

# CONCEPTS OF RESEARCH METHODS AND STATISTICS USED IN PROGRAM EVALUATION

Raluca (ANTONIE) GÂRBOAN

**Raluca (ANTONIE) GÂRBOAN**

Assistant Professor, Public Administration Department,  
Faculty of Political, Administrative and Communications  
Sciences, Babes-Bolyai University, Cluj-Napoca, Romania  
Tel.: 0040-264-431361  
Email: raluagarboan@apubb.ro

## Abstract

This paper 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.

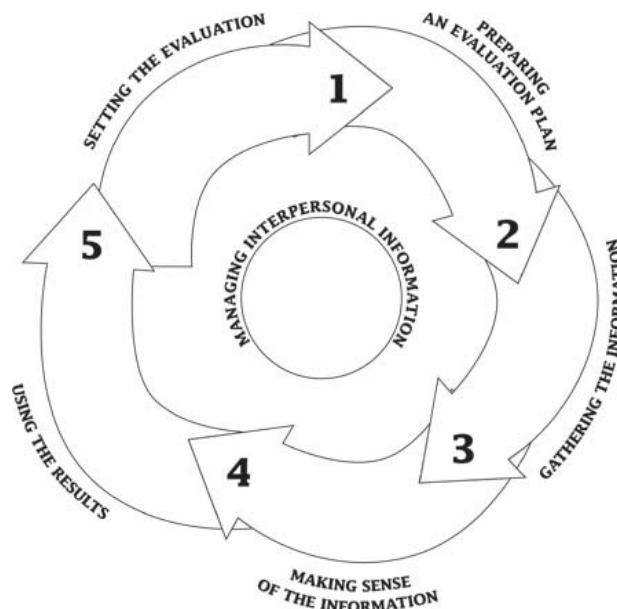
## 1. Introduction

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.

**Table 1.** Differences between evaluation and research

Criteria	Evaluation	Research
Aim	Practical solving of problems	Knowledge development
Area of interest	Established by the decision maker	Established by the researcher
Assessment	Comparison between what it is and what it should be	Usually focused on what it is
Priorities	The program, not the Evaluation itself	Research methodology, not the object of the research
Possible conflicts	Between the evaluators and the team implementing the program	Eventually between the researcher and the sponsor
Publishing	Generally no, but the findings are transmitted to the decision-makers	Crucial
Motivation	Improvement of the situation	Theory development and increased understanding
Audience (small in both cases)	Decision makers	Other researchers
Autonomy	Limited	Relatively high
Possibility to generalize the findings	Oriented to specific situations	Interest in generalization in time, space and different situations
Assessment criteria	Is it relevant for the decision makers? Are the results credible? Does it have any influence upon the program?	Did we test our research hypothesis? Do we have internal and external validity? Does the research open the way to other research problem?

Source: Weiss, C., 'Where Politics and Evaluation Research Meet', in D. Palumbo (ed.), *The politics of program evaluation*, Newbury Park: Sage Publications, 1987, p.167.

## 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).

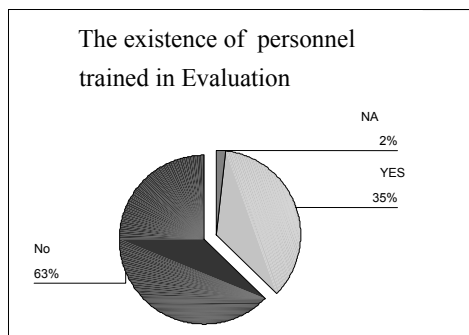
Still, the manner in which the variables (characteristics) are measured limits the number of statistics available to evaluators. For instance, in order to analyze a relationship between two variables, when the variables are measured at nominal and ordinal level, evaluators can use association tables (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 then 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 then 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.

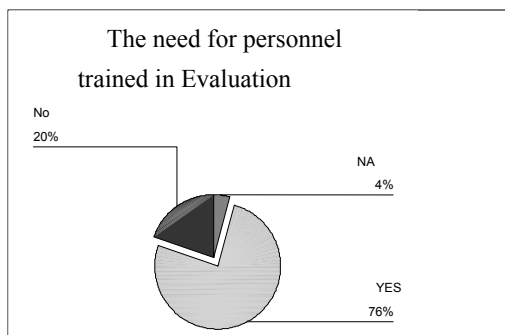
## **2. Evaluation, Research methods and Statistics expertise in the Romanian Public Administration**

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. What is more, we tried to see the perceived need for the personnel trained in the three fields.



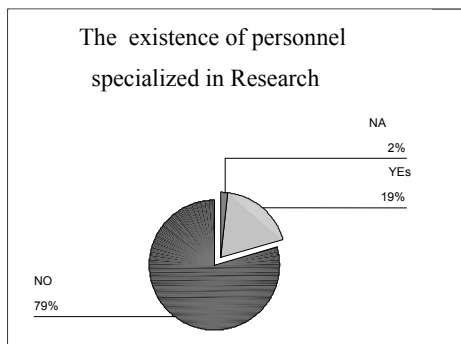
**Figure 2.** The existence of personnel trained in Evaluation



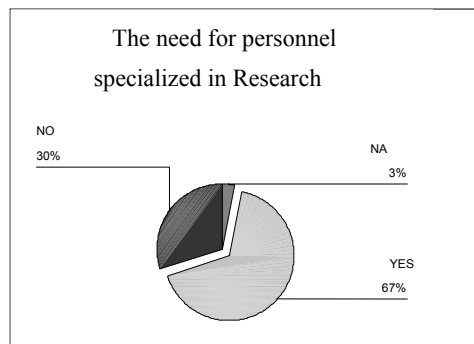
**Figure 3.** The need for personnel trained in Evaluation

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.



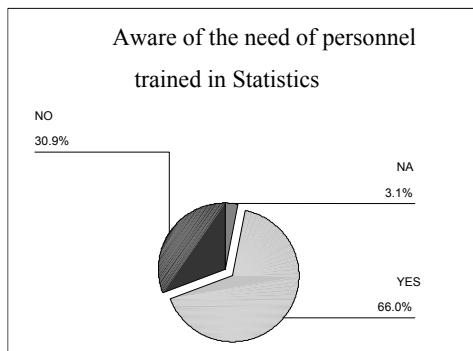
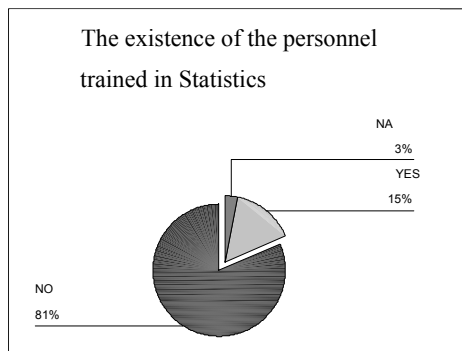
**Figure 4.** The existence in the institutions of the personnel trained in Research



**Figure 5.** The need for personnel trained 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.



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

**Figure 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.

### 3. 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>