Program Evaluation: Its Significance and Priority for Shaping and Modification of Public Policies: A Comparative Analysis

Zomorrodian, Asghar
Union Institute and University, USA
asghar.zomorrodian@myunion.edu
Matei, Lucica
National School of Political Studies and Public Administration, Romania
lmatei@snspa.ro

ABSTRACT
Although the primary purpose of evaluation is to improve the quality of programs by identifying their strengths and weaknesses, its roles becomes more instrumental in determining the fate of crucial policies and programs, when stakeholders and constituents are politically sensitive towards a given policy and allocation of scarce resources to them, particularly in bad economic situation. Policies like economic recovery, healthcare reform in U.S., and similar economic programs at the regional and global levels in other countries fall in that category. Assessing the organizational capability in achieving the desired policy outcomes and the gap between the program plan and its implementation is another vitally important result that can be obtained via sound evaluation process. This paper presents a case on the significance of policy and program evaluation and how a disciplined approach to such evaluation can provide factual evidences on the utility and impact of a given policy or program and gaining support from its sponsors and constituents. A review of evaluation traditions is followed by making reference to a few examples from United States and Romanian scenes.

INTRODUCTION
Policy and program evaluation had been around for quite a while and different public, private and nonprofit entities have spent lots of time and financial resources in doing such evaluation for both political and organizational purposes. This issue has become a priority recently as a response to more effective policy and program initiatives in order to cope with unprecedented global economic crisis by both public and private sectors. Such formidable problems call for more diligent and effective policy and program execution in order to optimize the use of scarce resources. As noted by an author recently, “The dire economic situation in the United States that was unfolding in the past two years with housing bubble bursting like a big explosion and stretching over like a tsunami over the financial market has caused the nation and the world an unprecedented shock due to its enormity, pervasiveness, and the requirement of huge amount of resources-in some cases even beyond the capability of governments- to cope with.” (Zomorrodian, 2009) The urgency of dealing with major problems and policy issues like unemployment crisis and healthcare reform in the U.S. that will require a series of sophisticated programs to actualize, and other regionally and globally oriented economic programs in other countries like Romania, all point out to the significance of effective and timely evaluation of policies and programs. This paper presents a case that systemic policy and program evaluation not only will help achieving more realistic assessment of the utility and effectiveness of policy intents via efficient program implementation, but also as a means to gain confidence of tax payers and policy constituents and targets by providing factual data and tangible results of program implementation as well as allowing for their participation in that process and eventually gaining their support. This in turn will prevent attempts to impede crafting
and implementing sound policies that otherwise will be subject the public to scare tactics, partisan politics and misinformation. Examples for policy and program evaluation approaches form U.S. and Romania will be addressed as a comparative look at this issue.

**POLICY ANALYSIS VERSUS PROGRAM EVALUATION**

While program evaluation is the bottom line for achieving the intent of a given policy and utilization of scarce resources in an optimal way, both the distinctions and relatedness of policy analysis and program evaluation are important to see how closely the two are connected together. Policy Analysis is a policy-oriented approach, method, and collection of techniques of synthesizing available information including the results of research in order to specify alternative policy and program choices and preferred alternatives. It aims at assessing organizational goals in terms of value inputs and to specify the requisite output criteria as a basis of goal determination and measurement of outcome performance and to determine needed additional information in support as a guide for future decisions. (Rossi & Williams, 1972, Bobbie, 2004). Program Evaluation on the other hand has to with “the use of social research methods to systematically investigate the effectiveness of social intervention programs that are adapted to their political and organizational environments and are designed to inform social action in ways that improve social conditions” (Rossi et al., 2004, p. 431) Such definition and distinction refer to the fact that the results of evaluation research can be an input to policy analysis and the higher the quality of a program evaluation's results, the more helpful they will be to analysts involved in key policy questions. Naturally program evaluation methods in terms of design, theory, method, data collection an analysis tools as well as basing the evaluation research on the major program intents by involving key stakeholders will lead to both reliable and useful results for program improvement and policy decisions.

**PROGRAM EVALUATION PROCESS**

The primary purpose of evaluation is to improve the quality of a program by identifying its strengths and weaknesses in order to determine whether the allocation of scarce resources should continue for a particular service. All programs, particularly the public ones, need to be assessed to see whether the stated goals and objectives have been accomplished. In general most authors refer to three primary purposes of program evaluation:

- **a) Program Improvement:** The objective is to make specific decisions about how to improve a program. This would be done during the “formative” stages of the program and the data needed at this point would include how the program is carried out, whether the program is practical and meets the needs of the target population, as well as any unintended effects are assessed.

- **b) Continuation of the Program:** The objective of such evaluation referred to as “summative: is to make an overall decision about whether or not to continue the program and/or whether or not to disseminate the program at other sites. The type of data needed at this point would include program outcomes and effects, what it would take to implement the program, and the extent to which the program serves the specified needs at a reasonable cost.

- **c) Accountability:** The objective is to provide data to the funding agent about the merits of continuing to fund the program and/or adjusting funding, to intervene in the management of the program if necessary, and to inform policy decisions. The data needed at this point would include how well the program was meeting the intended goals and its effects on the participants of the target population. Accountability overlaps with the two other purposes.

**TYPE OF PROGRAM EVALUATION**

generally speaking program evaluation falls into two categories of Formative and Summative types
depending on the specific intent of the evaluation. While there are some commonalities and overlaps between the two types, each serves distinctive purposes for program sponsors, stakeholders, and targets.

**Formative**: this type of evaluation normally is being done during different phases of program implementation and is process oriented. The focus is on how the program is progressing in terms of resource sufficiency, utilization, effective use of organizational support systems, achieving activity outputs among others.

**Summative**: This evaluation is normally emphasizes the overall purpose of the program in terms of outcomes and intended impacts. It is a kind of impact analysis that normally serves as main vehicle for program improvement, strategic direction, and continuation as well as policy decisions that may go beyond a single program. Figure (1) shows a brief comparison of components and the emphasis of the two modes of evaluation.

**Figure 1. Comparison of Formative and Summative Evaluation**

<table>
<thead>
<tr>
<th>Formative Evaluation</th>
<th>Summative Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong>: Determine whether the program is being implemented as planned, and</td>
<td><strong>Purpose</strong>: Determine whether the program has done what it set out to do in the short- and/or long-term. Used to assess whether the program's goals and objectives, were met.</td>
</tr>
<tr>
<td>whether interim results are turning out as expected. Geared toward program</td>
<td><strong>Impact Evaluation</strong>: An assessment of program outcomes and the impact the program was supposed to achieve.</td>
</tr>
<tr>
<td>improvement.</td>
<td></td>
</tr>
<tr>
<td><strong>Needs Assessment</strong>: An assessment of the social conditions a program is intended to</td>
<td></td>
</tr>
<tr>
<td>address and the need for the program.</td>
<td></td>
</tr>
<tr>
<td><strong>Context Evaluation</strong>: An assessment of the politics, community resources, and</td>
<td></td>
</tr>
<tr>
<td>broader climate in which a program will or does operate.</td>
<td></td>
</tr>
<tr>
<td><strong>Implementation Evaluation</strong>: An assessment of whether the program is being</td>
<td></td>
</tr>
<tr>
<td>implemented as anticipated.</td>
<td></td>
</tr>
<tr>
<td><strong>Process Evaluation</strong>: Looks at the effectiveness of the process through which the</td>
<td></td>
</tr>
<tr>
<td>service is delivered, regardless of whether it is in keeping with the original</td>
<td></td>
</tr>
<tr>
<td>strategy.</td>
<td></td>
</tr>
<tr>
<td><strong>Program Process Monitoring</strong>: A process evaluation that is conducted repeatedly</td>
<td></td>
</tr>
<tr>
<td>over time emphasizing time, cost and quality targets.</td>
<td></td>
</tr>
<tr>
<td><strong>Progress Evaluation</strong>: A look at short-term outcomes to see if the program is on</td>
<td></td>
</tr>
<tr>
<td>track, prior to the point at which a full-fledged summative evaluation would be</td>
<td></td>
</tr>
<tr>
<td>appropriate.</td>
<td></td>
</tr>
</tbody>
</table>
PROGRAM EVALUATION DESIGN
A typical program evaluation consists of several major components that must be put together, each focusing on aspect of assessment for the program that is the target of the evaluation. A brief explanation on each of those components will show how a well-rounded evaluation plan can work to achieve the evaluation objectives whether, formative or summative in nature.

a. Program Theory: Program theory is defined as a set of assumptions about the manner in which a program relates to the social benefits it is expected to produce and the strategy and tactics the program has adopted to achieve its goals and objectives. Rossi et al, indicate that program theory can be looked at as two intertwined theories of “impact theory” that relates to the nature of the change in social conditions brought about by program actions, and “process theory” that depicts the program’s organizations plan and service utilization plan, (Patton 2002). In other words the process theory is a plan or a blueprint which can identify the characteristics of what the program considered to be the focus point for accomplishing a successful program goal. Program impact theory, theory is considered to be a casual theory. It describes a cause-and-effect sequence in which certain program activities are instigating causes and certain social benefits are the effects they eventually produces. In a well articulated program to evaluate, the program theory identifies the crucial program’s activities and relates to the expected process. A well designed program theory model, as foundation to which the program rests, contain three required steps that include program description, program goals and objectives and program activities.

b. The Log Model: It is a systematic and visual way to present and share your understanding of the relationships among the resources you have to operate your program, the activities you plan, and the changes or results you hope to achieve. The purpose of a logic model is to provide stakeholders with a road map describing the sequence of related events connecting the need for planned program with the program’s desired results, (W.K. Kellogg, 2004). The logic models help the structure and organize the program design that builds the self-evaluation which is based upon the shared understanding of what is to take place. During the planning phase, the logic model requires that the stakeholders examine the best practice research. The program implementation is considered the logic model that forms the core the focused management plan. Figure (2) shows the components of a typical Logic Model and how one component lead to the next.

**Figure 2. Logic Model for Program Evaluation**

<table>
<thead>
<tr>
<th>Resources</th>
<th>Activities</th>
<th>Outputs</th>
<th>Short- &amp; Long-Term Outcomes</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>In order to accomplish our set of activities we will need the following:</td>
<td>In order to address our problem or asset we will conduct the following activities:</td>
<td>We expect that once completed or under way these activities will produce the following evidence of service delivery:</td>
<td>We expect that if completed or ongoing these activities will lead to the following changes in 1-3 then 4-6 years:</td>
<td>We expect that if completed these activities will lead to the following changes in 7-10 years:</td>
</tr>
</tbody>
</table>

e. The Stakeholders. A key step in the evaluation process is determining who the stakeholders are. For most social programs, the pivotal stakeholders are policymakers, program managers and/or directors, donors and funding agency representatives, program beneficiaries, other citizens, local leaders, and representatives from collaborating agencies who might be interested in the outcome of the evaluation project. Involving stakeholders in the process of evaluation is important and if that happens from the beginning it will ensure that the evaluation addresses their concerns.

Stakeholders can provide input into the evaluation plan may include the following:

i. Agency administrators who can determine whether the evaluation plan is consistent with the agency's resources and evaluation objectives.

ii. Program staff that can provide feedback on whether the evaluation will involve an excessive burden for them and whether it is appropriate for program participants.

iii. Advisory board members who can assess whether the evaluation will provide the type of information most important to know.

iv. Participants and community members who can determine if the evaluation instruments and procedures are culturally sensitive and appropriate.

While some stakeholders may have interests in both areas of the program and in its evaluation, it is important for the evaluator to think about these two interests as discrete topics. 

Program Interest has to do with getting input from a variety of stakeholders about how a program works, what it does well, what it does not do well, what needs to be evaluated, the best way to do the evaluation that can give the evaluator invaluable insights into the operations of the program. Anyone who is involved with the program in any capacity could potentially have an interest in the program, and information to share with an evaluator.

Evaluation Interest deals with stakeholders with an interest in the evaluation itself that is typically a subset of those interested in the program. This group cares specifically about the evaluation and what it shows. These are the people who will (or should) actually use the evaluation's results. Some stakeholders with an interest in the program may have little interest in the evaluation itself.

Program beneficiaries, for example, generally have little interest in evaluation results. They want the program to serve them, and would like to see it improve. But program beneficiaries would not generally do anything with the evaluation results, nor would they be directly affected by those results. (Wield and Sockeye, 1995) An advocacy group working on behalf of such beneficiaries, on the other hand, might be very interested in the evaluation results. One has also to bear in mind that all aspects of program evaluation are inherently political! Some evaluations are blatantly political in a party politics sense. Finally, in addition to finding that balance, the evaluator needs to communicate the evaluation design clearly to the stakeholders. This is a crucial step, since the results of the study will only be put to use if the stakeholders believe that the results come from a credible study that addresses their concerns.

d. The Evaluation Questions. Evaluation questions are the vehicle which lead us to gather relevant information as to whether the outputs and impact of the program have been achieved. There are a couple important points to consider in that respect. First, is the distinction between an issue area with policy implications and a specific program. This distinction between issues/topics and programs remains important as the evaluation questions are articulated. When we want to investigate an issue, we may ask questions such as what is known about the topic, what techniques have been shown to be effective in addressing certain types of problems, what conceptual frameworks have been used to understand the issue, and so on. The researcher turns to the literature, investigates what is known, and is able to answer many of these important types of questions. The evaluator may find it useful to investigate these topics in order to develop a strong foundation from which to understand the program and its results. But none of these questions get at the goal of evaluating a specific program. Evaluation research questions must tie directly to the logic model. Second, when we want to investigate an issue, we may ask questions such as what is known about the topic, what techniques have been shown to be effective in addressing certain types of problems, what conceptual frameworks have been used to understand the issue, and so on. The
researcher turns to the literature, investigates what is known, and is able to answer many of these
important types of questions. Such questions are issue based in contrast to the questions of facts
that focus on minute details or facts about your program used in the evaluation process. Naturally
the questions during the program implementation are formative in nature, focusing on the process
and how the program is progressing according to the plan and at the later stage, at the end or for a
program that has been in place for a while, the questions will be more focused on the quality of
outputs, the impact, and if the program has really fulfilled what was expected from it by key
stakeholders and program sponsors.

f. Evaluation Method. In the final analysis research questions are driving forces for the type of
evaluation to be conducted. For example if we approach questions from a broad perspective, we
will almost certainly see multiple ways to collect information that could be used to answer them.
The nature of the questions normally determines what type of research approach should be
selected. The three main approaches are qualitative, quantitative, and mixed, each with its own
advantages and disadvantages as well as fitness for a given type of program and stakeholders’
interests.

The Quantitative Tradition is a common way of thinking by focusing on numbers. While this might
be a good rule of thumb, one may often get confused by this. A survey, for example, seems on the
face of it to involve words rather than numbers. But ultimately the survey will be used to figure out
how many people responded in which ways by conducting a statistical analysis. The analysis of
the survey data will be quantitative. One way to avoid this sort of confusion is to focus not so much on
the method of obtaining the data, but rather on how the analysis will be done. If the analysis will
involve statistics, you’re in the quantitative realm. If not, you’re looking at a qualitative method.
Quantitative methods grew out of the physical and natural sciences. McNabb (2002) captures it
well with his quote of the common mantra of quantitative researchers: "If it can't be measured, it
can't be studied" (P. 82). Quantitative methods generally are used to answer research questions that
involve the concepts of "How much?" or "How many?" If a question involves a comparison of
differences between groups, quantitative methods are often the right answer. If we want to look at
the significance of the association between two variables, quantitative methods are the way to go.
Although in recent years, a strong qualitative tradition has grown up to challenge the dominance of
the quantitative tradition, to some it is still the only "real" way to do research.

The Qualitative Tradition: One way of thinking about qualitative methods is that they focus on
words, pictures, or actions – not numbers. This, too, is generally a good rule of thumb. However,
there are exceptions. Consider an analysis of how a particular minority group is pictured in the
media, and how that has changed over time. Clearly, this analysis would involve images. But the
analysis could be a statistical one, looking at the frequency of the various presentations. In this
case, it would be a "How much?" or "How many?" question, despite the fact that the raw data come
from images. The typical qualitative questions are "How?" and "Why?" These are not questions
that can be answered with a statistical analysis. To understand and explain how or why something
happens is a very different skill than being able to assess how often it happens, or how many times
it happens to people who are members of different subgroups. And while the quantitative analysis
can be of significant value, it is often just as valuable to decision makers to obtain a qualitative
analysis of why those quantitative figures look the way they do. The qualitative tradition is not as
well understood, and its methods are not taught or cited nearly as frequently. However, the basic
skills required for this type of analysis are taught and used (Creswell, 1998) In spite of the fact that
looking for patterns in information, or interpreting underlying themes are taught in classes on
history, literature, philosophy, and similar fields of study, until recently, the scientific method as
conceived by physical and natural scientists has not supported the use of these skills in research.
This is changing. Qualitative research methods are becoming better understood and more widely-
used based on generally agreed-upon standards for conducting it well. It is a tradition that is
maturing and that offers many fruitful avenues for research. Fig (3) shows comparison between the
qualitative and quantitative approaches.
Figure 3. Comparison of Quantitative and Qualitative Approaches

<table>
<thead>
<tr>
<th>Qualitative</th>
<th>Quantitative</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Characteristics</strong></td>
<td><strong>Key Characteristics</strong></td>
</tr>
<tr>
<td>▪ Answers how and why</td>
<td>▪ Answers what and how much</td>
</tr>
<tr>
<td>▪ Open-ended format</td>
<td>▪ Closed-ended format</td>
</tr>
<tr>
<td>▪ Researcher as instrument</td>
<td>▪ Instruments or secondary data</td>
</tr>
<tr>
<td>▪ Used to provide rich descriptions, as well as to enhance our</td>
<td>▪ Used to &quot;slice up&quot; a phenomenon into manageable and discrete</td>
</tr>
<tr>
<td>understanding of the context of a phenomenon.</td>
<td>elements of an overall conceptual framework</td>
</tr>
<tr>
<td>▪ Seeks to identify patterns among variables and to make</td>
<td>▪ Relies on conceptualization, statistical tools, and operationalization of variables of interest, valid and reliable measurement</td>
</tr>
<tr>
<td>distinctions</td>
<td></td>
</tr>
<tr>
<td><strong>Approach</strong></td>
<td><strong>Approach</strong></td>
</tr>
<tr>
<td>▪ Narrative data (words)</td>
<td>▪ Countable data (numbers)</td>
</tr>
<tr>
<td>▪ More subjective, based on researcher perceptions</td>
<td>▪ More objective, with researcher efforts to remain 'invisible'</td>
</tr>
<tr>
<td><strong>Limitations and Benefits</strong></td>
<td><strong>Limitations and Benefits</strong></td>
</tr>
<tr>
<td>▪ Typically has a small sample size</td>
<td>▪ Often has a large sample size</td>
</tr>
<tr>
<td>▪ Labor intensive; data are expensive to collect and analyze</td>
<td>▪ Less labor-intensive; data are cheaper to collect and analyze</td>
</tr>
<tr>
<td>▪ May lack important data (e.g., quantifiable differences observed)</td>
<td>▪ May miss important data (e.g., perceptions, values, and unintended outcomes)</td>
</tr>
<tr>
<td><strong>Common Techniques</strong></td>
<td><strong>Common Techniques</strong></td>
</tr>
<tr>
<td>▪ Observation</td>
<td>▪ Questionnaires</td>
</tr>
<tr>
<td>▪ Document review</td>
<td>▪ Experimental</td>
</tr>
<tr>
<td>▪ Focus groups</td>
<td>▪ Tests</td>
</tr>
<tr>
<td>Interviews (open-ended and semi-structured)</td>
<td>Secondary Data (e.g., existing databases)</td>
</tr>
</tbody>
</table>

The Mixed Methods Tradition: There are benefits and drawbacks to any research method. One underlying theme is that the best approach to evaluation is to collect information from multiple sources, using multiple methods and multiple perspectives. Why is this broader perspective important? In the early 1970s, program evaluation became increasingly quantitative due to changes in technology and demands for greater accountability by funding sources. Nevertheless, in spite of the sophistication of the evaluation designs and methods, experienced evaluators found that important “parts of the story” were being missed (Frechtling et al., 2002). Adding qualitative methods to the well-established quantitative ones allowed researcher to answer a broader array of questions. Moreover, mixing methods allowed for triangulation, which could enhance our confidence in the results. Used properly, each method can inform the other and strengthen an evaluation’s findings. Mixed method evaluations often use qualitative research at the beginning of the process. Observations, document reviews, and semi-structured interviews might be used, for example, to inform the development of a survey instrument. Once the survey is conducted and quantitative results obtained, a follow-on qualitative piece might be used to explore the reasons that certain unexpected findings emerged. Those follow-on interviews can then inform the development.
of further quantitative work. One important caveat about using mixed methods is that it can be the equivalent of conducting two separate studies, thus consumes more time and resources.

**Evaluation Implementation and its Challenges.** As a practical endeavor, program evaluation intention is to find tangible ways to improve social programs. However insightful the evaluation's findings may be, they will be worth little without successful program changes to address those findings. The challenge is to find ways to initiate and follow through on changes when the status quo must be altered. In the end, a successful program evaluation will produce information that can be used to develop a detailed plan that sets forth explicitly what is to be done, who is responsible for what, and when each step must be completed. Depending on the evaluator's role in the process, the evaluation may or may not move to the stage of providing an implementation plan. An outside evaluator, for example, may suggest general directions for change, but not be in a position to develop concrete change plans. The evaluation report should, however, lay the foundation for internal staff who will develop such plans. In other situations, the evaluator may be involved in developing the implementation strategy. In either case, the evaluation report should provide a clear roadmap to direct the efforts of those attempting to implement needed changes. A well-conceived implementation plan will contain many of the same elements as a program evaluation. The similarity between the first four implementation steps identified by Nicholas et al (2001) and the steps we have taken in developing our program evaluation designs are striking:

- Identify the **target group** who is the focus of the program.
- Stipulate the **goals** of the program or the desired **outcomes**.
- List the program **objectives**; identify specific measures that will be used to assess whether the goal(s) or outcomes have been achieved.
- Identify the **activities or strategies** for achieving program goals or objectives.
- Identify **tasks** to be completed and **target dates** for putting the strategies/activities into place.
- Identify **authority figures and resources** needed to accomplish the tasks.
- Develop a plan for **monitoring the implementation** of the program and the outcome evaluation

One of the primary challenges evaluators face is ensuring that adequate resources are available to enact and carry out the plan. The commitment of adequate resources and staff is essential to successful implementation. Mazmanian and Sabatier (1989) have written extensively on implementation and note that one of the biggest factors in program or policy implementation failure is the lack of adequate resources. Another critical issue involves the management team. If key staff do not support the changes being implemented, then success is virtually impossible. In addition, administrative processes need to be reviewed and adjusted continuously to accommodate the changes and make sure that continuing support of all stakeholders and targets of the program in the evaluation will lead to building confidence in the obtained evaluation results as noted earlier.

**LEVELS OF PROGRAM EVALUATION AND PUBLIC POLICY DECISIONS**

As mentioned before while the goal of evaluation is to determine the worth, effectiveness or efficiency of some program, procedure, project, process, or product, well-designed evaluations provide information that can help explain the findings that are observable to stakeholders and aid to decision makers to become more effective and efficient. Formative evaluation delves into more tacit research at a project level, while summative research looks at the macro strategic issues at the overall program level after the inputs integrating the outcome from multiple studies to arrive at judgment on an evaluation question. Kellogg foundation has developed three levels of evaluation that helps understanding the broad context of evaluation and together they maximize the collective understanding and ability to strengthen individual and group projects in grant-making in their case that seems to be equivalent to policy making decision ( )
Project-Level Evaluation: Project-level evaluation is the evaluation that project directors are responsible for locally. The project director, with appropriate staff and with input from board members and other relevant stakeholders, determine the critical evaluation questions, decides whether to use an internal evaluator or hire an external consultant, and conducts and guides the project-level evaluation. The Foundation provides assistance as needed. The primary goal of project-level evaluation is to improve and strengthen Kellogg-funded projects. Ultimately, project-level evaluation can be defined as the consistent, ongoing collection and analysis of information for use in decision making (P. 16). Project-level evaluation should not be a stand-alone activity, nor should it occur only at the end of a program. Project staff should think about how evaluation can become an integrated part of the project, providing important information about program management and service delivery decisions. Evaluation should be ongoing and occur at every phase of a project’s development, from preplanning to start-up to implementation and even to expansion or replication phases. For each of these phases, the most relevant questions to ask and the evaluation activities may differ. What remains the same, however, is that evaluation assists project staff, and community partners make effective decisions to continuously strengthen and improve the initiative.

Cluster Evaluation: Increasingly, we have targeted our grant-making by funding groups of projects that address issues of particular importance to the Foundation. The primary purpose for grouping similar projects together in “clusters” is to bring about more policy or systemic change than would be possible in a single project or in a series of unrelated projects. Cluster evaluation is a means of determining how well the collection of projects fulfills the objective of systemic change. Projects identified as part of a cluster are periodically brought together at networking conferences to discuss issues of interest to project directors, cluster evaluators, and the Foundation. Project directors typically know prior to receiving a grant whether they will be expected to participate in a cluster; but occasionally clusters are formed after grants have been made. Therefore, it is important to be familiar with cluster evaluation even if you are not currently participating in a cluster. In general, we use the information collected through cluster evaluation to enhance the effectiveness of grant making, clarify the strategies of major programming initiatives, and inform public policy debates (P. 18).

Programming and Policymaking Evaluation: Program and policymaking evaluation is the most macro form of evaluation at the Foundation. Conducted by the Foundation’s programming staff, it addresses cross-cutting programming and policy questions, and utilizes information gathered and synthesized from both project-level and cluster evaluation to make effective decisions about program funding and support. This type of evaluation also supports communities in effecting policy change at the local, state, and federal levels. Taken together, the three evaluation levels provide multiple perspectives, multisource, multilevel data from which to strengthen and assess individual and groups of projects. The interaction of professionals that occurs across all three levels of evaluation encourages creative and innovative thinking about new ways to evaluate programs and deliver information, which we hope will ultimately lead to sustained positive change at the community level. At the same time, evaluation information from multiple levels, when examined in holistic ways, helps the staff members make effective and informed decisions regarding our programming and policy work. (Evaluation Handbook, P. 15) The key to collecting data is to collect it from multiple sources and perspectives, and to use a variety of methods for collecting information. The best evaluations engage an evaluation team to analyze, interpret, and build consensus on the meaning of the data, and to reduce the likelihood of wrong or invalid interpretations.

THE U.S. EXAMPLES
Since United States is now facing with some serious national policy decisions and dilemmas like healthcare, economic recovery, environment, energy, war, and host of others long term implications both financially and otherwise, a reference to healthcare reform as main challenges
Facing the nation and how program evaluation can help to validate the policy, create opportunities for improvement, and provide a realistic picture as program relevant to such policy implementation is in order. The following are four excerpts from the General Accounting Office (GAO) on different aspects of healthcare polices and statutes that show a) even at a very large and macro level, a policy can be divided into program components, and b) each program can be subject to continuous, or periodic (that is formative) and or final/impact (summative) evaluation. These examples show how evaluation can cover a host of different aspects of complex policy system and its subsystems like providing healthcare to specific group (HIV affected, the CARE Act), use of feedback from service providers for Medicare (The Medicare Improvements for Patients and Providers Act of 2008), health insurance enrollment for high risk individuals at the State level (HRSAs Health Plans), and finally to improve the health protection by more control over contracting authorities (BioShield Act of 2004).

**Figure 4: Ryan White CARE Act: Health Resources and Services Administration's Implementation of Certain Provisions Hampered by Lack of Timely and Accurate Information**

**Source:** GAO-09-1020 September 29, 2009

**Summary**

Under the CARE Act, funds are made available to assist over 530,000 individuals affected by HIV/AIDS. Grantees directly provide services to individuals (clients) or arrange with service providers to do so. The Department of Health and Human Services’ (HHS), Health Resources and Services Administration (HRSA), which administers CARE Act programs, is required to cancel balances of grants that are unobligated after one year and redistribute amounts to grantees in need. HRSA began to collect client-level data in 2009. Under the CARE Act, states and territories receive grants for AIDS Drug Assistance Programs (ADAP), which provide HIV/AIDS drugs. GAO was asked to examine elements of the CARE Act. In this report, we review: (1) HRSA's implementation of the unobligated balance provisions, (2) HRSA's actions to collect client-level data, and (3) the status of ADAP waiting lists. GAO reviewed reports and agency documents and interviewed federal officials, officials from 13 state and 5 local health departments chosen based on location and number of cases, and other individuals knowledgeable about HIV/AIDS.

The lack of timely and accurate information reporting by grantees has delayed HRSA's distribution of certain grants and has placed at risk HRSA's ability to obligate these funds. The late submission of actual unobligated balances for the 2007 grant year delayed HRSA's ability to determine grantees' unobligated balances and redistribute these funds to other grantees. A number of grantees were late in their submissions. For example, 21 of the 56 metropolitan areas submitted their information beyond the date initially set by HRSA. Additionally, some grantees reported inaccurate unobligated balances, which required HRSA staff to correspond with grantees and request revised information, creating additional delays. HRSA is authorized to obligate fiscal year 2007 funds for a 3-year period and is at risk of losing the authority to make grants from these funds. HRSA officials said they have made changes to how they implement the unobligated provisions in an effort to avoid these issues in the future. HRSA has taken actions to collect client-level data by implementing a new data collection and reporting system. However, some grantees and service providers did not submit the initial reports by HRSA's deadline. HRSA set a July 31, 2009, submission deadline for grantees' initial reports, but 100 of 638 grantees did not meet this deadline. Client-level data includes information such as the dates clients were served, the types of services provided, and the clients' health status. HRSA has implemented a system to collect data on the number of unique clients from grantees and service providers that will allow HRSA to determine the services each client received and the outcomes of these services. In order for HRSA to collect this information, grantees and service providers must first collect the data using their own systems, and HRSA has provided technical and financial assistance so that they can develop these systems. For example, under a project initiated in 2009, HRSA awarded approximately $4 million to CARE Act grantees for the development of their own client-level data collection systems. The number of ADAPs with waiting lists and the number of individuals on those lists is increasing. In the first quarter of grant year 2008 (April 1, 2008, through June 30, 2008), 2 ADAPs had waiting lists with a total of 55 people on those lists; this grew to 3 ADAPs and a total of 112 people in the fourth quarter of the year, and increased to 4 ADAPs and 136 individuals in August 2009. Kentucky, Montana, Nebraska, and Wyoming were each maintaining a waiting list for ADAP services in August 2009; Nebraska had the largest number of individuals (71), and Wyoming had the smallest number (5). ADAP officials expressed concern that they will have to establish or expand waiting lists or implement other cost-control measures, such as limiting the number of drugs they make available.

**Recommendations for Executive Action**

**Recommendation:** To help ensure that HRSA is able to implement the unobligated balance provisions in a timely manner, and to obtain timely and accurate information on grantees' unobligated balances, the Secretary of HHS should instruct the administrator of HRSA to identify the causes of grantees' difficulties in providing a timely and accurate accounting of their unobligated balances.

**Agency Affected:** Department of Health and Human Services
Summary
The Medicare Improvements for Patients and Providers Act of 2008 directed the Secretary of Health and Human Services to develop a program to give physicians confidential feedback on the Medicare resources used to provide care to Medicare beneficiaries. GAO was asked to evaluate the per capita methodology for profiling physicians—a method which measures a patient's resource use over a fixed period of time and attributes that resource use to physicians—in order to assist the Centers for Medicare & Medicaid Services (CMS) with the development of a physician feedback approach. In response, this report examines (1) the extent to which physicians in selected specialties show stable practice patterns and how beneficiary utilization of services varies by physician resource use level; (2) factors to consider in developing feedback reports on physicians' performance, including per capita resource use; and (3) the extent to which feedback reports may influence physician behavior. GAO focused on four medical specialties and four metropolitan areas chosen for their geographic diversity and range in average Medicare spending per beneficiary. To identify considerations for developing a physician feedback system, GAO reviewed the literature and interviewed officials from health plans and specialty societies. Further, GAO drew upon literature and interviews to develop an illustration of how per capita measures could be included in a physician feedback report.

Using 2005 and 2006 Medicare claims data and a per capita methodology, GAO found that specialist physicians showed considerable stability in resource use despite high patient turnover. This stability suggests that per capita resource use is a reasonable approach for profiling specialist physicians because it reflects distinct patterns of a physician's resource use, not the particular population of beneficiaries seen by a physician in a given year. GAO also found that our per capita method can differentiate specialists' patterns of resource use with respect to different types of services, such as institutional services, which were a major factor in beneficiaries' resource use. In particular, patients of high resource use physicians used more institutional services than patients of low resource use physicians. GAO identified four key considerations in developing feedback reports on physician performance. To illustrate how per capita measures could be included in a physician feedback report, we developed a mock report containing three types of per capita measures. Although the literature suggested that feedback alone has no more than a moderate influence on physicians' behavior, the potential influence of feedback from CMS on Medicare costs may be greater, in part because of the relatively large share of physicians' practice revenues that Medicare typically represents. CMS reviewed a draft of this report and broadly agreed with our findings.

Summary
A growing number of states—35 as of June 2009—have created high-risk health insurance pools (HRPs) primarily to provide coverage to individuals whose health status limits their access to coverage in the private individual health insurance market. HRPs—typically state-run nonprofit associations—often contract with a private health insurance carrier to administer the pool and offer a range of health plan options to such individuals, who are commonly referred to as medically uninsurable. Plan options vary within pools and from state to state, and like the private individual market, HRPs typically impose waiting periods for coverage of preexisting conditions to discourage medically uninsurable individuals from foregoing health insurance until they require care. Because of the higher health care costs typically incurred by medically uninsurable individuals, all pools operate at a loss. Premiums for HRP health plans are higher than for plans offered to healthy individuals in the private health insurance market; however, these premiums are capped to limit enrollees' costs and are thus insufficient to cover the costs of enrollee health care claims. As a result, all HRPs supplement their revenues through various funding mechanisms, such as assessments on health insurance carriers and state general revenues. Federal grants are also awarded to establish and fund HRPs. As part of the Trade Adjustment and Assistance Reform Act of 2002, Congress established a program to provide grants to HRPs to offset losses and establish HRPs—commonly referred to as operational and seed grants, respectively. Subsequent legislation authorized funding for the program through 2010, including grants to be used for supplemental consumer benefits—commonly referred as bonus grants. Since 2003, the grant program has awarded nearly $286 million to state HRPs for various purposes. The Centers for Medicare & Medicaid Services (CMS), with the Department of Health and Human Services (HHS), administers this federal grant program. Recent health care reform proposals call for an expanded role for HRPs to enhance health insurance options for the medically uninsurable. Because of the federal funding provided to HRPs, you expressed interest in obtaining data on several aspects of each state HRP. In this report, we describe (1) HRP enrollment and enrollee demographics; (2) HRP plans' cost-sharing provisions, coverage restrictions, and premiums, and comparable information for certain private market health plans; and (3) HRPs' governance, expenditures, and funding.

HRP enrollment—enrollees and their dependents—totaled 199,418 in the 34 HRPs in 2008. We estimated nearly 4 million additional individuals to be potentially eligible for enrollment in an HRP based on their uninsured status and preexisting health conditions. In 2008, the average annual deductible for the most popular plan offered by each of the 34 HRPs was $1,593—almost
three times as high as the average annual deductible of $560 among employer-sponsored health insurance plans. About 63 percent of enrollees in these most popular HRP plans had deductibles of $1,000 or greater. In comparison, almost 88 percent of enrollees in employer-sponsored plans had a deductible of under $1,000 or no deductible. Collectively, in 2008, HRP governing boards included representatives from health insurance carriers (41 percent of board members), state government (13 percent), medical providers (9 percent), enrollees (7 percent), and employers (3 percent), among others. Total claims paid by HRPs in 2008 were about $1.9 billion, accounting for almost 95 percent of total HRP expenditures. The average claims per enrolled individual totaled $9,437 in 2008, an increase of about 39 percent since 2003. In 2008, premium revenue contributed 54 percent of HRP funding, and insurance carrier assessments contributed about 23 percent. HRPs were awarded a total of approximately $286 million in federal grants between 2003 and 2008, according to CMS. Almost 83 percent of these funds were operational grants, almost 15 percent were bonus grants, and less than 3 percent were seed grants. Federal grants comprised less than 2 percent of total HRP funding in 2008.

**Figure 7: Project BioShield: HHS Can Improve Agency Internal Controls for Its New Contracting Authorities**

**Summary**
The Project BioShield Act of 2004 (BioShield Act) increased the federal government's ability to procure needed countermeasures to address threats from chemical, biological, radiological, and nuclear agents. Under the BioShield Act, the Department of Health and Human Services (HHS) was provided with new contracting authorities (increased simplified acquisition and micropurchase thresholds, and expanded abilities to use procedures other than full and open competition and personal services contracts) and was authorized to use about $5.6 billion in a Special Reserve Fund to procure countermeasures. Based on the BioShield Act's mandate, GAO reviewed (1) how HHS has used its purchasing and contracting authorities, and (2) the extent to which HHS has internal controls in place to manage and help ensure the appropriate use of its new authorities. To do this work, GAO reviewed contract files and other HHS documents, including internal control guidance, which GAO compared with federal statutes and federal internal control standards.

Since 2004, HHS has awarded nine contracts using its Special Reserve Fund (Fund) purchasing authority under the BioShield Act to procure countermeasures that address anthrax, botulism, smallpox, and radiation poisoning. HHS may procure countermeasures that are approved by the Food and Drug Administration and ones that are unapproved, but are within 8 years of approval. Of the nine contracts, one was terminated for convenience and the remaining eight are valued at almost $2 billion. HHS officials told GAO that additional contracts are likely to be awarded in the near future as the Fund provides funding through fiscal year 2013. In addition, HHS has used one of its new contracting authorities, simplified acquisition procedures, although it has not used this authority since 2005. HHS has established internal controls on its new purchasing and contracting authorities. In addition to the language in the BioShield Act, which sets up a broad framework of controls over the use of the Special Reserve Fund, the internal controls for this purchasing authority are documented in a variety of internal policy and procedure documents and interagency agreements, which provide guidance on roles and responsibilities for how the controls are to be implemented. In response to BioShield Act requirements, HHS also established internal controls for three of the contracting authorities: the increased simplified acquisition threshold and its use with Special Reserve Funds, the increased micropurchase threshold, and the use of personal services contracts. Federal internal control standards state that, among other things, management needs to comprehensively identify risks, analyze them for possible effect, and determine how risks should be managed. Although some of the risk statements in a memo HHS issued identify some risks and one mentions possible negative consequences that could occur without proper controls in place, the risk statements for using the increased micropurchase threshold and increased simplified acquisition procedures lack analysis of specific risks. In particular, the memo does not discuss a key risk associated with using simplified acquisition procedures—namely, that an agency is prohibited from obtaining cost or pricing data for acquisitions at or below the simplified acquisition threshold. Without this data, the agency may not be able to determine if the price of a contract is fair and reasonable. Moreover, not having adequately documented and appropriately communicated risk assessments potentially results in future employees not knowing or understanding the risks or trade-offs involved in using the authorities. With employee turnover, HHS' reliance on the knowledge of current personnel to appropriately implement key controls will not enable future employees to make sound, informed, and consistent decisions.

**Status:** In process

**ROMANIAN SCIENCE: PROGRAM EVALUATION AND PUBLIC POLICY DECISION**
Generally speaking public policy in Romania can be explained through the following stages: I. Policy Specs and Development. This stage follows the principles applicable in the European space and include components like: the existence of a general legislative framework; valid and coherent for formulating public policies; autonomy of the ministries in elaborating own public policies, their standards and priorities; inter-ministerial relationship; avoiding re-
organization/reforms with unpredictable changes; and internationalization of governmental policies marked by accession to the EU, development of globalization for economic processes, NATO membership, and European Council membership.

II. Levels of Public Policy Process and decision making. The public policy system as the sum of instruments, procedures and institutional mechanisms to improve the quality and efficiency of the decision-making process assumes a good collaboration between the four levels identified as areas for public policy actors:

1. The political level which refers to the content of strategies and political programs assumed by the Government and ministries, in sectoral issues.
2. The legislative level which refers to the sum of international regulations to be assumed by Romania.
3. The executive level represented by the Government and includes also the managerial level which refers to the matters of functionality of ministries and public institutions.
4. The consultative level which refers to the relations developed by the Government and ministries with civil society, media and citizens.

III. Public Policy Approaches. The institutional and legislative approach to public policies is based on two premises: a) institutional management using strategic resource planning and efficient institutional models like the one used in private sector, and b) adopting a modernized legislative framework. The Romanian public policy process is based on several principles, some relate to administrative mechanisms and other relate to the expectations of key Stakeholders (Guide for Elaboration, 2004)

Principle of participation and transparency
Principle of continuity and coordination
Principle of responsibility (all political and administrative levels)
Principle of good governance
Principle of subsidiarity
Principle of cooperation and coherence

Figure (8) shows the role of stakeholders in substantiating, elaborating, implementing and assessing the public policy.

<table>
<thead>
<tr>
<th>Defining S: the stakeholder</th>
<th>S’s interest in policy substantiation</th>
<th>S’s interest in policy elaboration</th>
<th>S’s interest in the policy implementing</th>
<th>S’s interest in policy assessment</th>
<th>Resources available to the S</th>
<th>S’s resource mobilization capacity</th>
<th>S’s position to policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concerned ministry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public sector</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private sector</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional associations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Implementation and Assessing the public policies  The implementation of the public policy is supported in some cases also by the regulation process. It imposes quality requirements such as consultation, impact studies, requiring high costs, sometimes higher than the value of the estimated benefit, even leading to the critics concerning the efficiency of the respective regulation. Assessment emphasizes the capacity to forecast and approximate the outcomes of implementing the public policies, the sizes of efficiency, effectiveness and degree of saving, reflected in the results, comparable and measurable with the forecasted ones, and capacity to comply with the content of the activities achieved before the deadlines specified in the action plans for the public policies.  The activities of monitoring and assessment are taking place during and after the implementation of public policies. They follow the degree of achievement of the public policy’ objectives and take place at the level of every central government authority. Their object consists in the activities and results of the policy making process as indicated in Figure (9).

The impact analysis of the public policy allows political decision makers to formulate a perspective regarding the consequences of the actions to be accomplished and the assessment of the effectiveness of actions to be achieved as sown in Figure.

Figure 9: Typology of the assessments in the public policy cycle

Example, Policy Formulation: One example in formulating a public policy is the one of public debt initiated by the Ministry of Economy and Finance, supervised by the experts of the World Bank and IMF and those of the PHARE Project RO 02 586/03.04.03 titled: "Enhancing the management system of the state Institution in charge for assessing the public policy.  As shown n figure (10) this public policy process shows Target groups, General objective, Specific objectives, Accomplished objectives, Activity strategy and action plan as well as Identifying the authority and the necessary resource as well as interactions among the policy and process elements.
Figure 10: Process for public policy making

Institution with initiative
Ministry of Finance

Formulating the problem
Implementation of Law on public debt no. 313/2004:
Approaching at the portfolio level, the process of administrating the governmental public debt, as a result of maintaining the possibility of contracting loans by the Ministry of Finance and their sub-loaning to central public government authorities (1).

Naming the problem
Public debt management
(Administering the governmental public debt, authorizing and monitoring the local public debt)

Aim
Maintaining, in sustainable limits, the level of public debt and correlating it to the policies of reaching convergence

General objectives:
Increase of efficiency, transparency and predictability of governmental public debt management

Specific objectives
- reducing, on long term, the costs associated to the contracted governmental public debt and to the risks of non-payment of the obligations to be paid by central public government
- Development of the market of state bonds
- Optimizing the structure of the governmental public debt

Beneficiaries:
- Government, represented by the Ministry of Finance;
- Authorities of local public government;
- Economic operators;
- Financial and credit institutions

Alternatives in solving V1, V2, V3
Economic, social impact
Budget

Process of consultation

Presenting the recommended alternative:
- benefits, risks, impact, target groups, modalities and deadlines for monitoring / assessment, indicators of performance
- Deadlines involved
- Estimated budget

Implementation

Beneficiaries:
- Government, represented by the Ministry of Finance;
- Authorities of local public government;
- Economic operators;
- Financial and credit institutions

Alternatives in solving V1, V2, V3
Economic, social impact
Budget

Process of consultation

Presenting the recommended alternative:
- benefits, risks, impact, target groups, modalities and deadlines for monitoring / assessment, indicators of performance
- Deadlines involved
- Estimated budget

Implementation
Example, Policy Assessment: A good example of policy assessment is the one used by the Ministry of Economy and Finance concerning aiming at “Assuring the stability of the banking system and protecting the deponents” as presented in Figure (11)

![Figure 11: Utility and Sustainability](image_url)

The policy assessment has used financial indicators (the annual contributions’ quota of the credit institutions, the volume of the credit institutions’ contributions to the Fund, the equity volume of the Fund, the volume of the guaranteed deposits, the guarantee ceiling), quantitative indicators (the number of the credit institutions participating to the deposits’ guarantee diagram, the number of guaranteed deponents), analyzing them as dynamics during 03 – March 2008. The positive effects of reducing the contribution quota objective:

a. reducing the financial effort of the credit institutions participating to the Fund for Guarantee, with favorable consequences in the saving process;
b. increasing the taxes’ amounts owed by the credit institutions;
c. increasing the medium term guaranteed deposits;

CONCLUSION
While the intention was not to verify the validity and reliably of any of approaches discussed in this paper for policy analysis and program evaluation, the purpose was to show how a complex policy like National Healthcare reform, when get to implementation level, can be broken down into manageable programs with specific goals and impact and make them subject to continuous
evaluation in both formative and summative fashion. Agencies like GAO are positioned nicely to take charge of such tasks and provide policy makers with timely feedback and constructive feedback for revising, changing and reformulating the policies. The full report on all four evaluation reports done by GAO referred in this paper are available on GAO web site for further study and to see how such evaluation can practically help policy makers if tied to specific policy mandate. Also, such evaluation discipline will insure the two important concerns of any public policy and program that is accountability and following ethical standards. As an example ASPA code of ethics that includes items like: serve the public interest, exercise discretionary authority to promote the public interest, involve citizens in policy decision-making, respect the constitution and the law, prevent all forms of mismanagement of public funds by establishing and maintaining strong fiscal and management controls, conduct official acts without partisanship, subordinate institutional loyalties to the public good, promote organizational accountability through appropriate controls and procedures (ASPA Code of Ethics) are all considerations that can be included in a well designed program plan though the stages motioned as well as evaluating their achievements via the program evaluation plan.

REFERENCES


GAO. Medicare: Per Capita Method Can Be Used to Profile Physicians and Provide Feedback on Resource Use. Source: GAO-09-730R July 22, 2009 GAO Official Website

GAO. Health Insurance: Enrollment, Benefits, Funding, and Other Characteristics of State High-Risk Health Insurance Pools: GAO-09-820 July 21, 2009 GAO Official Website

GAO. Project Bioshield: HHS Can Improve Agency Internal Controls for Its New Contracting Authorities GAO-09-820 July 21, 2009 GAO Official Website


Romanian Government. Ministry of Economy and Finance, supervised by the experts of the World Bank and IMF and those of the PHARE Project RO 02 586/03.04.03”Enhancing the management system of the state Institution in charge for assessing the public policy: Ministry of Economy and Finance – Public Policy Unit


