Evaluating the Implementation Process: An Exploratory eGovernment Case Study

Maddalena Sorrentino
State University of Milan, maddalena.sorrentino@unimi.it

Katia Passerini
New Jersey Institute of Technology, pkatia@njit.edu

Follow this and additional works at: http://aisel.aisnet.org/amcis2009

Recommended Citation
http://aisel.aisnet.org/amcis2009/529

This material is brought to you by the Americas Conference on Information Systems (AMCIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in AMCIS 2009 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
Evaluating the Implementation Process: an Exploratory e-government Case Study

Maddalena Sorrentino  
State University of Milan  
maddalena.sorrentino@unimi.it

Katia Passerini  
New Jersey Institute of Technology  
pkatia@njit.edu

ABSTRACT
This paper approaches implementation evaluation as an integral part of the public administration organizational action. Drawing on contributions from organization theory, we assign a dual role to evaluation: valuable cognitive resource and accountability tool for the policymakers. This exploratory case study, supported by a review of secondary data sources, seeks to enrich the literature on implementation evaluation with a scientifically founded opinion on the launch of a recent e-government project by the City of Milan. As no attempt has yet been made to evaluate the effects of the initiative, our paper aims to help fill this gap. Our preliminary findings confirm the heuristic potential of an evaluation approach where interdisciplinary inputs can enlighten not only the results, but also the process of design, adoption and use of the e-services.

The paper discusses the role of the various evaluation stages, with a specific focus on process-based (in itinere) evaluations.

Keywords  
e-government, implementation, evaluation, organizational processes

INTRODUCTION
Evaluating the effects of both the operational management of the services provided and the innovation rate of policies, programs and projects launched is a crucial factor for public administrations (PA). Being accountable for the use of collective resources based on verifiable results aligned with the resources public institutions receive from society is becoming a recognized need in several spheres (Rebora, 1999).

In this paper, evaluation is defined as an opinion on the capacity of the policy (i.e. the set of actions designed to address a collective problem) to transform the state of the situation-problem in the desired direction (Dente and Vecchi, 2001). Public opinion tends to take implementation for granted, essentially because the collective interest’s main focus is on the decision-making process. Implementation is seen as a “technical” phase and as such is erroneously considered neutral and devoid of discretionary power. Contrary to this widespread opinion, implementation is an uncertain phase in which discretionary power cannot be eliminated through ex-ante standardization. This happens for various reasons. First, the operational launch verifies if what has been decided moves closer to (or further away from) the goal. Second, the implementation process itself can be dogged by unexpected events – such as cuts in available resources - placing the original project at risk. Ultimately, implementation is ambiguous because it involves the mobilization of numerous actors who may have personal agendas that do not necessarily coincide with the goals of the other stakeholders.

Figure 1 shows the evaluation can focus on: a) the “products” generated by the policy (outputs) or the effects on the recipients (the outcomes and impacts); b) the implementation, that is the actions and decisions culminating in the launch of a policy, such as the provision of a service, the enactment of a law; or c) the phases in which the political agenda takes shape (issue-making and decision-making). Depending on when it is carried out, evaluation can be ex post, when it analyzes the results (output, outcome and impacts); in itinere, when it is ongoing and conducted during implementation; or, ultimately, ex ante, when it is carried out before the implementation phase. Typically, the ex ante and ex post evaluations aim to confirm or revoke decisions already made. The in itinere evaluation – which generally responds to a broader cognitive need – seeks to account for what happens as the implementation of the policy unfolds.
This paper sets out to analyze the dynamics that precede the production of the outputs, outcomes, and impacts. It does not focus on the evaluation of the results of public policies. We maintain that the obstacles faced when evaluating implementation (considered the “black box” of policymaking) lie in the conceptual realm and need to be addressed starting there. We seek to demonstrate how organizational knowledge can help us shed new light on the evaluation conundrum. These preliminary reflections aim to offer a scientifically founded opinion on the implementation of an e-government project launched by the City of Milan (Italy). Our discussion will refer mainly to the concepts inspired by Policy Studies and Organization Theory. Due to space limitations, we will not examine the methods or measures employable in the practice nor will introduce an extensive review of evaluation procedures. For the purpose of this paper, the term ‘e-government’ signifies the use of information and communication technologies (ICT), and in particular the Internet, as a tool to achieve better government (OECD, 2004).

We illustrate our empirical case in exploratory form (Yin, 2003). This methodological choice is based on the fact that the policy cycle in question—a 12-month pilot scheme to test the City of Milan’s new air pollution regulation—was still ongoing at the time of writing. We conducted our analysis between July 2007 and January 2009 using a variety of official documentation found on the websites of the agencies involved together with press reports. We collected and reviewed several opinions and viewpoints written on the topic to build a realistic picture of the events. The analysis of the various sources of information enabled us to reconstruct several decisional processes. This analysis is only a first step in a more holistic case review that we plan to conduct in future research. We intend to supplement the analysis of secondary data with primary data collected through structured interviews and to compare the results with our preliminary assessment. Despite logistical constraints, we are keen to start the interviews within the implementation timeframe (now extended through 2009). The objective is to capture the in itinere drivers in a timely manner and identify the initial lessons learned.

The results of the analysis of secondary data sources, herein discussed in summary form, are presented as follows. After the introduction, we give a brief overview of some theoretical proposals on public-sector evaluation. The limitations of these proposals will be traced to their underlying concepts. We then illustrate the reasons that led the City of Milan to implement a pollution charge system and comment on the methods used to date to evaluate the new regulation. Finally, we discuss the implications of using a theoretically founded “toolbox” in the evaluation practice.

**EVALUATION IN THE LITERATURE: THE ISSUES**

Evaluation can take many meanings depending on the context and the disciplinary perspective adopted by the observer. According to Rossi and Berk (1981), cited in the seminal work of Pressman and Wildavsky (1984, p. 185):

> “Evaluation research may be conducted to answer questions that arise during the formulation of policy, in the design of programs, in the improvement of programs, and in testing the efficiency and effectiveness of programs that are in place or being considered. Specific policy questions may be concerned with […] whether any program can be enacted that will ameliorate a problem, whether programs are effective, whether a program is producing enough benefits to justify its cost […]”

Rossi and Berk underscore that evaluation research is an applied discipline. It is not a generic study of social change, but, conversely, addresses a specific intervention (our italics, Lippi, 2007) launched by an individual or collective client (policymaker, program executives, public-sector managers, interest groups). Five basic questions enable us to chart a Policy Evaluation research pathway: When? Where? For Whom? What? Why? (Browne and Wildavsky, 1983, p. 206). For example, in responding to the first question (When?), the evaluation must establish whether it is appropriate to conduct an evaluation that is retrospective (ex post), ongoing (in itinere), or pre-implementation (ex ante). The question Where? concerns the choice of settings to examine, e.g., the implementing agency or a target population. Responsibility to a controlling institution or the need to lend support to specific interest group orient the identification to the recipients of the research (For Whom?). Identifying the “object” on which to focus (e.g., the program, the service) replies to the question What? and further asks “how does this change the status quo?” (Regonini, 2001). Finally, the reasons (as understood by the client) for making
the evaluation are summed up by the question Why?. The evaluation is either formative or summative (Scriven, 1991) depending on whether the object is to improve the program or to furnish a summary opinion on its specific aspects.

Policy Evaluation owes the clarity of its assumptions to its primary merit. Designing an evaluation based on this approach must be specified beforehand. In addition, Policy Evaluation makes a marked distinction between the policymakers (politicians and senior officials) and the analysts, drawing a clear line that reassures both parties: to the former, the division of tasks, to the latter the identification of the best means to realize them (ibidem, p. 235). Despite these indisputable merits, the most orthodox version of Policy Evaluation is the target of several theoretical proposals (grouped under the name of Policy Inquiry) that raise doubts on its assumptions. Policy Inquiry delivers numerous concepts, but here we will only cite a few common aspects, useful for our proposed discussion.

When and where. The evaluation can be conducted at any time (Browne and Wildavsky, 1983). However, the distinction between the ex ante, the in itinere, and the ex post evaluation can prove to be unrealistic when put to the factual test. Further, the analysis of a problem does not always precede the identification of the solution, but vice versa. For example, PA often decide to solve a problem because they believe they have the best solutions. Depending on the temporal horizon and the context examined, this implies that the effects of a program can overlap or be barely distinguishable.

For whom. As Regonini (2001, p. 382) observes, often “the problems to be dealt with have the power to attract actors” instead of vice versa. Social conditions are indispensable factors in determining the value of a public policy (Lippi, 2007). The results of the evaluation acquire meaning only in relation to the responses expected by the stakeholders. Given that each category has a different agenda (Gregory and Martin, 1994), the evaluation thus needs to unravel the following knots: which are the most important interests? Those of the politicians? Those of the beneficiaries of an intervention? Those of the agency paying for it?

What. Policy evaluation assumes that the purposes of the policies are a given and, therefore, that they can be easily deduced from official documents. Because of the typical verbal vagueness of legislative interventions (especially in Italy’s case) (Rebora, 1999), the question then becomes: must the evaluation restrict itself to analyzing the explicit objectives of the intervention or is it more appropriate to also analyze the unexpected effects and the unforeseen consequences? Is it possible – ask Pressman and Wildavsky (1984) – to express an opinion disregarding such effects?

Why. Rossi, Freeman and Lipsey (1999) underscore how hard it is to clearly define the evaluation goals, especially as the programs often produce the expected results in the longer term or in an undefined timeframe. We then need to consider that in most cases, the managers are unable to draw on information about current performance or on the methods to enhance it. The business literature also highlights the ambiguity of the notion of performance in the public sector. Therefore, any kind of measuring system is bound, fatally, to be put to the debate (Stewart and Walsh, 1994).

In brief, Policy Evaluation seeks mainly to identify the conditions that enable the adoption of the best solutions and the setting up of data-reporting systems to monitor the results. The main limitation of these studies is their inability to realistically capture the social dimensions of the policies, the plurality of interests, and the strategic games between “partisan actors” (Lindblom, 1959).

Policy Inquiry studies seek – unlike the linearity of the mainstream approach – to shed light on the social interactions triggered by the implementation of a public policy. These studies assume that any policy is a “polyhedron with a great many faces” (Regonini, 2001, p. 371). Consequently, the possibility of adopting a rational design approach to evaluation is denied (Guba and Lincoln, 1987) and that is obviously a limitation. Policy Inquiry considers evaluation as a process that creates “the reality” (ibidem, p. 141). It is precisely because of its unpredictability that implementation is called the “black box” of policymaking (Barrett and Fudge, 1981).

Basically, the clashing and polarized positions make it tough to interpret the concrete situations that the evaluation practice must address. Of course, each of these efforts has improved our understanding of policy implementation, but we now need to make the next leap forward. First, by “opening the black box”, that is, developing the tools needed to understand the causal nexus that should “enable the attribution of the effects to the realization of the program itself” (Stame, 2007, p. xiv). Second, by translating the theoretical proposals into operational terms, that is to “transform this learning into an appropriate intervention tool” (Regonini, 2001, p. 184).

In addition, our e-government focus means we need to also consider fundamental IS/IT evaluation studies (Smithson and Hirschheim, 1998; Ward et al, 1996; Jones and Hughes, 2001), which we will introduce into future research as we prepare the interview canvas for the collection of primary data.
THE MILAN ECOPASS PROJECT

The City of Milan (CM) introduced an entry charge (called the Ecopass) in July 2007 to restrict city center access (7.30am-7.30pm Mon-Fri) to the most polluting vehicles. The regulation (the first of its kind in Italy) came into force as a pilot test on January 2, 2008. Implementation of the new system – mandated to several municipal units (local police, public transportation company, IS Dept.) – led to numerous broad-spectrum interventions on the city’s traffic flows, ICT platforms, administrative procedures, staff training programs, and institutional communication.

To equip the regulated area (8.2sqkm or about 10% of Milan’s total urban landscape), CM installed 43 remote-control/CCTV access points. As a vehicle drives through these, the number plate is matched against information stored in the national motor registry database to determine the “pollution category” and the corresponding charge for the entry permit. A widespread service network for sales-transactions processing was created with newsagents, garages, and banks across Milan and the province to deal with the issue, payment, and renewal of Ecopass permits. Several changes were made to the information systems (IS), including the development of new application interfaces to connect to the vehicle database kept by the Ministry for Transport, the web platforms and the bricks and mortar info-points.

The Ecopass regulation (a pollution not a congestion charge like that implemented in London in 2006) sparked intense debate also among many experts. The discussion between the regulators and the detractors was particularly heated. For weeks, the various forums were a hive of bitter attacks, further aggravated by technical malfunctions. For instance, the now-fixed glitch in the automatic reading devices triggered the erroneous levying of some 160,000 fines. This culminated in work overloads for the call center, the municipal front-offices, the prefectures, and the Justices of the Peace.

Ecopass also envisions forms of citizen consultation (as announced on the municipal website) combined with the monitoring of the regulation’s environmental and socio-economic impact. Nevertheless, up to now the municipal Mobility Agency has limited itself to issuing solely periodic reports on traffic flows and the quality of the city center air. At the time of writing, the City had confirmed that it will continue to apply the regulation through 2009. This extension will also enable us to collect further in itinere evaluation data via interviews with various stakeholders.

DISCUSSION

An initiative as controversial as the Ecopass regulation places the evaluator in a tough position, specifically that of connecting “the context and the content of what is implemented” (Pressman and Wildavsky, 1984). One can form different opinions about evaluating such initiatives depending on the different cognitive positions taken (Lippi, 2007). The widespread opinion is that of common sense, which is fairly unhelpful when it comes to explaining the effects of a specific regulation. An example of the common sense evaluation (classified as level-zero) is the debate that involved the citizenry around the Ecopass project. This was the first time a regulation aimed to influence the behavior of a specific group (i.e. citizens who do not reside in the center of Milan and those who own an earlier generation of vehicle).

In addition to the common sense opinion, evaluation may use the expert opinion of those with direct field experience. If a stakeholder is a political body, the evaluation (classified as level-one) may reflect ideologically oriented interests. The Ecopass debate also drew opposition from many public institutions, while some citizen committees and various industry guild representatives (retailers, professionals, businesses based in Milan province) also voiced their needs. The opinions above level-one move outside the policymaking arena and are classified as level-two. This is the ideal position to achieve a dual goal, that is knowledge and accountability. CM indicated that it would verify the level to which Ecopass objectives had been achieved at the end of the pilot. However, no attempt has yet been made to conduct a level-two evaluation to the best of our knowledge and according to the press.

How were the evaluation issues described earlier dealt with in the Ecopass project? Adopting Lippi’s proposal, the opinions formulated by the diverse stakeholders can be clustered between level-zero and level-one. For example, in July 2008, Milan’s Traffic and Health councilors published a joint document that allegedly correlates the launch of the Ecopass with the downward trend of hospital admittances for pollution-related diseases among historic center residents using 2007 as the comparative basis. Other data disclosed by CM for January-September 2008 indicated a reduction in airborne micro-particles, fewer road accidents in the traffic-restricted area, and an increase in the number of public transport passengers.

Based on the Policy Evaluation frameworks discussion, CM wanted to show the effectiveness of the regulation by presenting limited data on its output (or the products of the activities implemented) and on its outcome (or the direct results on the recipients of the policies adopted). However, CM made no attempt to analyze the impact, the trickiest evaluation because it aims to determine whether the program has influenced the situation-problem in the way expected (Dente and Vecchi, 1999).
Policy Inquiry suggests that a delivery benchmarking logic prevailed in the Milan case as it favored a work-in-progress reporting of scattered information produced by the implementing administration itself. We refrain from discussing the scientific relevance of the indicators chosen to demonstrate the achievement of the Ecopass goals (CM chose the slogan “Less Traffic, Cleaner Air”). However, we note that the attempt to evaluate – kept completely within the policymaking arena – essentially boiled down to the selective documentation (i.e. “without evidence of lack of success”, Thompson, 1967, p. 93) of some non-contextualized results. Therefore, CM moved on the grounds Rebora (1999) has defined as “pre-evaluation methods”. Interestingly, between end-2008 and end-February 2009 the traffic-reduction trend reported immediately after the launch of Ecopass was reversed.

CM adopted a deterministic view of the regulation and its effects (expected and actual), excluding from the analysis the organizational processes related to program delivery. If we want to take a step forward in the conceptual clarification - we remind readers that we are discussing a pilot scheme, which by nature should allow the administration to “learn” about the impact of both the methods and the effects of the intervention - it might be useful to adopt the analytical tools offered by Organization theory. Organizational theory may help overcome some of the limitations mentioned above, as organizational knowledge is not antithetical to other disciplines (e.g., economics, political science, sociology, law) but, on the contrary, is reciprocally enriching. Neither should we forget that organizational theory presides over the logic that steers the goals and technical actions in organizational settings (Maggi, 2003).

We saw earlier how the prevailing theories used to address the policy evaluation are based on the opposing perceptions of absolute rationality, a priori or a posteriori. However, if we adopt theories that see the organization as a process of actions and decisions guided by intentional and bounded rationality (Simon, Thompson and Smithburg, 1991), the implementation of a regulation and the evaluation of its effects appear in a different light. In particular, the Theory of Organizational Action (Maggi, 1990; 2003) suggests approaching the evaluation in terms of: i) analysis of the choices of action that unfold over time; and ii) assessment of the relations between the organizational action plans, i.e. “structural, technical, and institutional”. The evaluation is then rationally bounded – because conditioned by the availability of knowledge and the (limited) capacity of the decision-makers to predict –, but also includes the evaluation of the ‘reciprocal congruity’ of the processes that occur at diverse levels. We are keen to continue our analysis at multiple points in time (in itinere and various ex-post) as it will help identify short-term and long-term implications.

The Ecopass project led to the development of new ICT platforms that enabled CM to fulfill the mandate of the regulation and supervise compliance with the 2007 resolution implementing EU antipollution legislation. The new ICT developments can be interpreted by looking at the regulating role of technical platforms in the organizational action. CM – through its IS Department – wanted to maintain strategic supervision of the design choices by directly controlling the ICT standards and architectures. In parallel, the decisions on the details were left to the discretion of actors outside the IS Department. The administration maintained its control powers by outsourcing the application development to a provider in which CM holds ownership shares (ATM, the city’s public transportation company).

We then need to consider the decisions on the adoption of the technology, i.e. those aimed at integrating the new ICT platforms into the existing work processes. We can easily infer (from the local press) that those adoption processes triggered a confrontation, and even conflict, between the policymakers, the IS developers, and the program executives. Time constraints were a major factor in these dynamics. For example, tight deadlines led CM’s IS Department to reduce the timeframe for transition to the new system. This ‘big bang’ approach is deemed the riskiest because irreversible (Alter, 1996). Mistakenly, CM expected to ensure coordination through plans and programs, assuming that simply following the process workflow would ensure the initiative’s success. Instead, Ecopass required a continuous coordination by mutual adjustment between the higher ranks and the operators, especially when it came to correcting malfunctions and software glitches, modifying the new e-services, or changing one-way streets and the relative traffic signs. Most of which were eventually remedied; therefore, we could say that the ICT infrastructure ultimately “held”.

Finally, we need to examine the decisions of use of the ICT artifacts. Management sought to make these decisions by investing substantially in staff training and external communication, interpretable as initiatives aimed at changing the decisional premises based on which the end-users choose to act. In particular, the operational training given to the users of the ICT artifacts helped to improve their technical rationality. The external communication can be considered a tool used to limit the recipients’ (employees and interested motorists) scope for self-regulation.

Retrospective reconstruction of the evaluation practice – only summarized here – highlights a sometimes tortuous unraveling of decisions and objectives, which changed partly along the way. The overall result cannot be understood without taking into account – alongside the constraints – also the opportunities inherent the new artifacts. Here we will only mention two. Primarily, the control system based on the 43 access points enables the Milan municipality to gather a significant quantity of information on vehicle circulation. This seems especially useful for designing future urban interventions as it provides a
cognitive base – which, in turn, can be analyzed through IS tools – that didn’t exist prior to the Ecopass. The second opportunity lies in the upgrading of the applications portfolio, which delivered some new front-end services to all the citizens via the municipal website.

The case of the Ecopass is an example of a process of change with a major impact on the citizens, where the implementation of ICT-based solutions, despite being an important objective, is not by itself capable of determining the results. ICT artifacts are organizationally relevant because they introduce new rules into the various processes (analytically identified) and impose them on others. The analysis of the secondary data sources does not yet enable us to identify how the constraints and opportunities - stemming from the choices of the design, adoption, and use of the e-services - have influenced the Ecopass outputs, outcomes and impacts. However, we can reasonably assume that the origins of the problems afflicting the project were also due to an underestimation of the organizational effort.

CONCLUSIONS AND IMPLICATIONS

This paper maintains the usefulness of adopting an approach to implementation evaluation that differs from research founded on the analysis of variance, such as output studies. It suggests using a processual approach to evaluation that may lead to useful heuristics. Policy studies help us to wipe the slate clean of the assumption that e-government is neutral. These recognize that the initial plan is prone to shifts and swings, so clearly we must also consider the results on the recipients (demand side) of the regulations, not just the regulations. On the other hand, organizational analysis can help shed light on the supply-side, i.e. on the processes of action and decision that led to specific effects.

What are the benefits of adding the organizational dimension to the evaluator’s conceptual toolbox? We cite only a few of the more noteworthy aspects that will guide our future research and analysis. First of all, focusing our attention on the bounded-rational processes of action and decision enables us to surpass the sterile opposition between the objectivist and the subjectivist approach and also factors in time as a key variable. We reiterate that it is the in itinere evaluation that offers the policymakers and the managers the most feedback to reduce uncertainty and correct or (re)orient the public action.

Secondly, shedding light on the processes helps increase the body of common knowledge available on the organization and its dealings with the environment. Enlightenment (Weiss, 1972), in turn, creates the conditions that favor external accountability. We point out that the evaluation generates knowledge that is evidence-based, not generic. That knowledge becomes a way to learn about the organizational interdependencies, the importance of which is underestimated by the current managerial discourse. The feedback from this learning can help change management’s decisional premises.

Thirdly, seeing the organization as a process instead of a reified context enables us to use the same conceptual tools also to interpret situations in which change goes beyond the “boundaries” of a specific institution. That aspect is significant given that policies are unlikely to target only one administration: see, e.g., the growing tendency in the associated management of locally based services (Sorrentino and Ferro, 2008).

We offer these considerations as food for thought. The paper’s biggest limitation is that of proposing only a preliminary tentative evaluation of the Ecopass implementation. A full evaluation study would have required an in-depth analysis of primary data in line with case-study research methods, which will be part of future research. A second limitation is the lack of discussion of this effort alongside the notion of “public value” and the notion of “common good,” an important evaluation perspective in the context of e-government initiatives. This aspect will be considered in a future review of the literature. We must remember that evaluation – like many other empirical research methods used by the social sciences – is not an exact science. The scientific basis lies in the procedure through which the estimative data has been constructed (Weiss, 1972).

REFERENCES