged reveals for the same adjective a meaning related to Philosophy or Logic “expressing what has not happened but could, would, or might under differing conditions”. The most relevant synonyms are: *contrary to fact* and *conditional*. The concept has been successfully imported in the field of Program Evaluation through the Counterfactual Impact Evaluation (CIE). CIE mainly refers to an evaluation methodology that compares the effects of a program or of an intervention to the estimated effects of a scenario where the program and intervention are not present.

### **Counterfactual Theories**

There are several main theories explaining the concept of *counterfactual*. Most of them have their basis in Philosophy.

#### *The first attempts*

The first explicit definition of causation in terms of counterfactuals was formulated as early as 1748 by Hume. He refers to counterfactuals when defining cause and effect relationships:

“We may define a cause to be *an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second.* Or, in other words, *where, if the first object had not been, the second never had existed.*” (1748, Section VII). This early definition is a synthesis of what is meant nowadays by counterfactual, in Program Evaluation as well as in other fields of research and study. But, few empiricists have tried to explain causation via counterfactuals mainly because they have felt mainly uncertainty and subjectivity. A counterfactual statement of the form “If it had been the case that A, it would have been the case that C” is true if and only if there is an auxiliary set *S* of true statements consistent with the antecedent A, such that the members of *S*, when conjoined with A, imply the consequent C. The set *S* generated much controversy. (Goodman 1947.) Most empiricists agreed that *S* would have to include statements of laws of nature, while some thought that it would have to include statements of singular causation. (Menzies, 2009)

### *Rigorous counterfactual analyses*

The late 1960's brings the first rigorous counterfactual analyses. (Lyon 1967) This is a fruitful decade for the research and practice of program evaluation as well. For this timeline it is relevant especially the contribution of J. L. Mackie with his book “*The Cement of the Universe*” (1974). Mackie brings into attention the concept of causation as intrinsically related to the background conditions.

Beginning with the early 1970s, David Lewis elaborates on the counterfactual theory of causation. In 1986 he collects all relevant articles in “*Philosophical Papers: Volume II*” published at Oxford University Press.

The original theory of David Lewis, published in 1973, directly approaches, among other subjects of great interest for the counterfactual impact evaluation the counterfactual and casual dependence, the asymmetry of casual dependence and chancy causation. (Lewis 1973a and 1973b).

### *Comparative similarity between worlds*

*Comparative similarity between worlds* (Lewis 1973a) stands as the central concept in the worlds semantics Lewis uses in explaining the counterfactual causality. According to this

theory, one world A is said to be *closer to actuality* than world B if the first resembles the actual world more than the second does. Consequently, any two worlds can be ordered with respect to their closeness to the actual world, while the actual world is closest to actuality, resembling itself more than any other world resembles it.

### *The causal dependence between events*

The causal dependence between events plays a central role in Lewis's 1973 theory. Schematically expressed, event number 1 (E1) and event number 2 (E2) are two separate possible events; E1 is the cause for E2 if and only if when E1 occurs, E2 occurs as well and if when E1 does not occur, E2 does not occur either.

In his theories, Lewis conceives "a cause as something that makes a difference, and the difference it makes must be a difference from what would have happened without it. Had it been absent, its effects — some of them, at least, and usually all — would have been absent as well." (1973b, p.161)

### **Counterfactual in Program Evaluation. Towards Building a Counterfactual Evaluation Model**

In the field of Program Evaluation, the counterfactual theories and analysis has been adopted in the Impact Assessment area.

Impact assessment refers mainly to (1) the effects of programs and projects on medium and long term and (2) the net effects of programs and projects as distinct from the effects of other factors, variables or events.

Whatever type of impact we may choose to measure, social, economic or environmental, related to a program, we have to assess effects. And effects are naturally related to causes. That is why, counterfactual analysis is fit for impact assessments. In this context the counterfactual analysis becomes a method of evaluation. Its instruments are the diverse scenarios that can be built as "different worlds".

In Program Evaluation in general and in Impact Assessments in particular we may use the images of different worlds as scenarios to compare.

The Counterfactual method of evaluation is infinitely generous in instruments and options from this standpoint. On the one hand we have the real world, scenario number 0 (S0), and on the other hand, we may have an infinite number of imaginary scenarios S1, S2, S3 ...Sn, many of which are possible and some of which are even probable. The great refinement of the counterfactual method is to be able to distinguish first between the impossible and the possible, and then, between the possible and the probable. Once this distinction is completed, the counterfactual method of evaluation can be a valuable information source for the funding entities, for the implementers and for the (potential) beneficiaries of programs and projects. The necessary distinctions are to be made in close relationship to the background and to other similar projects and programs.

What is more, derived from the counterfactual theories, not only a method, but even an evaluation model can be recognized. As we have shown in another article, (Gârboan, 2008: 45), an evaluation model stipulates the question or the set of questions that a specific evaluation seeks to answer. It also involves a certain methodology to set up the criteria for assessment (Hansen, 2005). The literature on programs' evaluation and that on organizational effectiveness offer several typologies of evaluation models. Hansen (2003, 2005) and Scriven (2003) propose some of the most recently appeared and comprehensive typologies. These mainly consists in six different categories of models that are common at some point to other authors as well such as Birkmayer and Weiss (2003). The six categories are: results models, process models, system models, economic models, actor models and program theory models. The counterfactual evaluation model is part of the seventh category of evaluation models: the causation models. These derive from causation theories in philosophy and logic. The counterfactual model relays on the counterfactual causation theories of which we have already mentioned Lewis's. The main set of questions to which an evaluation done in the framework of the counterfactual model is supposed to answer are related to the following: are the results of the program, project or intervention significantly different from the results of the non-intervention? What are the most plausible/probable scenarios in the situation of the non-intervention? Is there any possibility to deduce and approximately measure their results? What are the advantages and the disadvantages of each probable scenario (for intervention and non-intervention)? Which is the most desirable scenario? Which is the worst-case scenario? Where does the actual reality scenario situate on a continuum between worst-case and best-case scenario?

The evaluation criteria are set within the counterfactual model by all the participants in the evaluation process: evaluator and experts from different fields.

There are several methods to approximate the counterfactual and the consequences of every scenario: (i) comparing the effects observed on beneficiaries with those observed on *non*-beneficiaries; or (ii) using the outcome observed for beneficiaries *before* they are exposed to the intervention, (ii) logic modeling methods and bench marking . However, caution must be used in *interpreting* these differences as the “effect” of the intervention.

The building of a CEM starts from finding a feasible way to approximate the effects of counterfactual scenarios. Then, CEM involves the building of counterfactual scenarios and analyzing them. It ends with the writing of the evaluation report.

In the present chapter we will focus on the existing methods of approximating the effects of counterfactual scenarios in line with the classical experiment methodology: comparing the effects of an intervention observed on beneficiaries with those observed on *non*-beneficiaries.

The main difficulty of this method would be the correct selection of the two groups: the beneficiaries and the non-beneficiaries. The two groups should be as similar as possible. As there is a complex variable system, a number of steps should be followed to ensure the comparability:

**The first step:** Make a list of all possible variables relevant for the evaluation. There are going to be two sets of relevant variables: set number 1-socio-demographic variables-that helps in building the comparison group/groups and set number 2-program/intervention comparison variables-characteristics specific to the program or intervention relevant for measuring its results and impacts.

**The second step:** Order the variables in the two sets according to their relation to the investigated program or intervention. A strong relation would recommend the variable for the top of the list, while a weak relation would send the variable to the end of the list.

**The third step:** Make a list of the beneficiaries or of the sample of beneficiaries specifying for each of them the values of the relevant characteristics (variables)for comparison, using the set number one of variables.

**The fourth step:** Identify a group or several groups of non-beneficiaries as similar as possible to the group of beneficiaries. The greater the number of non-beneficiary groups, the more counterfactual scenarios can be determined and the greater the probability of reaching relevant conclusions in the evaluation process.

**The fifth step:** collect the data necessary to compare the values of the second set of variables for the group of beneficiaries and the group/groups of non-beneficiaries. For this step, an important concept should be considered: globalization. During this process, globalization can intervene as an important data source or as a wedge that stimulates change (Loessner, Hințea, and 2005:58). The impact of globalization can be small or large according to the type and specificity of the investigated intervention and of the constructed scenario. The variety of comparable outcomes „can be attributed to characteristics of local institutions and the adaptability and relative entrepreneurial character of their managements” (Loessner, Hințea, 2005:65). In collecting the necessary data an increasing role can be attributed to the narrowing of the digital divide. In an article presenting data from a research that tries to measure the level of the digital divide existing in Romania, Dan Șandor reveals that: digital divide is continually narrowing in terms of access to technology and communication, and also in terms of computer literacy (Șandor, 2006: 154). This means increased access to the necessary data for counterfactual program evaluation as well.

These five steps are the first five steps in the process of building a counterfactual evaluation model. To be complete, the model should also involve the following steps:

**The sixth step:** scenario-building-describe the actual reality scenario and the counterfactual scenarios based on the data collection realized at step number 5.

**The seventh step:** scenario-analysis. The analysis of the scenarios built at step number six. The analysis is based on the two sets of variables. According to the scenarios built, the variable systems can be completed.

**The eighth step:** writing the evaluation report.

**Practical use of CEM:**

The CEM can be used for the evaluation of programs, projects and interventions of socio-economic developments in all stages of implementation.

It can be of great help in assessing the quality of activities, programs and projects. CEM logic could also be applied in the assessment the effects of using other evaluation models, such as *Total Quality Management (TQM)*. „TQM is comprised of a set of principles, tools, and procedures that help accomplish the mission of the organization both from a qualitative and quantitative standpoint. TQM is a managerial philosophy that is accomplished within the framework of a managerial system that promotes a continuous improvement with regard to all the activities within an organization. The process of continuous improvement involves three key dimensions: focus on the client; betterment of processes; and total involvement” (Șandor, 2005: 88). CEM could be used in finding the extent and the nature of TQM application impacts.

Another possible use of CEM is to anticipate the desirable organizational change. “The mission of any organizational change process is to be successful (without successful results change processes are simply a waste of the organizations resources), meaning reaching the goal set by the change process, using resources as efficient as possible and perceiving the whole process as positive as possible by the entire organization” (Baba, Cherecheș, Țiclău, Mora, 2009a). What is more, CEM could also assess the effects of organizational change.

CEM should be a used in the governance process as well. “ Governments have been under increasing pressure to change the way they interact with citizens, open up and increase access to services provided” (Baba, Cherecheș, Țiclău, Mora, 2009b) CEM can be perceived as a driver of change, inspiring governments to find increasingly better scenarios in facing citizens’ requests.

Another possible use of CEM is in the process of designing and creating new public structures such as those necessary for public marketing. As Țiclău, Mora, Țigănaș and Bacali argue, creating the structure in the public field is the condition for every new paradigm to be implemented “because we are talking about public administration, for a successful implementation of public marketing the necessary organizational structures needs to be created. Without a marketing bureau/department on the organizational chart no funding can be



allocated legally, thus even being open and willing to carry out marketing activities public managers have to rely on financial “tricks” in order to fund these activities. (Țiclău, Mora, Țigănaș and Bacali, 2010). The use of counterfactual logic in the design of the new structures refers to the conception of several scenarios of the creation and evolution of the structure, based on the available data and experience.

The main advantage of using this CEM is its comprehensive approach. It helps answering an extremely relevant question for every program: does it make a difference? It contribute to estimating casual effects of programs, projects and interventions, measuring intended and unintended effects, for different actors and in diverse circumstances.

In order to add to the accuracy of the analysis, and to the benefits of the counterfactual method of evaluation, step number 6 can be further developed and enriched with step 6.1: building the best case scenario and the worst-case scenario. This artifice will help creating a continuum an which all the other scenarios can find a place. What is more important is that on this continuum, we can establish the *average treatment effect*, especially because *is the basis for cost effectiveness calculations*. (White, 2009)

### **Limitations and pitfalls**

One of the main limitations is the subjectivity of the model. This is because the different scenarios compared with the actual reality are constructed in a hypothetical manner. Subjectivity can be limited to a certain degree by using reality-based data from different program evaluations or case-studies.

But as Stryczynski mentions, even with these data, collected from reality, we need to work with caution: “We will need our more qualitative, "traditional" evaluation techniques to understand to which interventions these findings can be transferred and what determines the degree of transferability” (Stryczynski, 2009).

Another important limitation of the counterfactual model is the lack of data. Especially in countries without a well-established evaluation culture and capacity such as Romania (Malan, 2004, Curley, Perianu, 2006, Gârboan, Șandor, 2007), the lack of data from other evaluations or from other case-studies related to programs or projects, could be a pitfall in the way of

using counterfactual evaluation model. Data from other countries can be used only with great care, if the situations are comparable from different relevant perspectives.

## **Conclusions**

The CEM offers a multimethod toolkit to perform program evaluation. It involves the qualitative and quantitative paradigm, experimental and non-experimental evaluation designs. A comprehensive and cultural effort is needed for a change to occur at all levels of the public administration. (Mora, Țiclău, 2008: 96) This effort can be made even more fruitful by using the counterfactual evaluation model. Recent evaluation theory and practice has proved that the main counterfactual theories find an extensive application in the Program Evaluation field. (EVALSED:[http://ec.europa.eu/regional\\_policy/sources/docgener/evaluation/evalsed/sourcebooks/method\\_techniques/counterfactual\\_impact\\_evaluation/index\\_en.htm](http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/sourcebooks/method_techniques/counterfactual_impact_evaluation/index_en.htm) (31.01.2012). *It does make sense to go counter to the facts in Program Evaluation.* But extensive attention should be rendered to the limitations and pitfalls of CEM.

## References

- 1.\*\*\*, EVALSED: [http://ec.europa.eu/regional\\_policy/sources/docgener/evaluation/evalsed/sourcebooks/method\\_techniques/counterfactual\\_impact\\_evaluation/index\\_en.htm](http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/sourcebooks/method_techniques/counterfactual_impact_evaluation/index_en.htm) (accessed on February, the 31<sup>st</sup> of January, 2012)
2. Birckmayer, J. D., Hirschon Weiss C., „Theory-Based Evaluation in Practice”, 2003, *Evaluation Review* 24(4): 407–31.
3. Baba C. O., Cherecheș R. M., Țiclău T.-C., Mora C.M., *Organizational diagnosis - a management tool for change. Case study in Satu Mare county hospital*, Social\_Sciences\_Citation, *Transylvanian Review of Administrative Science* , 25E, 2009a, P.32 – 39
4. Baba C. O., Cherecheș R. M., Țiclău T.-C., Mora C.M., *Public Participation in Public Policy Process Case Study in Seven Counties From North-Western Region of Romania*, Social\_Sciences\_Citation, *Transylvanian Review of Administrative Science* , 26E, 2009b, P.5 – 13
5. Collins English Dictionary – Complete and Unabridged © HarperCollins Publishers 1991, 1994, 1998, 2000, 2003
6. Curley, H., Perianu, E., *Assessment of the Evaluation Culture in Romania, 2006*, [http://static.anaf.ro/static/10/Mfp/evaluare/Assessment\\_Romanian\\_Eval\\_CultureEN.pdf](http://static.anaf.ro/static/10/Mfp/evaluare/Assessment_Romanian_Eval_CultureEN.pdf) (accessed on February, the 07<sup>th</sup>, 2012).
7. Gârboan R., Șandor, D., *Evaluation culture and capacity in Romanian public institutions*, 2007, in *Transylvanian Review of Administrative Sciences*, 19E/2007, pp. 71-78.
8. Gârboan, R., *Evaluation culture and capacity in Romanian public institutions at regional and local level*, 2007, in *Transylvanian Review of Administrative Sciences*, 21E/2007, pp. 47-60.
9. Goodman, N., “The Problem of Counterfactual Conditionals”, 1947, *Journal of Philosophy*, 44: 113–128. Reprinted in Goodman, 1983.
10. Hansen, H. F., *Choosing Evaluation Models, A Discussion on Evaluation Design*, 2005, in *Evaluation*, Vol. 11, No. 4, p. 447-462, SAGE Publication

11. Țiclău T.-C., Hințea, C. E., Mora C. M., *Leadership and Management in the Health Care System: Leadership Perception in Cluj County Children's Hospital*, Transylvanian Review of Administrative Sciences, 27E, 2009, P.89 – 104
12. Hume, D., *An Enquiry concerning Human Understanding*, 1748.
13. Lewis, D., *Counterfactuals*. Oxford: Blackwell, 1973.
14. Lewis, D., *Philosophical Papers: Volume II*. Oxford: Oxford University Press, 1986.
15. Loessner, A., Hințea, C., *Impact of Globalization in the Central and Eastern Europe and the Role of Higher Education. Experience of Babeș-Bolyai University, 2005*, Transylvanian Review of Administrative Sciences, 15E/2005, pp. 57-66.
16. Lyon, A., "Causality", 1967, *British Journal for the Philosophy of Science*, 18: 1–20
17. Malan, J., *Benchmarking Evaluation Capacity in the New Member States*, Centre for Strategy and Evaluation Services, 2004.
18. Menzies, P, "Counterfactual Theories of Causation", 2009, *The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2009/entries/causation-counterfactual/>.
19. Mora C. M., Țiclău T.-C., *Managerial approach to public administration*, 2008, Social\_Sciences\_Citation, Transylvanian Review of Administrative Sciences, 24E , 2008, pp. 87 – 97
20. Șandor, D., A Framework for the Selfevaluation of the Functioning of Public Institutions: Incentive Toward an Increased Quality or Merely a Formality?, 2005, Transylvanian Review of Administrative Sciences, 15E/2005, pp. 87-95.
21. Șandor, D., Romania's Digital Divide and the Failures of E-government, Transylvanian Review of Administrative Sciences, 16E/2006, pp. 154-162.
22. Stryczynski K., *Rigorous impact evaluation: the magic bullet for evaluation of Cohesion Policy?*, European Commission, Bruxelles, 2009.
23. The American Heritage® Dictionary of the English Language, Fourth Edition copyright ©2000 by Houghton Mifflin Company. Updated in 2009. Published by Houghton Mifflin Company, 2009.

24. White H., *Theory-Based Impact Evaluation: Principles And Practice*, Working Paper n. 3, International Initiative for Impact Evaluation, New Delhi, 2009.

## **CHAPTER VI. INTRODUCING A MODEL FOR SOCIAL IMPACT ASSESSMENT OF PUBLIC ADMINISTRATION REFORM IN ROMANIA**

*Social Impact Assessment (hereafter, SIA) is a specific type of evaluation extremely useful in public management, a research technique and a public policy instrument successfully used all over the world by those responsible with institutional and organizational management, with the coordination of projects and programs financed from public or private funds. Governments use the SIA in order to notice in time the effects of the interventions they implement or they intend to implement. The aim of a model for Social Impact Assessment of Public Administration Reform is to see in time and mitigate the unwanted effects of public administration reforms on the groups of people, on communities and on society, as well as to encourage the positive elements of the impact. In short, Social Impact Assessment of Public Administration Reform can be used in order to minimize losses and maximize the benefits of the reform interventions upon small or large social groups.*

*Key-concepts: social impact assessment models, public administration reform*

## **The problem addressed**

The main problem addressed by the present chapter is the weak data support regarding the social impact assessment of the public administration reform programs in Romania. This is an issue that may have a crucial long-term influence on the quality of public administration reforms. One of the great problems of public administration reform in Romania is the lack of experience in treating the citizen as customers. This is where social impact assessment research can have a positive impact. Another problem addressed by the study is the data gap concerning the previous, present and future effects of reform programs upon population and communities, and, in this respect, the weak data support for the decision-making process concerning the future reforms. One of the expected results of a model for Social Impact Assessment of Public Administration Reform is to support public policies and decision-making process for central and local governments in order to minimize losses and maximize the benefits of reform programs upon small or large groups of people and communities.

The main difficulty of the problem under discussion consists in the lack of a coherent and systemic monitoring system in Public Administration reform programs and, consequently, the gaps in empirical data collected over time. Most gaps can be covered by a complex data collection process both retrospectively and in real-time. The present project proposes an elaborate indicator system that can be used both as a monitoring system and for the future evaluation studies of the social impact assessment of Public Administration reform programs.

## **Necessity and opportunity**

The main limit of the current approaches in the context of the state of the art in the field is the fact that the social impact assessment model that we design and develop will have a national focus as the impacts of Public Administration reform program are considered in the Romanian social, economic, organizational and cultural framework. But this limit is going to be irrelevant as a strong component of benchmarking and comparative perspective can be used in the design of the social impact assessment methodology as well as in the presentation of the results of the impact evaluations carried out with the help of this model. Becker and Vanclay (2008) and A.M.Esteves, D.Franks and F.Vanclay, (2012), propose some of the most recently appeared and comprehensive typologies regarding impact assessment models. Researchers such as C. Barrow (2001, 2003), F. Vanclay, A.M.Esteves (2011), H. Becker (1997), Rabel J.

Burdge (2003) and many other authors promote the study of SIA theoretically and practically through the many publications in the field, through the IAIA-International Association for Impact Assessment, and inside this, through numerous activities such as teaching, organising workshops, conferences, discussion lists, editing professional publications, through permanently updating the domain web site : [www.iaia.org](http://www.iaia.org). A special feature of recent contributions is the stress placed on the practical applicability of the information proposed. That is why, a good part of the newly-appeared publications focus on definitions, justification, and they come with methodologies that, followed step by step, lead to the practical implementation of SIA. An example example in this way is the book of Christopher Barrow: “Social Impact Assessment: An Introduction”, published in 2004, at Oxford University Press. Endowed with more practical aims than theoretical ones, SIA has at least three generally accepted objectives: to inform about changes in norms, beliefs, perceptions, values and their effects, to anticipate possible impacts of actions both negative and positive, to suggest development alternatives to avoid. In short, it is meant to reduce or mitigate problems and maximize benefits.(Barrow, 2004:3)

Frank Vanclay is situated in the same paradigm. Together with other authors from IAIA, he published *The International Handbook of Impact Assessment (2003, reprinted in 2008)*. “Today, the objective of SIA is to ensure that the developments(or planned interventions) that do occur maximize the benefits and minimize the costs of those developments, especially those costs borne by the community”( Vanclay, 2003:1). He mainly refers to externalities, costs that are not properly taken into account. The decision makers, regulatory authorities and developers fail to consider them partly because they are not easily quantifiable and identifiable (Vanclay, Esteves, 2011).

The evaluation model that we develop in order to assess the social impacts of public administration reform is presently absent from the main flow of publications both nationally and internationally.

### **The concrete objectives of building a model for the Social Impact Assessment of reform programs in public administration**

1)Performing the evaluation of the Public Administration reform programs, from the social impact perspective.



2) The possible use of the social impact assessment model developed in the evaluation of the impact of a Public Administration reform program in the 1990-2012 interval (Retrospective (ex-post) social impact assessment).

3) The implementation of the social impact assessment model developed in the evaluation of the social impact of a Public Administration reform programs that are in the implementation stage (interim social impact assessment).

4) The implementation of the social impact assessment model developed in the evaluation of the social impact of a Public Administration reform program that is going to be implemented in the future (ex-ante social impact assessment).

5) The possibility of elaboration packages of public policy proposals based on social impact assessment results in order to diminish negative social impacts of the Public Administration reforms and to encourage positive social impacts.

The degree of originality and innovation of a Social Impact Assessment of Public Administration Reform is high to extremely high. At present there exist no methodological tool for impact assessment study of reform programs adjusted to the Romanian institutional realities. In the impact assessments performed by international organisation in Romanian institutional environment the use of a methodology insufficiently adjusted to the institutional environment have distorted results. The Romanian public institutions beneficiate so far by a weak methodological toolkit that could help in designing and implementing social impact assessments.

A model for Social Impact Assessment of Public Administration Reform involves the design of extremely useful methodological tools for the impact assessments of Public Administration reforms in the Romanian institutional environment. This will give Romanian responsible with the design and implementation of Public Administration reforms the opportunity of more accurately calculating the impact assessment of programs and projects supported from public money. The new model would offer the public officials and the citizen an accurate image and idea of the potential and actual success of certain investments or investment opportunities from public money. As a consequence, the intuitive factor in the decision making process in public administration reform would significantly.

The evaluation model that we develop in order to assess the social impact of public administration reform is presently absent from the main flow of publications both nationally and internationally.

The elements of originality and innovation brought by the present project is the creation of a methodology that can be used for the social impact assessment at the level of complex reform programs. Some elements of the proposed model have already been included in several studies regarding public administration or public services reform by Mora, C. & Ticiu, T. (2012), Antonie, R. (2012), Hințea, C., (2011) and Țigănaș A., Țiclău, T.C., Mora C. M., Bacali L., (2011). But a consistent social impact assessment model for the public administration reform is still missing from the field literature.

The Social Impact Assessment of Public Administration Reform would significantly influence the scientific field by introducing, developing and applying an social impact assessment model and methodology especially designed for the evaluation of Public Administration reform programs in the context of Romanian socio-economic and cultural environment. This new concept and approach will definitely open new themes and research directions especially towards the social impact assessment of reforms in other domains and towards the use of social impact evaluation findings in policy-making process - which is an approach quite new to the Romanian practice, especially due to the lack of the evaluation culture and capacity.

The potential impact of the model in the scientific environment consists in: the possible use of data collected with the help of this model by other members of scientific community, the multiple possible use of the results of ex-ante, interim and ex-post social impact assessment of large-scale programs such as Public Administration reform in comparative studies by scientists from Romania and from abroad, the use of the social impact assessment model created in order to evaluate new possible or effective impacts of complex programs in other fields, the potential multiplier effect: eventual new impact assessment models will be developed for assessing other complex programs etc. The public awareness of the actual social impacts of the Public Administration reform programs in Romania is another potential impact along with the possibility to design packages of public policy proposals in order to diminish negative impacts of the Public Administration reforms implemented so far and to encourage and accentuate the positive impacts at social, economic and cultural level.

## **Methodology**

Concerning the technical issues, the methodology used, there is a certain agreement among professionals. Social Impact Assessment involves the use of program evaluation methodology and that of sociological research methods, both quantitative (statistical) and qualitative (observation, interview, case-studies, etc)

For Social Impact a varied methodology developed, according to the socio-economic, cultural and organisational context, according to the nature of the intervention, the necessary variables to be measured, the available budget, and also according to the research capacity and a series of other factors involved.

Out of the classical sources (Backer, 1997; Rossi, Freeman, Lypsey, 1999) and of the most recent publications in the field: (Esteves, Franks and Vanclay, 2012), (Vanclay, Esteves, 2011), numerous research designs for impact assessment can be used, according to the intervention assignment, the type of controls used, and the data collection strategies. Therefore, we can use simple analysis before and after intervention, cross-sectional studies for non-uniform programs, panel studies: several repeated measures for non-uniform programs and time-series: many repeated measures. Simple before-and-after studies have a non-random and uniform intervention assignment, targets measured before and after intervention, while the output is measured on exposed targets before and after intervention. Cross-sectional studies for non-uniform programs have a non-random and non-uniform intervention assignment, targets differentially exposed to intervention compared with statistical controls. As data collection strategies, after-intervention output measures and control variables are used Panel studies: several repeated measures for non-uniform programs have a non-random and non-uniform intervention assignment where targets are measured before, during and after intervention. For data collection repeated measures are used, taken of exposure to intervention and of output. The time series case: many repeated measures, the intervention assignment is non-random and uniform, there are large aggregates compared before and after intervention. For data collection, many repeated before and after intervention output measures on large aggregates are applied.

The investigation in the area of Social Impact Assessment of Public Administration Reform have to integrate the newest approaches regarding methods and tools from program evaluation and social sciences research. From the social sciences research methodology, the multimethod paradigm is the most appropriate. Specifically, the methods that can be used are: analysis of documents (the document population consists in: documents and reports related to the Public

Administration reform from Romania and from abroad; the research instrument will be the document analysis grid), the secondary data analysis, in order to analyse the Public Administration reform programs, the interview (with the interview guide as instrument) applied to reform programs responsables (present and former ministeries, secretaries of state in the Ministry of Administration and Interior, public officials at local level and other important actors responsible with the reform of Public Administration. The evaluation types that can be used, are ex-post evaluation of the social impact, interim evaluation of the social impact, and ex-ante evaluation of the social impact. From the program evaluation toolkit, a wide variety of evaluation models, methods and instruments could be used: results models, program-theory model and process models, system models and one of the most recently debated in the scientific literature and in the evaluation practice at the level of the European Union: the counterfactual evaluation model. Along with the social research methods mentioned above, the researchers can also use the social survey with questionnaire and the interview, where necessary, to make sure the existent data gaps do not influence the results of the project. A survey can be conducted for each of the ex-post, the interim and the ex-ante evaluation of social impact. The population targeted for the Social Impact Assessment of Public Administration Reform would be: (1) the personnel involved in the development of the Public Administration reform programs at central and local level and (2) the citizens from across the country. The sample would be a random stratified one. The stratification variables possible to be used are: the administrative region (for the investigations performed both on citizen and on public clerks and public officials responsible with reform) : North-East, South-East, South Muntenia, South-West Oltenia, West, North-West, Center, Bucharest-Ilfov and the institution type (used only for the responsible with the design and implementation of Public Administration reform survey): Central government, Prefectures, City Halls, General Directions of Public Finances, Work and Social Protection Directions, Prefectures, County Councils. Beside these, according to the concrete reform domain, some other stratification variables might be used as well. Analysis of the documents, the secondary data analysis applied on data basis and outputs resulted from other research studies, and the methodology specific to public policy proposals: setting the priorities, assessing the alternative scenarios for action, assessing the anticipated results for every scenario can also be used for the accomplishment of the Social Impact Assessment of Public Administration Reform.

The steps involved by the model for Social Impact Assessment of Public Administration Reform in Romania include:

1. The analysis of the social impact assessment models used in the assessment of complex reform programs in public administration.
2. The analysis of the Public Administration reform programs from 1990 till 2012.
3. The analysis of the public administration reform programs from 2013 and of the intended Public Administration reform programs for the future.
4. The development of a social impact assessment model and methodology fit for the evaluation of Public Administration reform programs.
5. Testing the model developed at step number 4.
6. Methodological Design of the research (sampling, questionnaire and interview guide construction and testing) for the investigation of the Social Impact Assessment of Public Administration Reform
7. The collection and analysis of empirical data regarding the ex-post social impact assessment of the reform program in Public Administration from the 1990-2012 interval.
8. Methodological Design (sampling, questionnaire and interview guide construction and testing) for the investigation of the interim social impact assessment of a public administration reform program that is in the implementation stage.
9. The collection and analysis of empirical data regarding the interim social impact assessment of a reform program in Public Administration that is in the implementation stage.
10. Methodological Design (sampling, questionnaire and interview guide construction and testing) for the investigation of the ex-ante social impact assessment of a public administration reform program that is going to be implemented.
11. The collection and analysis of empirical data regarding the ex-ante social impact assessment of a reform program in Public Administration that is going to be implemented.
12. The development of recommendations and policy proposals based on the results of the analysis of the social impact assessment studies performed

## **Conclusions**

Scientifically, the problem is highly important as the developers of Public Administration reform programs could use the results of a social impact assessment of previous, present and future reform programs in order to make data-based decisions. Technologically, the issue is

highly significant as it develops a new social impact assessment methodology fit for the evaluation of the impact of Public Administration reform programs. From the socio-economic and cultural point of view, the SIA methodology presented is extremely relevant as it brings into focus the socio-economic and cultural impacts of the Public Administration reform programs. What is more, with the help of this methodology, the responsables with the reform in public administration can elaborate public policy proposals in order to diminish the negative social impacts of the Public Administration reform programs and to encourage the positive impacts.

The main potential risk are related to: **the data quality and data gaps** regarding the public administration reform programs; approaches for mitigation: collecting primary data, where possible, and signaling the gaps in data interpretation process where necessary.

## References

1. Antonie, R., Going counter to the facts in program evaluation. Towards a counterfactual evaluation model (CEM), 2012, *Revista de Cercetare si Interventie Sociala*, 36, pp. 224-233.
2. Barrow, C. J., *Social Impact Assessment: An Introduction*, Oxford University Press, USA, 2001.
3. Barrow, C. J., *Social Impact Assessment: An Introduction*. London: Arnold, 2003.
4. Becker, H. A., and Vanclay F., Editors, *The International Handbook of Social Impact Assessment*, United Kingdom: Edward Elgar Publishing, 2003.
5. Becker, H. A., *Social Impact Assessment: Method and Experience in Europe, North America, and Developing World*. UCL Press Limited, London, 1997.
6. Burdge, R. J., Guest Editor, *The Practice of Social Impact Assessment*, 2003, *Impact Assessment and Project Appraisal*, 21(2&3): 84-234 (two issues)
7. Esteves, A.M., Franks, D., Vanclay F., Social impact assessment: the state of the art, 2012, *Impact Assessment and Project Appraisal*, Volume 30, Issue 1, 2012, Special Issue: The state of the art of impact assessment in 2012, pages 35-44
8. Hîncea, C., “Reform and Management in Romania. Strategy and structural change”, 2011, *Revista de cercetare și intervenție socială*, vol. 34, pp. 177-196.
9. Mora, C., Ticlau, T., Transformational leadership in the public sector. A pilot study using MLQ to evaluate leadership style in Cluj county local authorities, 2012, *Revista de Cercetare si Interventie Sociala*, 36, 74-98.
10. Rossi, P., Freeman H., Lypsey M., *Evaluation, A Systematic Approach*, 6th ed., SAGE Publication, 1999.
11. Țigănaș A., Țiclău, T.C., Mora C. M., Bacali L., “Use of Public Sector Marketing and Leadership in Romania’s Local Public Administration”, 2011, *Revista de cercetare și intervenție socială*, vol. 34, pp. 212-233.
12. Vanclay, F., Esteves A., *New Directions in Social Impact Assessment: Conceptual and Methodological Advances*, Edward Edgar Publishing, UK, 2011.

## **CHAPTER VII. CONCEPTS OF RESEARCH METHODS AND STATISTICS USED IN PROGRAM EVALUATION**

*This chapter introduces aspects related to the relation between Evaluation on the one side and Research methods and Statistics on the other side. Because of the interdisciplinary profile of program evaluation as a theoretical and practical field, sometimes the importance of using the appropriate research methods and the adequate statistical methods is regarded as having a secondary importance. Based on our own observations and on some other assessments, we are able to state that the use of research methods and of statistical methods should be at the core of program evaluation.*

*Key-concepts: research methods, statistics, program evaluation, qualitative and cuantitative methods*



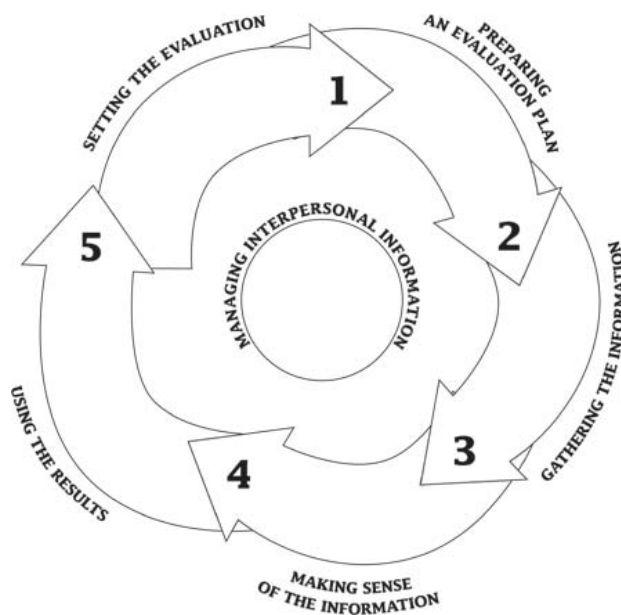
## **Introduction**

According to the European Union Commission, program evaluation can be defined as “a judgment of interventions according to the results, impacts and needs they aim to satisfy” (EU Commission). We also refer to program evaluation as to “the process of assessing the extent to which project, program or policy objectives have been achieved and how economically and efficiently” (Mulreany, 1999). More than that, the UK Treasury defines evaluation as “a critical and detached look at the objectives and how they are being met” (UK Treasury). Even if generically it is named “Program Evaluation”, it applies to policies, programs, projects and other types of interventions. Program evaluation usually involves judgement on basis of criteria based on data collected with the help of research methods and techniques. When numeric data are involved, the judgments relay on statistical arguments.

### **The link between evaluation and research methods**

Evaluation models are usually used to define the objectives of an evaluation, what variables and indicators to study, and the methods needed to collect and interpret the data. At the beginning of each evaluation study a model should be structured in order to carry out a program evaluation systematically and easily. There are numerous models that are being used. Synthetically, the majority use the following steps: (1) identifying the evaluation objectives/initial questions, (2) establishing the indicator system, (3) collecting the data, (4) analyzing the data, and (5) reporting the results.

An interesting five step model used by Community Action Resources for Inuit, Métis and First Nations is presented in Figure 1.



**Figure 1.** Evaluation model. Source: Community Action Resources for Inuit, Métis and First Nations, <http://www.hc-sc.gc.ca>

The diagram presents a dynamic version of the activities that take place during an evaluation. The activities involved are: setting the context of the evaluation, preparing an evaluation plan, gathering the information, making sense of the information and using the results. At the core of evaluation process is the idea of data or information.

Information is supposed to be used in order to improve the program, project or policy evaluated. Evaluation is one of the most important steps in Program Cycle Management, beside Programming, Identification, Formulation and Implementation. Its purpose is to learn through systematic data collection and analysis how to improve programs' and projects' design, how to properly implement interventions, the way we should address accountability concerns, how to make the best decisions concerning the allocation of resources.

As the result of an evaluation several types of decision could be taken: the continuation of the program according to the original design, the continuation of the program with more or less significant changes in the original design, the termination of the program or the changing of future programs or projects according to the lessons learned. Any of these decisions is based on data collected with the help of social research methods and interpreted either qualitatively or statistically, according to the type of the data.

Research methods are involved in every stage of the evaluation cycle as well. We collect and interpret data before the program is implemented (ex-ante evaluation), in order to improve allocation of resources and program design, during the implementation (interim evaluation), in order to analyze whether the program is reaching its objectives and the possibilities to improve the design and the management of the program or project. Data is needed to assess the project or the program after the implementation stage as well (ex-post evaluation) when we can see what the results of the program are, quantitatively and qualitatively.

### **Research methods in program evaluation**

We have already established that research methods are extremely useful in every model and in every stage of the evaluation cycle. Now we have to establish what the most useful research methods are, and when do we use them in the evaluation cycle?

Both qualitative and quantitative methods are used in Program evaluation. The accent is placed upon the complementary use of the two research paradigms and of their subsequent methods. Therefore program evaluation uses the multi-method research model and the preponderance of qualitative or quantitative is decided by several criteria such as: program implementation area, program dimension, number of beneficiaries etc.

Quantitative methods are used especially for the large-scale programs, when there are numerous beneficiaries and when the objectives of the evaluation involve finding out the perspective of the target group. The aim of using quantitative methods is to reach statistically significant results.

Qualitative methods are used mainly in medium and small-scale programs and sometimes in complex programs in order to refine instruments and to find out as many details as possible on different aspects of the program. Qualitative research methods such as individual interview, focus-group, qualitative observation and document analysis are frequently used as well in assessing the programs with a significant social component.

### **Differences between Evaluation and Research**

Even if a strong relation between evaluation and research can easily be perceived, as shown above, several differences must be stressed. As Palumbo had shown (Palumbo, 1987), Carole Weiss illustrated a series of criteria that help distinguishing between the two (Table 1). Some of the most important criteria are the aim, the area of interest, the priorities, the audience, the autonomy, the possibility to generalize the findings etc.

According to these criteria, Evaluation is oriented especially to practical problem- solving, while Research aims mostly at knowledge development. Their target is different even if they may use a common methodological toolkit. The area of interest of evaluation is decided either by the decision maker, by specific actors that might ask for the evaluation, such as the financing entity or the implementing unit.

### **Selecting Appropriate Statistics**

When quantitative analysis is used, several criteria must be considered to ensure selecting the most appropriate data analysis technique in the case of a specific program evaluation. The most frequently used criteria refer to questions, measurement and audience.

Question criteria refer mainly to the evaluation questions and stress whether they are about a casual relationship between a specific cause and effect, or they rely on quantitative variables.

Measurement criteria are concerned with the level of measurement of the variable used, and the level of precision of the measurements etc.

Audience criteria are related to the type of audience of the evaluation. Elements like the expectances of the audience regarding the presentation of data, the precision requested etc. are very important. A target group of the evaluation that is not highly qualified in statistics will expect to see graphs or simple frequency tables, while a statistics qualified target group will definitely expect to see more sophisticated statistical analysis.

### **Selecting a statistical technique to be used in evaluation**

When evaluators collect numerical data to address the evaluation questions, they may have to

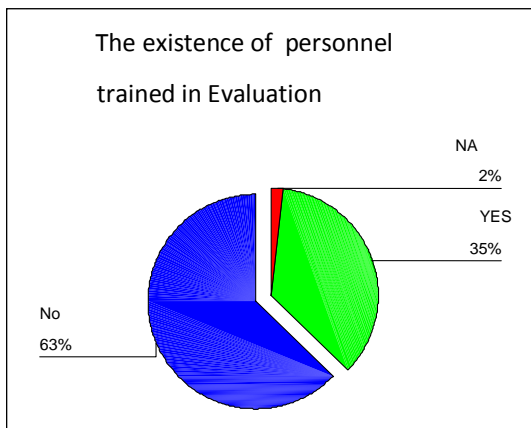
use statistical techniques to analyze the data and to reach reliable conclusions regarding the program. With the help of statistical techniques, evaluators can find information about the relationship between the program, as a cause, and an alleged effect (e.g. by using association). Evaluators may also find out whether and to what extent a group of beneficiaries has been reached by the program (e.g. by using frequency tables). Or, they may find out whether the results of the program are mainly due to one or another characteristics of the program (e.g. by using regression).

Still, the manner in which the variables (characteristics) are measured limits the number of statistics available to evaluators. For instance, in order to analyze a relationship between two variables, when the variables are measured at nominal and ordinal level, evaluators can use association tables (cross tabulation) and as a test for statistical significance, they can use Chi-square test with the computation of lambda or gamma coefficients respectively. But, in the same situation, when the variables are measured on a scale more complex than the ordinal one, on an interval scale, for example, beside the chi-square test evaluators can use the *t*-test.

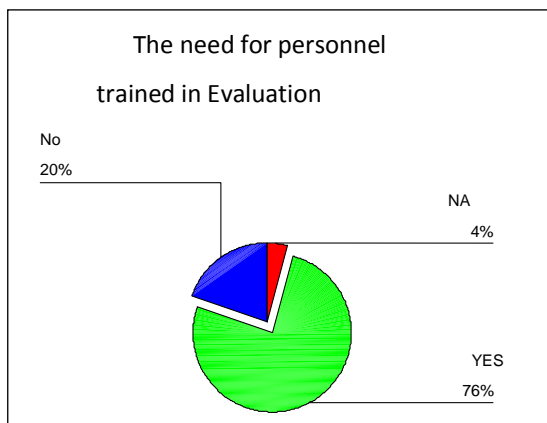
In order to assess a program impact, evaluators may use regression, but only with variables measured on a more complex scale than the ordinal one (e.g. interval). In this situation, the appropriate measure of magnitude of the relationships will be shown by R-square and beta weights.

### **Evaluation, Research methods and Statistics expertise in the Romanian Public Administration**

When talking about the relationship between evaluation and research methods and statistics, we would like to take a look at the way these fields relate in practice. We have measured evaluation capacity in Romanian public institutions at regional and local level (Gârboan, 2007) and, among other aspects we tried to find out real data about the existence of personnel trained in Evaluation, Research methods and Statistics.



**Fig. 2. The existence of personnel trained in Evaluation**

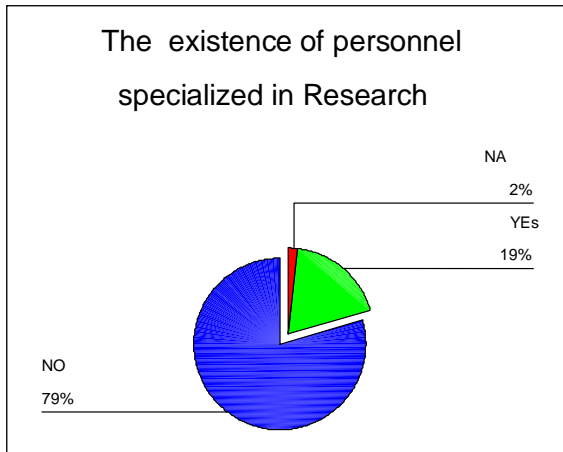


**Fig. 3. The need for personnel trained in Evaluation**

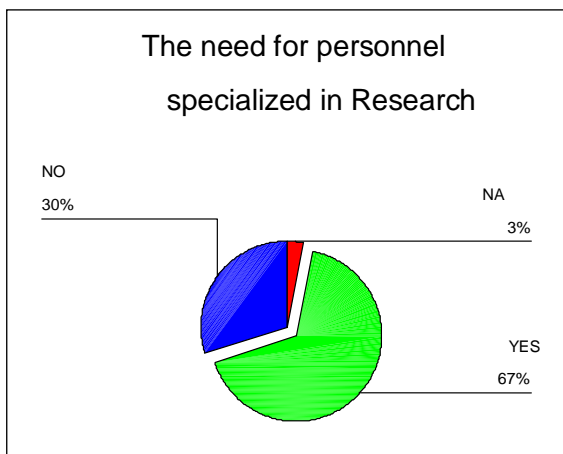
The fact that 63% of the public institutions which were questioned don't have in their structures specialized personnel in evaluation field (Figure 2) and 76% are aware of the existence of this need (Figure 3), shows the tendency to develop the capacity of evaluation in Romanian public institutions.

And because the capacity of evaluation doesn't require only human resources specialized in Evaluation, but also personnel which is specialized in social sciences, Research methods and in Statistics we measured the existence of specialists in these fields in the Romanian public

institutions.

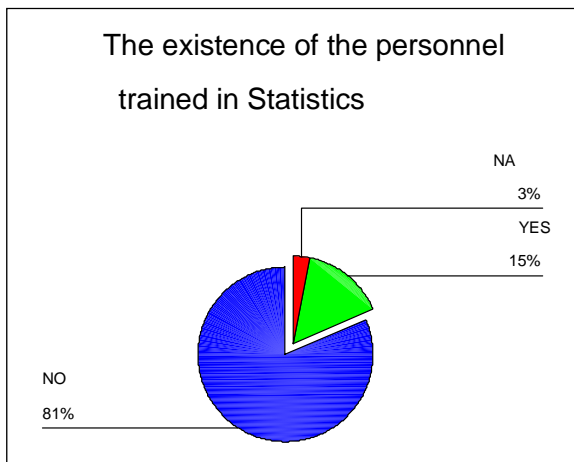


**Fig. 4. The existence in the institutions of the personnel specialized in Research**

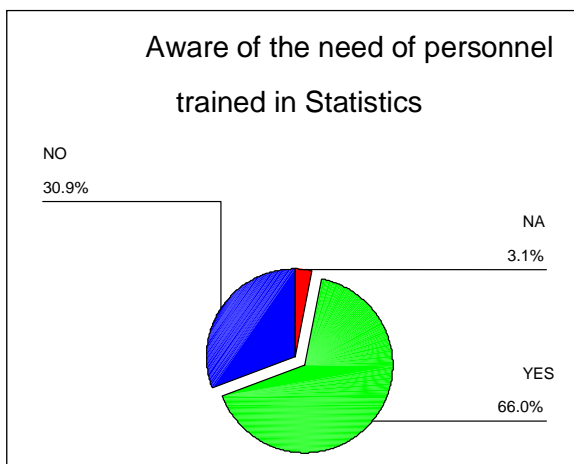


**Fig. 5. The need for personnel specialized in Research**

Regarding Research, 79% declared that they don't have employees trained in Research Methodology (Figure 4), but only 67% are aware of the need for this type of personnel (Figure 5), fact which reveals that Program evaluation field is not known in his essence. Programs which have the role to inform the institutions that there is no possibility to make evaluation unless they have personnel trained in research methods are very welcomed, this aspect being even more important in the public sector where the social impact must be considered a reference point.



**Fig. 6. The existence of the personnel trained in Statistics**



**Fig. 7. Aware of the need for personnel trained in Statistics**

This situation is even more visible in the case of Statistics. About 81% of the institutions realize the lack of trained personnel (Figure 6), but only 66% are aware of the real need for this type of personnel (Figure 7). Or it is known that evaluation of programs cannot be done without statistics, especially when we talk about complex programs.

### Conclusions

Based on our own observations and on some other assessments, we are able to state that the use



of research methods and of statistical methods should be at the core of program evaluation. The existing evaluation capacity cannot be improved without real commitment towards learning from evaluation. And in order to learn from evaluation and to see all its benefit we must fundament our evaluations on arguments that relay on real data collected with the help of research methods and analyzed, when numbers are involved, with the help of statistical methods. Without it our evaluation reports will stick to the “educated guess level or even at the common sense level which is not always quite convincing.

## References

1. Behn, R.D., 'Why measure performance? Different purposes require different measures', 2003, *Public Administration Review*, vol. 63, no. 5.
2. Bickman, L. and Rog, D.J., (eds.), *Handbook of Applied Social Research Methods*, New York: Sage Publications, 1998.
3. Box, R.C., 'An Examination of the Debate over Research in Public Administration', 1992, *Public Administration Review*, vol. 52, no. 1.
4. Caplow, T., *L'Enquête sociologique*, Paris: Armand Colin, 1970.
5. Chelcea, S., *Metodologia cercetării sociologice. Metode cantitative și calitative*, Bucharest: Editura Economică, 2001.
6. Chen, H. and Rossi, P., 'Issues in the theory-driven perspective', 1989, *Evaluation and program planning*, vol. 12, no. 4.
7. Cook, T.D. and Campbell, D.T., *Quasi-experimentation: Design and Analysis Issues for Field Settings*, Chicago: Rand McNally, 1979.
8. Gârboan, R., 'Evaluation Culture and Capacity in Romanian Public Institutions at Regional and Local Level', 2007, *Transylvanian Review of Administrative Sciences*, vol. 21, pp. 47-60.
9. Palumbo, D. (ed.), *The politics of program evaluation*, Newbury Park: Sage Publications, 1987.
10. Patton, M.Q., *How to Use Qualitative Methods in Evaluation*, New York: Sage Publications, 1987.
11. Patton, M.Q., *Qualitative Research and Evaluation Methods*, New York: Sage Publications, 2002.
12. Patton, M.Q., *Utilization-Focused Evaluation*, New York: Sage Publications, 1997.
13. Rotariu, T. and Iluț, P., *Ancheta sociologică și sondajul de opinie: teorie și practică*, Iași: Polirom, 1997.
14. Scriven, M., *Evaluation thesaurus*, 4th Edition, Newbury Park: Sage Publications, 1991.
15. Shadish, W.R. Jr., Cook T.D. and Leviton, L.C., *Fundamentele evaluării programelor: Teorii ale practicii*, Bucharest: FIMAN, 1997.
16. Shadish, W.R. Jr., Cook, T.D. and Leviton, L.C., *Foundation of Program evaluation: Theories and Practice*, Newbury Park: Sage Publications, 1991.
17. Silverman, D., *Interpreting Qualitative Data: Methods for Analyzing Talk, Text and Interaction*, 2nd Edition, New York: Sage Publications, 2001.

18. Stake, R.E., (ed.), *Evaluating the Arts in Education: A Responsive Approach*, Columbus: Merrill, 1975.
19. Stake, R.E., *Standards-Based and Responsive Evaluation*, New York: Sage Publications, 2003.
20. Wholey, J.S., Hatry, H.P. and Newcomer, K.E., (ed.), *Handbook of Practical Program Evaluation*, 1<sup>st</sup> Edition, San Francisco: Jossey-Bass, 1994.
21. Witkin, B.R and Altschuld, J.W., *Planning and Conducting Needs Assessments. A Practical Guide*, New York: Sage Publications, 1995.
22. <http://www.evaled.info>
23. <http://www.evalrom.ro>
24. <http://www.worldbank.org>
25. <http://www.hc-sc.gc.ca>

## **CHAPTER VIII. THE ROLE OF PROGRAM EVALUATION IN THE DECISION-MAKING PROCESS**

*The decision of not organizing a Program Evaluation System at country level government has many negative implications as far as the decision-making process is concerned. The lack of political responsiveness, fiscal discipline and institutional effectiveness are part of the effects. The government does not require a coherent, solid evaluation system and, in exchange, it gets 'Bleak House'- type reports. Program evaluation offers the adequate tools to do evidence-based decision-making on public policy priorities and public resource allocation.*

*Key-concepts: program evaluation, decision-making process*

## **Introduction**

The lack of culture and capacity in program monitoring and evaluation involves the lack of tradition to assess performance in the public sphere. Romania has been characterized by an increased dynamic of legislative change during the past 20 years. But, willingly or not, the regulations concerning Program Evaluation field has been left aside. The Romanian National Evaluation Strategy is an important step forward in acknowledging the importance of Program Evaluation in the decision-making process. But further steps are required in order to turn strategic thinking into reality. The Government should strive to introduce Program Evaluation as a component of decision-making process. Internationally, Results-Based Management and Results- Based Reporting are presently on focus. Successful reform of public administration necessarily involves the evaluation of programs and performance. There has been no regime able to manage its fiscal resources effectively if its programs and its performance were not constantly measured, evaluated and improved.

## **The paradigm**

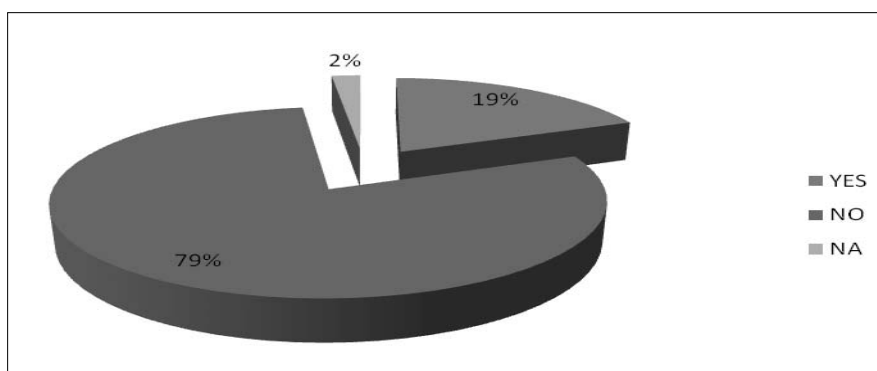
This paper is conceived according to a paradigm proposed by Michael Quinn Patton and illustratively called ‘the paradigm of the practical use of program evaluation’ (Patton, 2002). It focuses on the diverse possibilities of using evaluation results by different types of actors. Without practical use, there is no aim for program evaluation. Evaluation process should begin according to Patton’s paradigm when the design of the intervention (program, project, activity) is being created. Decision-making at all governmental levels is one very important field where evaluation results could be of great use. The condition is that relevant actors in the decision-making process understand the multiple benefits of a solid evaluation system.

## **Data evidence**

Evidence practically means data. And for the decision-making process there are two types of empirical data sources: systematic research and practical experience. There are several entities responsible for the accomplishment of systematic research in public administration: universities, research institutes (private or NGOs), other private

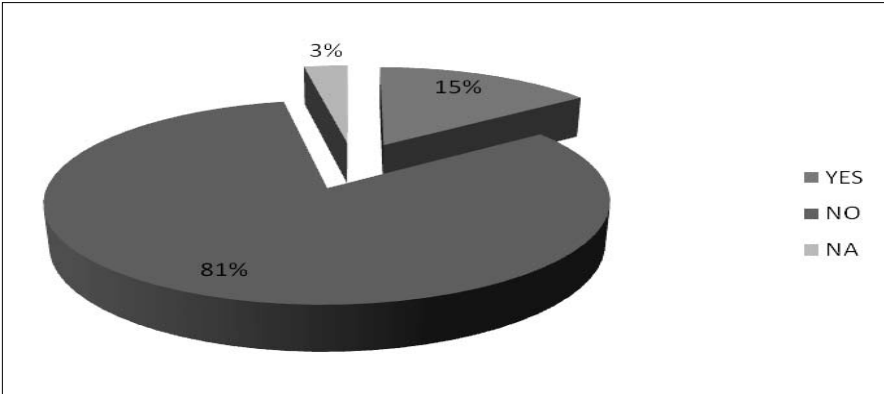
organizations and NGOs involved in the delivery of public services and public administration institutions themselves. The data sources from the practical experience should be delivered by all the actors involved in the wide process of Public Administration, beginning with the government (national, regional and local), the private entities and NGOs involved in public policy and public services. A coherent and comprehensive data system for both research and practical experience is extremely difficult to accomplish. But, a Data Management System for the public sector should be considered and efforts should be made in building it. Presently, the data gaps in this area are tremendous. They are due partially to the lack of strategic orientation and partially to the lack of infrastructure and human resources. The problem with the human resource trained in data gathering and interpretation is severe. In 2007 we conducted an initial research focusing on the evaluation culture and capacity in Romanian Public institutions at regional and local levels (The research method was the questionnaire survey and the instrument, the questionnaire).

In the pilot study the questionnaire was applied to five public institutions and the research itself, the sample was represented by 97 public institutions across the country, mostly mayors of cities, municipalities and county councils). Among the results there were statements referring to the lack of legislation and institutional support. Evaluation capacity requires not only trained human resource in Program Evaluation, but also in Social Science Research Methods and in Statistics. Starting from this premise, we included in the questionnaire a few items that relate to the existence in the Romanian public institutions of personnel trained in Research Methods and Techniques and in Statistics.



**Figure 1: Personnel trained in Research Methods and Techniques**

For the Research Methods and Techniques area, in 2007, 79% of the institutions admitted they do not have trained employees. This gap is even more serious when it comes to Statistics.



**Figure 2:** Personnel trained in Statistics

About 81% of the institutions notify the lack of such specialists. Without statistical expertise, there can be no program evaluation or performance measurement. The importance attached to the use of evidence in decision-making is beyond any discussion or interpretation especially when public funds are supporting the decision. Program Evaluation is a link between empirical data and the decision that is being made. It provides the necessary explanatory tools in order to perceive the different views towards administrative realities.

**Strategic assumptions**

The subsequent question to which this paper tries to provide an answer is what should be done at a strategic level in order to ensure that empirical data evidence of performance is systematically pursued and used as a guiding principle in governmental decision-making? Even if, in time, the benefits of program evaluation are sometimes under-estimated, the ‘new public management models were seen as providing fresh opportunities for evaluation to become a more significant element’ (Halligan, 2003, p. 80). Several questions should find their answers before our initial question can be given an answer. Some of those questions have been often transparent in the literature regarding public management or the governance reform. Some have already been given answers which are to be understood from the perspective they have been written in. For instance, to a question regarding the limits of Program Evaluation application in



the decision-making process, Di Francesco finds an explanation in the political pressure: ‘the onset of fatigue in the application of evaluation to policy advice finally acknowledged the external political constraints facing program evaluation at every level’ (Di Francesco, 2000, pp. 45-46). Some questions are still waiting to be asked and answered. This paper will try to highlight some of these questions and their answer as perceived in 2009-2010, in Romania, during a research in the framework of a governmental research program we participated in. These questions are: ‘Who are the actors at the origins of the decision-making process?’ and ‘Who influences the most the decision-making process?’.

During 2009 and 2010 we have been involved in a research conducted by the members of the Public Administration Department of Babeş-Bolyai University trying to analyze the decision-making process in the Romanian public institutions. The questionnaires were applied to top level civil servants and public officials. Among the results there are several data that could provide answers to the above-mentioned questions.

***Who are the actors at the origins of the decision-making process?***

The answers to this question were given on a scale from 1 to 5, 1 meaning ‘never’ and 5 meaning ‘very frequently’. We will present the results through the perspective of the Mean scores and the Standard Deviation.

**Table 1:** Actors at the origins of the decision-making process

			<b>Standard</b>
INT.1.	Local/County Councillors	3.45	1.1
INT.2.	The Mayor/The President of the County	4.41	0.8
INT.3.	The Vice-Mayors/The Vice-Presidents of	3.32	1.1
INT.4.	Political parties	2.48	1.2
INT.5.	Civil servants and public employees	2.91	1.1
INT.6.	Other public institutions at the central or	2.48	1.
INT.7.	Citizens	2.36	1.2
INT.8.	Mass-media representatives	1.	0.9
INT.9.	NGO representatives	1.91	1.0
INT.1	Civil servants and public employees	1.96	1
INT.11	Unions representatives	1.96	0.9
INT.1	EU and EU institutions	2.32	1.2

The results clearly show that the Mayor and the President of the County Council are the most prominent actors at the origins of the decision-making process in the Romanian public institutions, with a **Mean (X)** of 4.41 on a 1 to 5 scale. The next most important actors at the origins of the decision-making process are Local and County Councillors ( $X=3.45$ ) followed very closely by the Vice-Mayors and the Vice-Presidents of the County Council ( $X=3.32$ ). Therefore, in order to make sure that empirical data evidence of performance is systematically used in governmental decision-making these categories of actors should be made aware of the importance of using the data results of evaluation. Actors generally assumed as evaluation users and consumers are graded, unfortunately, as having among the lowest influence scores in originating the decision-making process: NGO representatives ( $X=1.91$ ), private sector representatives ( $X=1.96$ ) and unions representatives ( $X=1.96$ ). These categories of actors should be encouraged to participate more at the origins of the decision-making process. Generally, the decision-making process should become more participatory, especially in the public field, where public money is involved. The participation of more actors usually adds value to the decision-making process, as more perspectives are focused on the same generally important issue. The importance and benefits of multi-actor decision-making is also highlighted by Pierre and Peter (2005) in order to develop a common set of priorities for society, coherence, steering and accountability.

Interpreting the **Standard Deviation (SD)** scores we notice that the highest agreement between the investigated subjects has been reached in the case of the Mayor and the President of the County Council ( $SD=0.82$ ), while the lowest level of agreement is connected to the political parties being at the origins of the decision-making process. This indicates a high level of controversy towards the issue of the political parties' role. The low score of SD in the case of the Mayor and the President of the County Council stresses the importance of making these actors aware of the importance of using data and evaluation in order to make better decisions.

#### ***Who influences the most the decision-making process?***

It is important to know whether the actors at the origins of the decision-making process are also the most influential. The answer to this question should tell us what other actors should be made aware of the importance of the practical use of evaluation.

**Table 2:** The influence of actors in the decision-making process

			<b>Standard</b>
ILD.1.	Local/County Councillors	7.15	2.
ILD.2.	The Mayor/The President of the County	8.81	1.8
ILD.3.	The Vice-Mayors/The Vice-Presidents of	6.58	2.6
ILD.4.	Political parties	4.83	2.8
ILD.5.	Civil servants and public employees	4.98	2.
ILD.6.	Other public institutions at the central or	4.27	2.4
ILD.7.	Citizens	4.09	2.6
ILD.8.	Mass-media representatives	3.01	2.0
ILD.9.	NGO representatives	2.	2.1
ILD.1	Civil servants and public employees	3.12	2.2
ILD.1	Unions representatives	3.12	2.1
ILD.1	EU and EU institutions	4	2.9

The influence of actors in the decision-making process ranks again the Mayor and the President of the County Council on the first place with a Mean of 8.81 on a 1 to 10 scale. This is supported again by the lowest score of SD (1.89), meaning the highest homogeneity or the highest degree of agreement among the investigated subjects. The next score was obtained by Local and County Councillors ( $X=7.15$ ), while the other categories of actors scored far less (Table 2). This data set confirms the importance of involving the Mayors and the Presidents of the County Council as well as Local and County Councilors into the evaluation process. It also reveals the importance of increasing the relative importance of other categories of actors in the decision-making process.

## **Conclusions**

The Mayors and the Presidents of the County Council are at the origins of the decision-making process in Romanian public institutions and influence it the most. In order to make sure that empirical data evidence is used in governmental decision-making these categories of actors should be made aware of the importance of using evaluation toolkits. Actors generally assumed as evaluation users and consumers are graded, unfortunately, as having a very low influence in originating the decision-making process. The decision-making process should become more participatory and organizing a Program Evaluation System should be one of the priorities of governments at all levels for its improvement. It could mean an added value to political responsiveness,

fiscal discipline and institutional effectiveness if designed and applied properly.

## References:

1. Chelcea, S., *Metodologia cercetării sociologice. Metode cantitative și calitative*, București: Editura Economică, 2001.
2. Di Francesco, M., 'An Evaluation Crucible: Evaluating Policy Advice in Australian Central Agencies, 2000, *Australian Journal of Public Administration*, vol. 59, no. 1, pp. 36-48.
3. Halligan, J., 'Public Sector Reform and Evaluation in Australia and New Zealand', in Wollman, H. (ed.), *Evaluation in Public Sector Reform*, Cheltenham: Elgar, 2003, pp. 80-97.
4. Majone, G., 'Agenda Setting', in Moran, M., Rein, M. and Goodin, R.E. (eds.), *The Oxford Handbook of Public Policy*, New York: Oxford University Press, 2006, pp. 228-250.
5. Palumbo, D. (ed.), *The Politics of Program Evaluation*, Newbury Park: Sage Publications, 1987.
6. Patton, M.Q., *How to Use Qualitative Methods in Evaluation*, New York: Sage Publications, 1987.
7. Patton, M.Q., *Qualitative Research and Evaluation Methods*, New York: Sage Publications, 2002.
8. Patton, M.Q., *Utilization-Focused Evaluation*, New York: Sage Publications, 1997.
9. Pierre, J. and Peters, G., *Governance, Politics and State*, New York: Macmillan, 2000.
10. Pierre, J. and Peters, G., *Governing Complex Societies: Trajectories and Scenarios*, New York: Macmillan, 2005.
11. Rotariu, T. and Iluț, P., *Ancheta sociologică și sondajul de opinie: teorie și practică*, Iași: Polirom, 1997.

## **CHAPTER IX. EVALUATION CULTURE AND CAPACITY IN ROMANIAN PUBLIC INSTITUTIONS AT REGIONAL AND LOCAL LEVEL**

*This paper presents a research study conducted in 97 Romanian public institutions at regional and local level. Two dimensions are analysed: evaluation culture and evaluation capacity, the first one being at the core of the second. Based on our own observations and on some other assessments, the existing evaluation capacity cannot be improved without real commitment towards learning from evaluation. Even if a strong evaluation culture and capacity cannot be perceived, Romanian public institutions at regional and local level show specific features in order to be optimistic about a development of a strong evaluation culture in the future.*

**Key-concepts:** *evaluation culture, evaluation capacity*

## **Introduction**

Program evaluation is connected in Romanian public institutions at regional and local level mainly to the programs and projects financed from pre-accession funds (which have had well-established monitoring and evaluation systems. This led to an initial development of the reporting activities towards monitoring and evaluation. The legal basis for evaluation has been analysed by Local Short Term expert Victor Canghizer<sup>3</sup>. The main conclusion of the research has been that by comparison with EU evaluation regulations, Romanian legislation has specific provisions that explicitly require evaluation only for the co-financing budgets for EU and other donor funding and for research programs. As the other types of activities are concerned requirements to conduct evaluations are either not clear enough or not enforced properly.

Evaluation culture is considered to be "the institutional commitment to learning from evaluation"<sup>4</sup>. In practice, evaluation culture expresses itself through systematically assessing how well programs and projects are working, what changes need to be done in the design and implementation techniques. The evaluation capacity involves, beside a strong evaluation culture, other elements such as: monitoring systems, analytic expertise and good communication networks<sup>5</sup>. Evaluation culture is sometimes seen as a pre-condition to having a well developed evaluation capacity.

In a recent study, „*Assessment of the Evaluation Culture in Romania*”<sup>6</sup> Hilary Curley and Eugen Perianu tried to figure out the evaluation culture in Romania, from a different perspective. It is not necessarily seen as a key element in the constitution of evaluation capacity at country level.

To analyze the culture of evaluation and the capacity of evaluation on Romanian public institutions at regional and local level, we implemented a research.

## **Research objectives**

---

<sup>3</sup> Victor Canghizer (2006)- *Report on the existing legal framework for evaluation*

<sup>4</sup> United States General Accounting Office, Report to Congressional Committees, *Program Evaluation- An Evaluation Culture and Collaborative Partnerships Help Building Agency Capacity*, 2003, p.3

<sup>5</sup> ibidem

<sup>6</sup> Hilary Curley, Eugen Perianu (2006)-*Assessment of the Evaluation Culture in Romania*

The purpose of the research is to find out whether we can talk about a capacity of evaluation and about a culture of evaluation at a regional and local level.

### **Research methodology**

The research was preceded by a pilot study. The research method was the sociologic investigation and the research instrument was the questionnaire. During the pilot study the questionnaire was applied to five public institutions, and during the research the sample was made of 97 public institutions from the whole country, most of them being town halls and districtal councils. The sampling and applying the questionnaire was made in collaboration with The Ministry of Administration and Interns( the actual Ministry of Interns and Administrative Reforms). The selected sample was an exhaustive one, the questionnaire had been sent to all town halls from the country. Beside this, the questionnaire was applied to another smaller town halls, considered significant from the perspective of the inhabitants number.

The institutions investigated on the pilot study are: The prefect's office Cluj, Satu Mare town hall, Babes-Bolyai University and the Teritorial Inspectorate of Labour from Cluj.

The institution's managers were asked to select the most adequate employee of the institution that can provide or collect needed information. Later, the designated person collects information and fills the questionnaire on behalf of the institution.

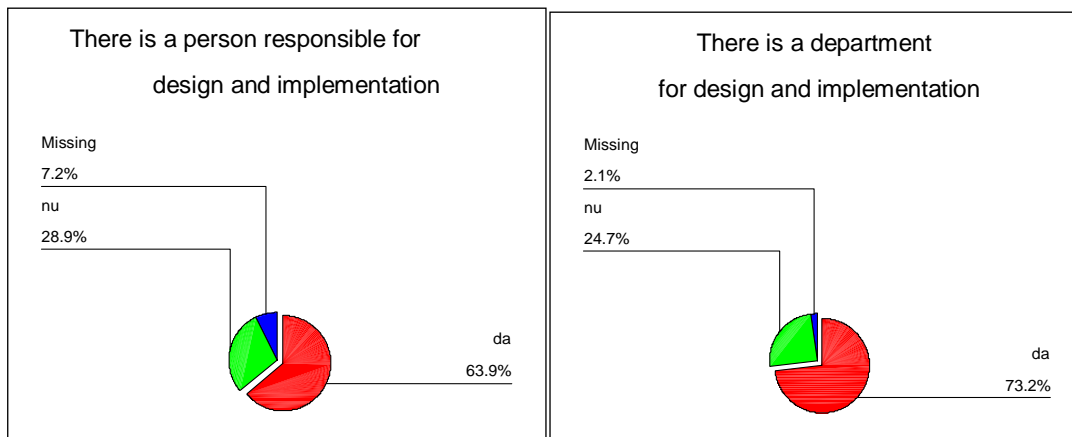
The pilot study was realized on March 3<sup>rd</sup>-March 6<sup>th</sup>, 2007 and the research was on April-May 2007. After the pilot study we decided that the research should focus on town halls and district councils following to make another research with different instruments for decentralized and deconcentrated institutions because of their specificities. The questionnaires were sent electronically and the answer rate was 85% ( 97 answers from 117 questioned institutions – thanks to Ministry of Interns and Administrative Reforms support represented by the State Secretary, Univ.Lector, Mr. Liviu Radu).

### **Results**

Asked if there is a person or a department responsible for design and program implementation, institutions' officials give conclusive answers. 63.9% of the respondents declare that in their institution there is a person responsible with design and project implementation, while 73.2% declare that there is a department responsible for design and



project implementation. This nearly 10 percent difference shows confusion regarding the dissemination of responsibility in project design and implementation. 10% declare they have such a department, but not a person responsible with project and program implementing, or this person can not be identified.



**Fig. 1**

**Fig. 2**

10% means a quite big proportion of public institution that are not yet obviously oriented to a program budgeting system, even if efforts are made for more than 10 years<sup>7</sup>. However, from the received answers to this question, we can tell that 90% of investigated institutions have this tendency and that represents a begin for developing a solid capacity of evaluation and an equally evaluation culture.

Evaluation culture was operationalized by 12 variables:

1. The existence inside the institution of a detailed implementation plans of interventions
2. The permanent effort of finding best modalities of implementing interventions
3. The measurement degree of objectives
4. Collecting informations regularly for each efectuated intervention
5. The use degree of collected information
6. Information accesibility as part of institution
7. The fulfill degree of objectives is known during the intervention

<sup>7</sup> Law 189/1998 regarding the local public finances

8. Permanent monitoring of activities
9. Management expenses during the intervention will be known
10. At the institution level there serious concerns regarding the quality of activities
11. The existence of periodic analysis of registred performances
12. The degree of evaluating results

Each of these variables were measured from 1 to 4 ( 1 means “at a low level” and 4 means “at a high level”). After that we created a new variable by adding scores from the previous 12 variables. The maximum score is 48 points, meaning that the respondent selected the maximum score at all 12 characteristics. Later, we realised a scale on which we can represent exactly the level of culture evaluation in regional and local romanian administrative institutions. In this way, between 0 and 10 points it’s situated the culture “zero stage”; in that case thr respondent gets the minimum score for each of the 12 characteristics and he is indecised regarding some variables. “*Zero stage*” culture suppose not only the absence of specific elements of an evaluation culture, but also the absence of some elements that prefigure or prepare the evaluation culture. Between 11 and 20 points we can talk about a *pre-culture of evaluation*; in that case we can not talk about an evaluation culture, but some characteristics are present or easy developped, and that announces a possibility of developping a culture like that in the future. A *weak culture* supposes scores between 21 and 30 points. We can identify here specific elements of culture evaluation, but not enough consolidated. We will meet fragmented monitorizations of activities and even incipient forms, predominantly formal of evaluation. *The developped culture* (between 31 and 40 points) distinguishes in program budgeting system oriented institutions in which exists monitorising activity systems, evaluation costs can be identificated, there are evaluators and trained personnel on the research methodology specific to social sciences. *The expert culture* (between 41 and 48 points) is characteristic to institutions that not only works on program budgeting system and have well organised informational management systems, but the results of evaluation are used to improove future activities and are shown as examples of training for others.

As far as evaluation culture is concerned, we registered 85 valid answers from 97 investigated institutions.

**Table 1. Evaluation culture. Central tendency**

		Evaluation culture (absolut scors)	Evaluation culture (grouped values)
N	Valid	85	85
	Missing	12	12
Mean		22.33	2.65
Median		22.00	3.00
Mode		26	3
Sum		1898	225

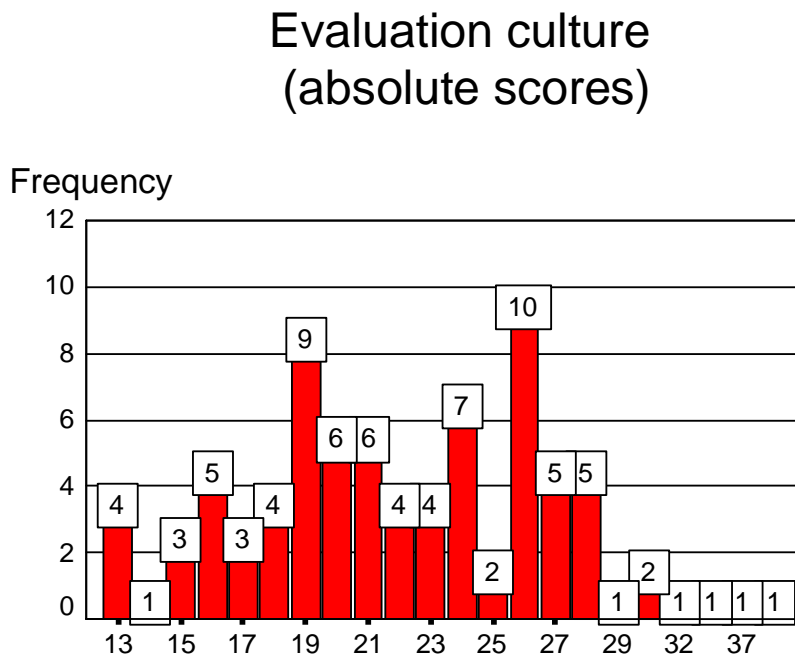
From these institutions, many of them present a weak evaluation culture. 46 institutions accumulated between 21 and 30 points at the multi-criterial analysis that have been realised. It means that Romania's public administration passed the pre-culture stage of evaluation and it already registered some progress in developping an evaluation culture.

**Table 2. Evaluation culture. Distribution of frequencies**

		Evaluation Culture			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2 preculture (11p-20p)	35	36.1	41.2	41.2
	3 weak culture (21p-30p)	46	47.4	54.1	95.3
	4 developed culture (31p-40p)	3	3.1	3.5	98.8
	5 very developed culture (41p-48p)	1	1.0	1.2	100.0
	Total	85	87.6	100.0	
Missing	System	12	12.4		
Total		97	100.0		

However, as it results from the table, it exists a quite big proportion (41.2%) from investigated institutions that are situated on the pre-culture zone of evaluation. These still have to make progresses on data management systems, using previous experiences in improving general activities and even in organising their activities based on programs.

**Fig. 3. Evaluation culture**



The figure represents the repartition of investigated public institutions on the scale used to measure evaluation culture. We notice that modal value is score 26 (10 cases), but very near situates the score 19 (9 cases), that means rather a pre-culture tendency than a developed culture tendency.

The evaluation capacity involves, beside evaluation culture specific characteristics, 6 more variables:

1. the existence of at least an employee trained in evaluation
2. the existence of a person/a department responsible for strategies, policies, programs and projects design and implementation
3. the number of employees trained in research methods/ statistics
4. the degree of difficulty to collect information from other institutions
5. the measure in which in performance assessments there are usually involved specialists from the outside of the institution
6. the measure in which there are constant preoccupations to improve assessment capacity

At his turn, the evaluation capacity can be appreciated on a scale from 1 to 5, based on scores obtained at all 16 characteristics. At 1<sup>st</sup> level situates “zero stage” capacity (0-15 points). Institutions included in this category don’t have evaluation capacity, do not organize their labour based on program budgeting principles, do not have data management systems, monitoring and evaluation systems, do not have specialists trained in socio-human sciences methods and technics, neither evaluation specialised staff, they can not identify costs supposed by an evaluation, benefits resulted after monitoring and evaluating programs and they never took part in evaluating activities. *Very low capacity* of evaluation suppose the presence of the 16 characteristics in a very small measure: 16 to 30 points. *Low capacity* is recorded by scores between 31 and 45 points. Between 46 and 55 points we have a *high capacity* of evaluation and between 56 and 64 points we identify the *expert capacity*.

**Table 3.** Evaluation capacity. Central tendency

		Evaluation capacity absolute values)	Evaluation capacity (grouped absolute values)
N	Valid	82	82
	Missing	15	15
Mean		31.88	2.57
Median		31.00	3.00
Mode		26	3
Suma		2614	211

We can observe that regarding the capacity of evaluation, they are 82 valid answers and the mean of the absolute values is about 31 points (31.88), which represents the middle part of the scale, but with a slow tendency to the lower part of the scale (closer to 0 p than to 64p).

The following table of results shows that 45.1% of the respondents, which are public institutions, have a very low capacity of evaluation, while 52.4% have a low capacity. We have to remark that they are not any cases in which the capacity is at the 0 level.

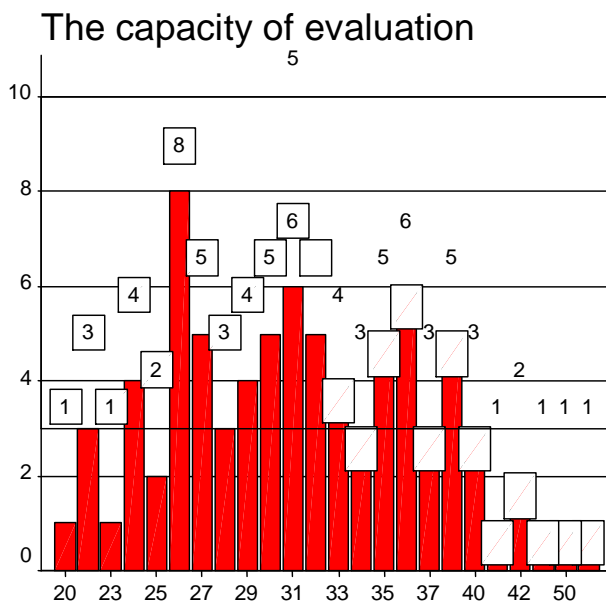
Although, we can meet high capacity values in two cases.

**Table 4.** The capacity of evaluation. Frequencies' distribution.

**capacity of evaluation\_grouped together values (alt\_4.7 + con\_4.13 + ext\_4.14 + cap\_4.15 + cultev)**

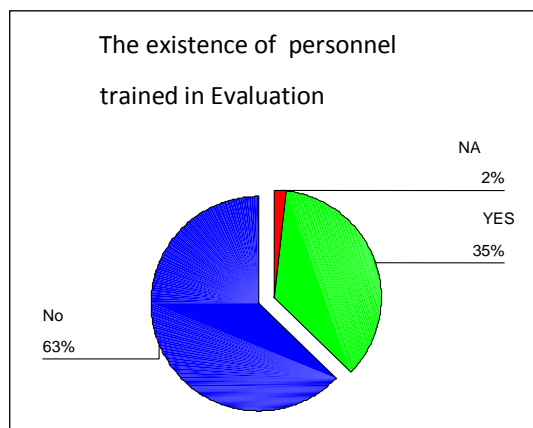
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2 very low capacity (16p-30p)	37	38.1	45.1	45.1
	3 law capacity (31p-45p)	43	44.3	52.4	97.6
	4 high capacity (46p-55p)	2	2.1	2.4	100.0
	Total	82	84.5	100.0	
Missing	System	15	15.5		
Total		97	100.0		

By watching the graphics of the absolute values which are not grouped together, we can clearly see that the mode value is 26 (Table 3), which is similar to the mode value discovered at the culture of evaluation. Still, we can observe that the capacity of evaluation mode value is located more left, under the median, while the mode value of the culture of evaluation is situated above the median. This is a positive aspect, because the culture is an basic ingredient of the capacity of evaluation.

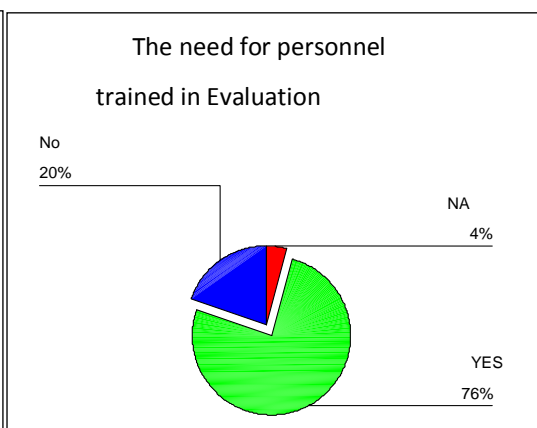


**Fig. 4.** The capacity of evaluation (absolute values)

For having a complete image of the capacity of evaluation in regional and local administration in Romania, we included in the questionnaire some questions regarding their expertise in evaluation, but also regarding the human resources who received training in this domain, or have practical experience in programs' evaluation.



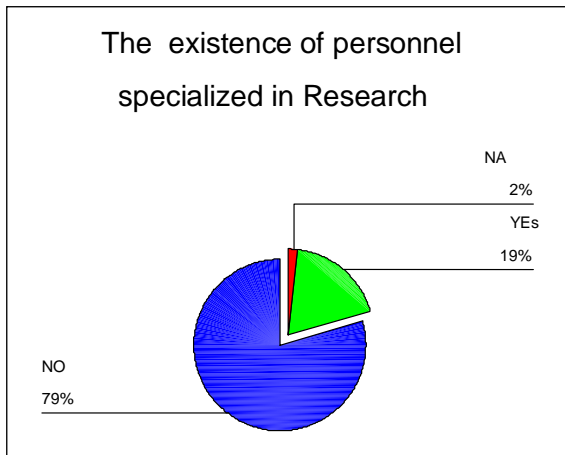
**Fig. 5.**



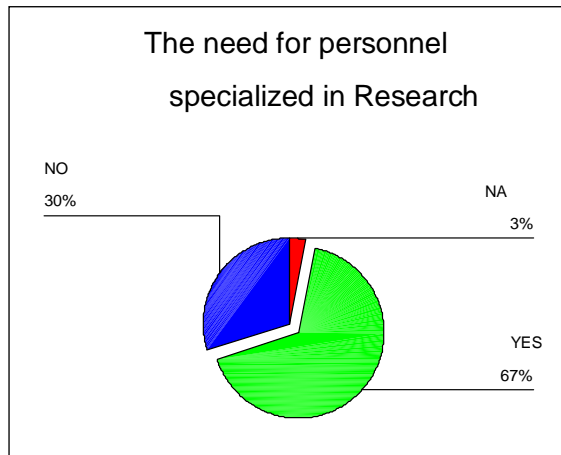
**Fig. 6.**

The fact that 63% of the public institutions which were questioned don't have in their structures specialized personnel in evaluation field and 76% are aware of the existence of this need, indicates that 13%, even if they have specialists in evaluation, are conscious of the fact that the need for evaluation specialized personnel is even bigger. This fact shows the tendency to development of the capacity of evaluation and the unquestionable existence of the evaluation culture in romanian public institutions.

The capacity of evaluation doesn't require only human resources specialized in valuation, but also personnel which is specialized in social sciences' research methods and techniques and in statistics. Starting from this premise, we included in the questionnaire some items with reference at the existence of specialists in these fields in the Romanian public institutions.

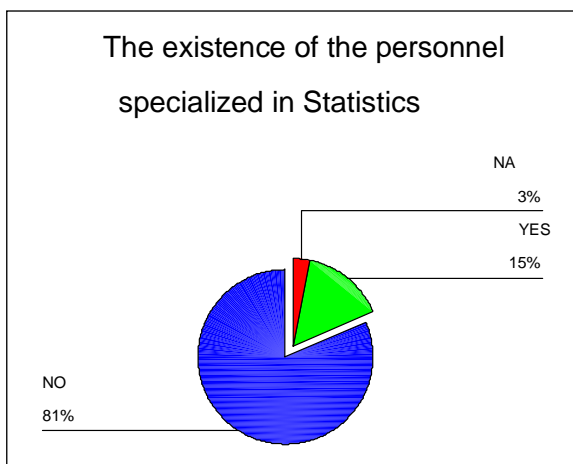


**Fig. 7.**

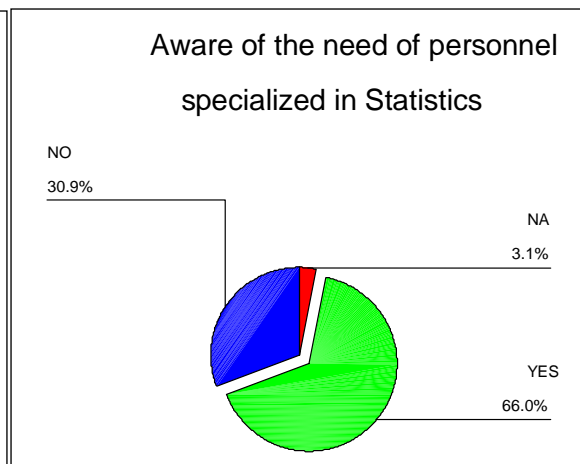


**Fig. 8**

Regarding Research 79% declared that they don't have employees specialized in Research Methodology, but only 67% are aware of the need for this type of personnel, fact which reveals that the programs' evaluation domain is not known in its essence and here from it comes the need for programs which have the role to inform the institutions that there is no way to make evaluation unless the personnel knows very well the research methods, this aspect being even more important in the public sector where the social impact must be considered a reference point.



**Fig. 9.**

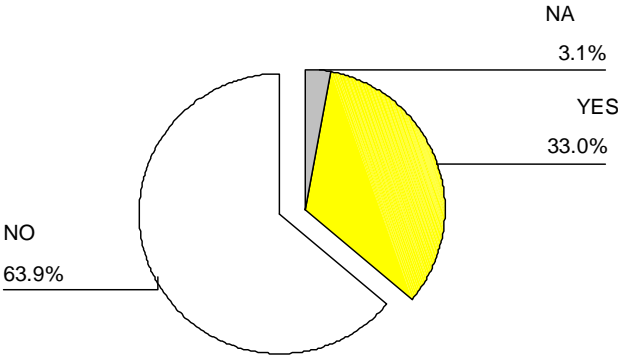


**Fig. 10.**

This situation is even more visible in statistics case. 81% of the institutions realize the lack of specialized personnel, but only 66% are aware of the real need for this type of personnel. Or it



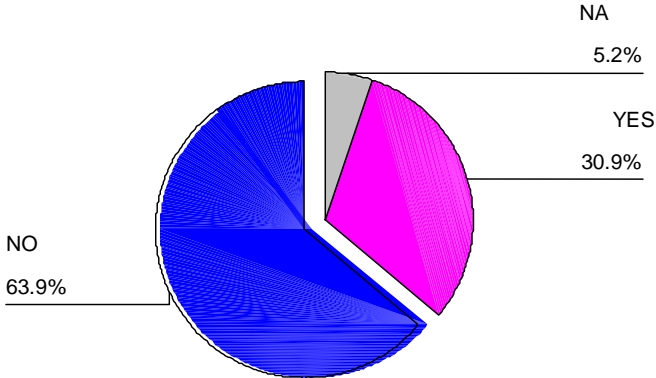
is known that it cannot be done evaluation of programs without statistics, especially when we talk about complex programs.



**Fig. 11.** The capacity to estimate the resources

The evaluation capacity is also given by the estimation capacity of the necessary resources, which, as the present study indicates, is very low, only one third of the respondents being able to estimate the costs of a evaluation.

### The involvement in the programs/projects evaluation



**Fig. 12.** The involvement in the programs/projects valuation

The study's results also reveal that the institutions' involvement in valuations of projects and programs is extremely low (only 30.9%), even though the most of the respondents have participated in such evaluations and have evaluated themselves when it came to projects which required European financing.

The expertise gained through training/practical activities

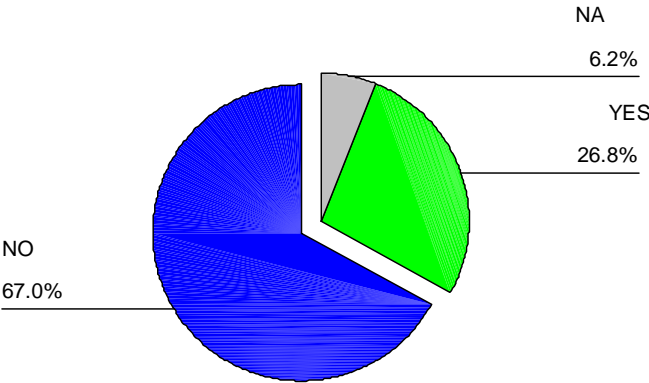


Fig. 13. The expertise gained through training/practical activities

If we choose to talk about the expertise gained through training and/or practical activities, the situation is even more dramatic. Here from it comes the need for training in evaluation and research methods fields, and the emphatic need for research projects which aim at projects and programs evaluation with the implications of the public sector employers.

The existence of some periodic analyses of the performances is dependent on the existence of concrete and measurably objectives, on collecting regularly data for each intervention and on finding the proper utility of this data for discovering how the things are in the institution. This assumption was confirmed. These variables can be integrated in a regression model.

Table 5. The regression model

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.531 <sup>a</sup>	.282	.257	.60

- a. Predictors: (Constant), UTIL\_4.5 the collected data is used for finding out how the things are  
 OB\_4.3 the objectives are very concrete and measurably  
 INFO\_4.4 for each individual intervention there is a regularly data update

This regression model is significant from a statistical point of view (model significance is: 0.000), but it has not a very important explanatory role (R-square is 0.282). Only 28,2% from the dependent variable's variation, *the existence of some periodic analyses of the performances*, is explained by the variation of the independent variables *the existence of concrete and measurable objectives, collecting data regularly for each intervention and finding the proper utility of this data for discovering how the things are in the institution.*

**Table 6. Anova**

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	12.519	3	4.173	11.495	.000 <sup>a</sup>
	Residual	31.948	88	.363		
	Total	44.467	91			

- a. Predictors: (Constant), UTIL\_4.5 the collected information is used  
 In order to realize how things are, OB\_4.3 the objectives of the interventions are very concrete and measurable, INFO\_4.4 for each intervention there are information collected regularly
- b. Dependent Variable: AN\_4.12 there are systematic analysis of the registered performances

The independent variable with the strongest explicative power is *utilizarea informatiilor colectate pentru a vedea cum stau lucrurile* (using collected information to realize how things are) (with standardized beta coefficient of 0.414), and the variable with the weakest explicative power is *colectarea informațiilor în mod regulat pentru fiecare*

*intervenție efectuată* (systematically collected information for each intervention) (with standardized beta coefficient of -0.013)

**Table 7. Coefficients**

<b>Coefficients</b>						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.788	.219		3.601	.001
	OB_4.3 the objectives of the intervention are very concrete and measurable	.210	.114	.198	1.846	.068
	INFO_4.4 for each intervention we have information collected	-1.185E-02	.117	-.013	-.102	.919
	UTIL_4.5 collected information is used to realize	.442	.132	.414	3.352	.001

a. Dependent Variable: AN\_4.12 there are systematic analysis of the registered performances

So, in order to develop a strong evaluation capacity, we need a system for collecting information, and once collected this information must be used in a constructive way.

### **Conclusions**

The results, after analyzing the evaluation capacity shows that the hypothesis confirms. The public institutions from Romania don't beneficiate yet of a developed evaluation capacity, but there certainly exists a promise about this matter.

After the analysis of the evaluation culture in Romanian public institutions at regional and local level it becomes obvious that the measurement and assessment of projects results and impact is something new. The poor expertise in program evaluation in Romania is a cause of a limited understanding of the benefits of the evaluation among public clerks. Those who manage and supervise the implementation of the programs are the first individuals that need

to learn about the need of evaluation and about the facts that the benefits of an evaluation should always outweigh the costs.

The Ministry of Public Administration Reform and Internal Affairs consulted the Finance Ministry and it has appreciated the opportunities that structural funds give for developing a strong evaluation culture, as a final step in the process of implementing projects and programs.

The interest in evaluation, as an academic research, is growing and it distinguishes itself as an academic discipline on different levels. The language that is currently used to describe evaluation is confusing, which leads to perception of the evaluation as being different depending on different decision levels. Some of those problems are caused by the interchangeability of terms like “măsurare”, (“assessment”), “audit” (“audit”) and “evaluare” (“evaluation”), from Romanian language. Also there’s insufficient literature to define the dimensions of evaluation: a control instrument or a management instrument, internal vs. external, different types of evaluation. Evaluation is still seen as a control instrument used to supervise, and detect the errors from a system.

Getting used with the meanings of evaluation and monitoring is continuously developing in Public Administration, especially when it comes about projects management or public politics. There is nothing much to tell about evaluation in other Public Administration sectors.

At the half of the year 2005, the Management Authority<sup>8</sup> of Public Finance Ministry, began developing a national evaluation strategy based on a technical assistance contract financed by PHARE 2003. The National Evaluation Strategy is an important instrument that could gather evaluation’s tendency of development in Romania. It also represents an opportunity of developing evaluation culture which will efficiently sustain the government of the state. As a final purpose, the National Strategy of evaluation tries to be a “national functional evaluation system, where the parts of the system help each other, includes public and private sector and also civil society; to participate in public intervention management and responsibility of politics and public managers”. This Evaluation Strategy was designed in November 2006 and it tries to guide on a single way the standards of evaluation that are used in Romania today. The strategy was build as a stage of building and developing an evaluation culture in

---

<sup>8</sup> Which coordinates the evaluation and monitoring of PHARE and the Operational Programs funded from EU structural instruments

Romania. In order to implement the strategy and improve the evaluation capacity of public administration in Romania ( politics and decision elements), of potential representatives in evaluation, supervisors and local potential evaluation companies, academic environment and participating organizations. Considering the importance of evaluation capacity, in public administration, the activities were concentrated on the structures that manage the EU funds.

The interest for evaluation is growing but the precise request for evaluation is just at the beginning. There aren't any examples presenting evaluation (ex-ante, interim or ex-post) as being built in lack of a list of programs financed by European funds. The evaluation request depends on the existence of a legislation which provides a controlled evaluation and general acknowledged necessities and utility of implementation politics, strategic management and defined budget course.

Nowadays, the problem is that the lack of law blocks the development of evaluation system. Evaluation was used for the first time to European funds or other programs financed by other countries. Today understanding evaluation is situated at a lower level compared with public national interventions. So there is no evaluation institution in public administration except the structures that manage EU funds.

It seems to be a malfunction between supervision and evaluation functions. Those who define the monitoring systems should be trained in evaluation because managing information means building the next evaluation.

## REFERENCES

1. Chen, H.T., *Theory-driven evaluation*. Thousand Oaks, CA: Sage Publications, 1990.
2. Chelimski, E., Shadish, W.R. (Eds.), *Evaluation for the 21<sup>st</sup> Century*, SAGE Publications Inc.,1997.
3. ECOTEC, *Thematic Report on Support to PAJC in Bulgaria and Romania*, 2006.
4. Fitzpatrick,J.L., Sanders, J.R., Worthen, B.R., *Program Evaluation*. (3<sup>rd</sup> Edition) Boston: Pearson, 2004.
5. Hughes, J., *Project Manager's Guide to Evaluation*, <http://www.evaluate-europe.net/handbook/handbookpdf>.
6. Kubish A, Schorr L. et al, *New approaches to evaluating community initiatives*, Washington DC: The Aspen Institute, 1995.
7. Kusek, J. Z., Rist, R. C., *Ten Steps to a Results Based Monitoring and Evaluation System*, World Bank Publications, 2004.
8. MWH, *Ex-post evaluation of the PHARE programme in Romania*, 2006.
9. Patton, Michael Quinn, *Utilization-Focused Evaluation: The New Century Text*, 3<sup>rd</sup> ed., Beverly Hills: Sage Publications, 1997.
10. Shadish, W. R. Jr., Cook, T. D., Leviton, L. C., *Fundamentele evaluării programelor: Teorii ale practicii*, FIMAN, 1999.
11. Stake, R., *Standards-Based and Responsive Evaluation*,2003.
12. Tavistock Institute et al., *The Evaluation of Socio-Economic Development: The GUIDE*, 2003.
13. United Nations, *World Development Report*, New York: Oxford University Press,1991.
14. United States General Accounting Office, *Program Evaluation- An Evaluation Culture and Collaborative Partnerships Help Building Agency Capacity*, 2003.
15. Weiss, C. H., “Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families.” In James P. Connell, Anne C. Kubish, Lisbeth B. Schorr and Carol H. Weiss (eds.) *New Approaches to Evaluating Community Initiatives. Vol.1: Concepts, Methods, and Contexts*. Washington, D.C.: The Aspen Institute, 1995.

## **CHAPTER X. LEVERS SUPPORTTING PROGRAM EVALUATION CULTURE AND CAPACITY IN ROMANIAN PUBLIC ADMINISTRATION: LEADERSHIP**

*Program Evaluation culture and capacity is at the very beginning of its development in Romania. In this chapter we highlight one of the fundamental, but not always obvious, connections that support a sustainable Evaluation culture and capacity building and development: the link between leadership and program evaluation. If properly used, program evaluation results can be a strong instrument in leadership, just as leadership can fundamentally encourage the development of evaluation culture and capacity. More precisely, we identify the ways in which different power sources can help leaders in developing evaluation culture and capacity.*

**Key-words:** *evaluation culture, evaluation capacity, leadership*



## **Introduction**

Usually, leadership in public administration is referred to as “capacity of persons holding senior positions in public organizations to influence the subordinates. This is assumed to be greater if based on one or more power sources.” (Hințea et al, 2010). Program evaluation refers to the process of assessing whether objectives of specific programs, project or interventions are reached, and whether the anticipated results have been achieved. Evaluation is highly explicative. It tries to elucidate why certain objectives have not been reached, why, different indicators have other values at the end of the interventions than those anticipated. A mutual relationship of support can be established between leadership and program evaluation through the evaluation culture and capacity. If properly used, program evaluation results can be a strong instrument in leadership just as leadership can fundamentally encourage the development of evaluation culture and capacity.

## **Evaluation culture and capacity**

Evaluation culture is considered to be "one of the institutional obligations to learn from evaluation." (USGAO, 2003: 3) The same source defines evaluation capacity as involving, beyond a strong culture of evaluation, elements such as monitoring systems, analytical expertise and communication networks (USGAO, 2003). Evaluation culture is sometimes seen as a prerequisite for the development of evaluation capacity. In practice, evaluation culture expresses itself through systematically assessing how well programs and projects are working, what changes need to be done in the design and implementation techniques. The relationship between the evaluation capacity and the evaluation culture one of subordination.

Evaluation capacity has been defined from different perspectives, starting from diverse sets of variables. In the study „*Assessment of the Evaluation Culture in Romania*”, Hilary Curley and Eugen Perianu perceive *evaluation culture* in Romania, from a different perspective. It is not necessarily seen as a key element in the constitution of evaluation capacity at country level. The variables they consider in measuring the evaluation culture are: the frequency of the commissioning of evaluations, the existence of Romanian evaluation experts, the dissemination of the evaluation outside the management group, the exposure of the evaluation findings, the extent in which evaluations make a significant impact on the accountability

debate or through “lessons learned” improvements in planning; the existence of institutionalizing factors (e.g. regulation) and strong non-formal drivers (e.g. civil action groups), the development of outcome/result based monitoring. (Curley, Perianu 2006). Another study benchmarked in 2004 the evaluation capacity in the EU new member states as compared to the 15 older member states. Jack Malan, a researcher in the Centre for Strategy and Evaluation Services reached a series of conclusions relevant for the Romanian public institutions, and defined *evaluation capacity* starting from the following variables: the number of evaluators with necessary skills, the existence of support from public authorities, the existence of an evaluation culture, the existence of information and support on evaluation, such as guidelines, methodologies and best practice examples, the presence of commercial incentives to improve evaluation capacity and expertise, the ability to ensure that evaluation results feed into policy making, the presence of baseline data and of the defined targets and performance indicators and the quality of evaluation reports. (Malan, 2004).

From the above – mentioned sets of variables, it is obvious that there is an overlapping and a subordination relationship between the evaluation culture and the evaluation capacity.

We have tried to figure out the levers, methods and instruments that could be used in order to stimulate the development of the evaluation culture and capacity. Leadership is one of these levers.

### **The link between Leadership and Evaluation**

Regarding the connection between leadership and evaluation our perception is based on the following principle:

*Power sources can help leaders in developing evaluation culture and capacity and evaluation culture and capacity can help leaders in fulfilling their mission.*

The underlying rationale for this principle is that leaders could use power-sources levers to help develop evaluation culture and capacity. The benefits leaders could have from the results of evaluation are multiple: knowledge-based thinking and decision-making is one of the most important.

Following the above-mentioned study, we came to many unanswered questions. One of those questions is: what levers should we use to develop evaluation culture and capacities in Romanian Public Administration? What methods would be most effective? The benefits of evaluation have been asserted so often (Patton,1997), (Shadish, Cook, Leviton, 1999), (Stake, 2003), (Weiss, 1995). Why, in the case of Romania, its development is delayed.

One of the answers at hand is the social, political, cultural, organizational history. We refer mainly to the lack of data collection and transparency traditions. We also have to consider the fear of punishment that was specific to every evaluation (perceived as control) process in the communist regime. But, we further investigate whether some controllable levers, methods and instruments could be used in order to prevent a further delay in the development of evaluation culture and capacity in Romania and in the countries with similar traditions.

In the present chapter, we propose a debate on whether leadership could be one of the levers that could help the development of the evaluation culture and capacity. Why leadership? Because it is one of the most flexible and most influential organizational realities.

Leadership capacity is greater if based on one or more power sources. These sources of power derive, on the one hand, from specific job title within the organization structure and, on the other hand, from the personal skills of the leader.

Let us present the way different sources of power could lead to the development of evaluation culture and capacity. In doing this, we'll use the definitions and classifications given by John French and Bertram Raven (1959).

Reward is one of the fundamental sources of power. It is based on the belief that a person has the ability to reward another person in exchange for loyalty and obedience. Speaking of organization, the reward may be materialized in the form of salary increases, promotion, or another form of recognition. Leaders can use reward to stimulate the subordinates. Can leaders contribute to the development of the evaluation culture and capacity? The answer is yes, they can. In the same time, they could stimulate the capacity of the subordinates to improve their performances and be more competitive for the rewards.

Leaders could encourage subordinates to:

- use detailed plans for operational activities,
- set concrete and measurable objectives of their actions and interventions,

- systematically collect data for every intervention they are involved in,
- use the information they collect to figure out how things work,
- make the information accessible for all those interested in,
- make transparent the degree of accomplishment of the objectives all across the intervention, or in charge of,
- permanently monitor activities specific to an intervention they are involved in or responsible of,
- make transparent the expenses of an intervention they are in charge of every moment of the implementation process. This way, leaders can encourage the development of the evaluation culture and capacity.

What would be the gain for such an effort? Leaders could use the information generated by the monitoring and evaluation system in order to reward their subordinates. Consequently, they can improve their leadership capacity and power. At their turn, subordinates could be more successful in finding the best ways to perform their activities and to compete for rewards.

Coercive power is based on the belief that a person has the ability to punish the other person to convince it to follow an order. This can be used to stimulate subordinates to do the actions mentioned above by limiting their privileges when they do not comply.

Expertise and information are two of the most important sources of power for leadership, especially in a complex and technically evolved environment as today's society. Program evaluation can deliver both expertise and information. Evaluation reports can reveal the necessity of continuing, interrupting or changing the implementation of an intervention. These are relevant information pieces for the leadership process. What is more, evaluation can filter pieces of valuable information from those unnecessary.

The leaders' authority is subjective, psychological and moral in nature as opposed to the forms of influence based on material resources or physical coercion. In an organizational framework, the development of an evaluation culture and capacity can both support and be supported by leadership.

People are willing to follow certain rules and to obey those in leadership. Leaders can use the information rendered by program evaluation system in order to increase their power. And they can use the power in order to support the development of the evaluation culture and capacity.

In Romanian public administration, the evaluation capacity is at the very beginning of its development. Leadership could be a fundamental lever in this development. Just as in Max Weber's typology time, precedent and tradition legitimize the leaders in the eyes of his subjects, program evaluation findings can legitimize the leaders not only in the eyes of the clerks they lead, but in the eyes of the citizen as well.

Personal qualities of the leaders (according to Max Weber, part of the charismatic authority) definitely have greater impact if supported by data and facts extracted from evaluation reports. When we refer to the third type of authority postulated by Weber as specific to the modern civilizations, namely rational legal authority it can easily be understood that program evaluation can bring valuable information and expertise to the leaders to increase their influence and power. Legal-rational authority comes, from a position based on internal rules of the organization. But when a person has a job that gives authority, but this position is overshadowed by factors such as lack of professional competence, we may speak only of official authority, and not a real one. In this case, program evaluation can be a good instrument to turn the official authority into a genuine one.

## **Conclusions**

Program evaluation is not the ultimate solution for any problem in public or private organisation, but its results can help influencing process in organizations. This is often associated with the use of forms of coercion (threats, sanctions).

Program evaluation and leadership does not imply the use of force, but the ability to make people really understand their mission and want to reach for their goals. There can be established a mutual relationship of support between leadership and program evaluation. If properly used, program evaluation results can be a strong instrument in leadership just as leadership can fundamentally encourage the development of evaluation culture and capacity.

Program evaluation can deliver both expertise and information which are two of the most important sources of power for leadership, especially in a complex and technically evolved environment as today's society.

## References

1. Curley, H. Perianu, E., *Assessment of the Evaluation Culture in Romania*, 2006, [http://static.anaf.ro/static/10/Mfp/evaluare/Assessment\\_Romanian\\_Eval\\_CultureEN.pdf](http://static.anaf.ro/static/10/Mfp/evaluare/Assessment_Romanian_Eval_CultureEN.pdf) (accessed on February, the 07<sup>th</sup>, 2012).
2. ECOTEC, *Thematic Report on Support to PAJC in Bulgaria and Romania*, 2006, [http://ec.europa.eu/enlargement/pdf/financial\\_assistance/phare/evaluation/zz\\_pajc\\_0536\\_final\\_version\\_revised\\_e4\\_en.pdf](http://ec.europa.eu/enlargement/pdf/financial_assistance/phare/evaluation/zz_pajc_0536_final_version_revised_e4_en.pdf) (accessed on February, the 07<sup>th</sup>, 2012)
3. French, J.R.P. and Raven B., "Bases of Social Power." in Ed. Dorwin Cartwright *Studies in Social Power..* University of Michigan, Ann Arbor, 1959.
4. Gârboan, R., *Evaluation culture and capacity in Romanian public institutions at regional and local level*, 2007, in *Transylvanian Review of Administrative Sciences*, 21E/2007, P. 47-60.
5. Gârboan R., Șandor, D., *Evaluation culture and capacity in Romanian public institutions*, 2007, in *Transylvanian Review of Administrative Sciences*, 19E/2007, pp. 71-78.
6. Hințea, C., Hințea, C., Țiclău, T., Jenei, G., Vasquez-Burguete, J., Gutierrez Rodriguez P., *Management si leadership in sectorul public in Managementul administratiei publice*, Accent, 2010.
7. Malan, J., *Benchmarking Evaluation Capacity in the New Member States*, Centre for Strategy and Evaluation Services, 2004.
8. Patton Q. M. (1997): *Utilization-Focused Evaluation: The New Century Text*, 3rd ed., Beverly Hills: Sage Publications.
9. Shadish W. R. Jr., Cook T. D., Leviton L. C., *Fundamentele evaluării programelor: Teorii ale practicii*, FIMAN, 1999.
10. Stake, R., *Standards-Based and Responsive Evaluation*, in Tavistock Institute et al, *The Evaluation of Socio-Economic Development: The GUIDE*, 2003.
11. United States General Accounting Office, *Program Evaluation - An Evaluation Culture and Collaborative Partnerships Help Building Agency Capacity*, 2003.
12. Weiss, C. H., *Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families*, 1995, in James P. Connell, Anne C. Kubish, Lisbeth B. Schorr and Carol H. Weiss

(eds.) *New Approaches to Evaluating Community Initiatives. Vol.1: Concepts, Methods, and Contexts.* Washington, D.C.: The Aspen Institute 1995.