Institutionalizing Evaluations in Mexico: The Role of CONEVAL
Gonzalo Hernández Licona

Abstract

Perhaps the best well known topics related to impact evaluations are the ones associated to methodological issues. There are many technical books and papers showing the proper way of estimating impacts. However, dealing with technical aspects is only one part of the reality of impact evaluations. In this paper we sustain that in order to have more impact evaluations that are useful for improving social policy, is essential to change the rules of the game of public policy: We need to institutionalize evaluations.

Mexico launched in the past 10 years the General Law for Social Development, the Law on Budget and Fiscal Responsibility, the General Guidelines for the Evaluation of Federal Programs, it has created an autonomous institution, the National Council for the Evaluation of Social Policy (CONEVAL), whose objective is the evaluation of social programs and policies, and it has moved to be a more democratic country. All these institutional changes have contributed to build a Results Based Monitoring and Evaluation System (M&E).

Now impact evaluations are demanded, conducted, reported and disseminated within this M&E framework. We believe that we can have more results in terms of improving public policy and accountability if we insert impact evaluations within an M&E system, than if we only have isolated impact evaluations.

Institutionalizing evaluations not only has generated more evaluations, but the results have improved the decision making process within the Mexican Federal Government. There have been changes in the programs’ operations, design and implementation because of the information generated by impact evaluations. Furthermore, just recently the Executive’s budget decisions took into account the evaluations’ results.

There are still many challenges before suggesting that this one has been an extraordinary exercise. For instance, Mexico’s local governments still have a long way to go in terms of transparency and accountability. Nevertheless, with better technical instruments, but especially by institutionalizing evaluations, Mexico has taken important steps to improve the performance of social policy based on evidence.

Introduction

When I was first appointed General Director for Evaluation at the Ministry of Social Development in Mexico, I was sure that the most important step we should take was to brush up the team’s Impact Evaluation methodologies. For this matter, we were fortunate to hire very good teachers directly from the Institute for Fiscal Studies in London. After

---

1 I thank Edgar Martínez for his important contribution to this document and Marina Hernández, Janet Zamudio and Brenda Valdez for the translation of this paper.
some months of permanent training, our staff was technically prepared to conduct almost any impact evaluation at the Ministry.

We then made an appointment with the director of a certain social program. We explained him that we were going to evaluate his program with the latest and most rigorous methodologies. We briefly talked about the self-selection problems and we argued that the information from our evaluation would be very useful for decision making. He didn’t understand most of our technical explanations, but the most important part of our meeting was that he certainly did not want to be evaluated. He thought that it was a mistake to have external evaluations since the program staff had better knowledge about the program than we did. Therefore, we did not get far with that evaluation.

In fact, we did not get far in evaluating any program. Managers did not like the idea, and we did not have enough power to convince them about the beauties of rigorous evaluations. The main problem was that there were no incentives to evaluate social programs in Mexico; hence there was a need to change the institutional settings.

This paper focuses, first of all, on the institutional changes that were put into practice in Mexico in order to build a Monitoring and Evaluation (M&E) System, also, on how these measures helped the Federal Government to carry out evaluations and, finally, on how these evaluations have contributed to change decision making in the country. Our opinion is that despite important efforts made by the Mexican Government to invest in social policy, the reasons for delaying the construction of an M&E System were both, the lack of technical knowledge to conduct high-quality evaluations, especially impact evaluations, and also, the lack of an appropriate institutional structure. Although both obstacles are important, the second one implies changing the rules of the game which is the key factor to have a sustained M&E System.

The Evaluation of the Social Development Policy

I. Changing the rules of the game

It was not until 1997 that Congress became a plural political entity in Mexico. For almost 70 years, the Institutional Revolutionary Party (PRI) had a clear majority in Congress and
held the presidency on every election. During that period, democracy remained just an elusive word. Evaluations usually don't flourish on authoritarian soils.

It was not a coincidence that in the year 2000, when the opposition parties took over the majority in Congress, the Chamber of Deputies demanded external evaluations of all federal programs. This mandate took the Executive by surprise and, as it could be expected, those first evaluations were not good at all. But it did a great deal towards changing the culture about evaluation in Mexico.

Several years later, Congress took another important step. Although the quality of evaluations had improved, it was difficult to find good examples of studies that used adequate methodologies and, more importantly, usually the institutions in charge of carrying out the evaluations were hired with little transparency. For this reason, in 2005, Congress approved the General Law of Social Development (LGDS) aimed to institutionalize various practices of social policy, including evaluation. As a result, the National Council for the Evaluation of Social Development Policy (CONEVAL) was created.

CONEVAL began operations in 2006 as an institution with technical and administrative autonomy; its objective is to evaluate the social development policy, their programs, funds, and actions; and to measure poverty at national, state, and municipal levels. Coneval coordinates the evaluation of programs and policies executed by public dependencies in the social sphere.

Even though CONEVAL is part of the Executive and the president of its Governing Body is the Head of the Ministry of Social Development, the rest of its Board members are 6 independent academics democratically elected by: the 32 governments of the Federal Entities, the Chamber of Deputies, the Senate, the representatives of the municipalities and the Federal Government. The governance of CONEVAL was designed in order to have credible and objective evaluations, as well as solid poverty estimates.

It is important to acknowledge that perhaps the main reason to create CONEVAL was the lack of trust that Mexicans usually have in public institutions. In this case, Congress did not trust that the evaluations required since 2000 were objective and transparent. They believed, and sometimes they were right about it, that program operators issued terms of reference and hired evaluation teams with the main goal of finding only favorable aspects.
about the program. These evaluations were sent to Congress without anything happening in terms of policy making. Besides, it was difficult to find these evaluations because they were not public. The way CONEVAL was conceived by Congress, with a majority of its board members being academic citizens not appointed by the Executive, changed many of the problems faced by evaluations since 2000.

Almost simultaneously with the creation of CONEVAL, Congress passed the Law on Budget and Fiscal Responsibility (LPRH) issued in 2006 and renovated in 2007. Under this law, the Performance Evaluation System (SED) was created. The SED aims to evaluate public programs and projects to help strengthen a results-based management system. The institutions in charge of developing the SED are the Ministry of Finance (SHCP) and the Ministry of Audit (SFP).

This new law was finally the result of many attempts during the 90s to improve indicators and to measure programs’ performance. There were some previous results from that initiative, one of them was a budget targeted by sectors. There was also a set of indicators which were common to all programs. Unfortunately, these indicators were basically about the budget expenditure and not about the programs’ results. The 2007 Law changed all this.

CONEVAL and the Performance and Evaluations System are cornerstones for the evaluation of Federal programs and policies, especially in the social sphere, yet there is a significant challenge: the coordination between them. The LPRH gives SHCP and SFP broad powers for evaluation. But CONEVAL can only evaluate the social sphere, as indicated by the LGDS.

For this reason, between 2006 and 2007, taking advantage of the beginning of a new administration, an important exercise of institutional coordination took place between CONEVAL, SCHP and SFP. These normative steps helped once again to change the rules of the game in order to promote evaluations in the Federal Government. The following normative results were obtained:

a) In the 2007 Federal Budget Decree (which all Federal Ministries should comply with) published by Congress, it was established that the coordination of the evaluation of
federal programs would be in charge of CONEVAL, SCHP and SFP. For that coordination, the General Guidelines for the Evaluation of Federal Programs were to be issued within 90 days.

b) The General Guidelines for the Evaluation of Federal Programs were issued the last day of March 2007 with the participation of CONEVAL, SCHP and SFP. Those guidelines are the foundations of a Results-Based public management culture. To this end, they encouraged the following elements:

It was requested from all federal programs to have a Logical Framework (ML), in order to: improve the programs’ internal planning, be clear about the contribution of each program to national objectives and have better results indicators.

An adequate planning reflected on the ML can help to modify the Rules of Operation of the programs, improving public management. If the link between the National Developing Plan and the programs’ Rules of Operation is stronger, there is a better chance to accomplish the national objectives.

To have a more coordinated and homogeneous evaluations of programs and policies of the federal government, the Guidelines suggested an Annual Evaluation Program (PAE) which contains the most important evaluations to be done year by year. This is a jointly program which will be released annually by the 3 instances.

Impact Evaluations are mandated by the PAE and therefore programs cannot easily refuse to cooperate. However is worth noticing that the Annual Evaluation Program does not only mandate impact evaluations. It also mandates processes, design, consistency and other types of evaluations, which have different objectives than impact ones but still complement each other. Since by law all programs must have annual evaluations, it would have been very difficult to have impact evaluations for all of them.
Finally, another piece of legislation was launched to promote transparency in a country where various decisions and pieces of information were never known. In 2004 Congress approved the Law of Transparency and Access to Information. This law mandated all public institutions to disclose their information. Every piece of information has to be delivered to all citizens who ask for it. This law, along with the requirement to have external evaluations for all programs, makes it rather difficult for program managers to keep their performance hidden from public eye.

This set of new rules, starting with a more democratic country and ending with the Law of Transparency, have moved the Federal Government towards a Results-Based Evaluation and Monitoring System (M&E), which has become the basis for the Performance Evaluation System, created by the Law on Budget and Fiscal Responsibility (LPRH). The information generated by this system has been used by the general public and by policy makers for the last couple of years.

II. Constructing a Results Based Evaluation and Monitoring System²

As we said before, coordination of almost all public evaluations has been done based on the General Guidelines for the Evaluation of Federal Programs. Besides being a normative instrument, these precepts have set up the foundations for a Results Based Evaluation and Monitoring System for the Federal Government since 2007.

This M&E System includes three major pillars, as depicted in Figure 1:

a) Better planning tools (including important elements for monitoring),
b) An improved evaluation scheme, and

² For this paper and especially for this section, I followed the paper “Advances of Mexico’s Monitoring and Evaluation System: Three years after initiation”, 2009 mimeo, by Gonzalo Hernández Licona and Edgar Martínez Mendoza.
c) A process in which decisions are made taking into account the information provided by the M&E System.

As it can be seen in Figure 1, the main focus of the new M&E system is that it is based on Results. Therefore, the information driven out from evaluations and monitoring as well as the decision making process should emphasize results (as opposed to just producing budget indicators).

In this way, M&E is part of the planning process of public policy, as shown in Figure 2. The information provided by M&E should be useful to improve the design of public actions and to enhance accountability, which citizens are increasingly demanding. The scheme makes clear that the planning of social policy must begin with the identification and quantification of the economic and social problems. Once this is done, the reason why these problems exist should be analyzed. And finally, public policy actions such as the budgetary allocation and the operation of the proposed tools, should be implemented.

Figure 1.
Results-based Monitoring and Evaluation System

PLANNING

- Strategic goals
- Logical Framework
- Monitoring systems

EVALUATION

- Annual Evaluation Program
- Consistency Evaluation
- Impact Evaluation
- Design Evaluation
- Strategic Evaluation
- Complementary Evaluation

Results

Using the information from M&E

Performance evaluation (Annual report)

---

Figure 2

Planning process of the public policy

- Identify the problem
- Analysis
- Design programs
  - Budget
  - Operation
- Monitoring and Evaluation

Accountability
Finally, the diagram shows that the M&E System has three objectives: to have frequent information on the implementation of programs, to evaluate if the results are achieved and, finally, to reinforce accountability. The arrows on the left side indicate that the information coming from the evaluation system makes sense only if it is used to make concrete policy decisions, such as improving the design of programs and policies, assigning budget in a better way and improving programs operation.

**a) Planning and Monitoring**

CONEVAL has several actions in the area of monitoring, such as approving the indicators for results, products or services, and managing social programs. CONEVAL also serves as the main link between federal offices and entities for the establishment and consolidation of an M&E System; and for the provision of consultation and technical assistance.

At the federal level, various actions have been taken to improve the design and management of programs. An inter-institutional coordinating entity between CONEVAL, SHCP, SFP and the President’s Office has been established, to generate standardized criteria for the development of indicators of social programs. The logical framework methodology was incorporated in coordination with the Latin American and Caribbean Institute for Economic and Social Planning, a department of the Economic Commission for Latin America and the Caribbean, in order to clarify and improve the design of objectives of those programs operated by federal offices, and to enhance the creation of results-based indicators. Training, technical assistance, and consultation were provided to improve planning and development of indicators; a link between national planning and programs objectives of sector-based programs was established; the use of diagnostic assessments that identify problems that should be addressed by programs was reinforced, and recently a scheme for approving indicators was initiated.

In addition, these efforts have also been promoted more gradually at the state level, focusing on the following activities:

---

3 These faculties are specified in the Fifth Chapter of the General Law on Social Development (*Ley General de Desarrollo Social*), in CONEVAL’s by-laws and the decree that created CONEVAL. All three are available on CONEVAL’s web page: www.coneval.gob.mx

4 Since 2007 various training workshops have been provided, and agreements on technical cooperation have been signed between CONEVAL and specific states.
Consulting for the implementation of evaluation and monitoring schemes,
Consulting for the improvement of the design of sector-based programs and strategies in the field of social development, and for the use of tools to monitor their performance, and
Consulting for the analysis of social development.

In the last four years, since the system was initiated, a significant progress in monitoring has been reported. Currently, approximately 520 budgetary programs have a logical framework matrix, which lends greater clarity to the objectives pursued and allows the organization of indicators for their measurement. Moreover, many of the programs are not only focused on having objectives for results, but also reflect elements of measurement (30 percent of these indicators seek to measure results).

Of the 520 programs, CONEVAL is more focused on those with operational rules and on those that lack of these rules but that make social transfers (in budgetary terms these are identified under modalities S and U); a total of 130 programs, with 1,773 indicators (nearly 26 percent of them are indicators of results). In addition, and as an example of the use that has been given to these information, the Ministry of Finance (SHCP) selected 234 indicators for budget monitoring of which 62 percent are results indicators. These indicators were sent to Congress and disseminated in an official report. The number of indicators per level is indicated in Table 1.

### The importance of Results Indicators

Results indicators (labeled as Goal or Aim indicators in Table 1), are not only important for planning purposes, but also for measuring final results in a country without such a culture. As stated before, in the past the only indicators that were considered important were those related to the budget (percentage of the budget spent during and at the end of the year). This practice made sense since the financial plan was defined (still is the case) only focusing on the resources spent at the end of the year. For instance, if a program managed to spend the whole budget, there was a good chance that it would have received the same amount the following year, or even, an additional bonus for achieving that goal. On the other hand, a program which was unable to spend its budget would have been
penalized with less resources on the following year. Thus, there was no need to measure final results and therefore very few programs used to have results indicators.

Nowadays there is another reason why results indicators are important. Having almost 130 social programs, we understand that it is almost impossible to carry out impact evaluations for all of them. Actually, it would be a great success, for our M&E System, if we could have around 15 to 20 rigorous impact evaluations in a period of 6 years. Under these circumstances the majority of programs would not have an impact evaluation. But still, the Law demands to have monitoring and evaluation instruments for all of them. For this reason, having appropriate goal indicators (or results indicators, or outcome indicators, as some people label them) for all programs is our second best option.

We understand that results indicators (in opposition to impact evaluations) do not measure properly the causality between the intervention (program) and the problem that is pretended to solve. But this indicators, along with other types of shorter evaluations (as the ones we will mention further down), can help us to have a better idea of the program’s overall performance, than having only budget indicators.

Two more things are important to mention about results indicators. After several evaluation seminars that CONEVAL has conducted for Federal Government institutions (and also for states institutions), now some programs are willing to have rigorous impact evaluations, using their own resources and receiving technical advice from CONEVAL. Although building a suitable methodology for impact evaluation is not easy, the use of results indicators coming from the logical framework to measure impact has been taken into account.

Finally, a proportion of the 2010, and especially, of the 2011 Budget, was defined using information from the M&E System, which included not only the results from impact evaluations, but also from the results indicators and other evaluations. This set of information was very useful to change the way in which budget is distributed. Budget indicators are still important for deciding the allocation of resources, but information about the programs’ results is playing, gradually, a more important role.
Table 1: Indicators contained in 2010 MIRs

<table>
<thead>
<tr>
<th>Level of indicator</th>
<th>Indicators in PASH (Finance and Public Credit Secretariat information system)</th>
<th>Indicators S and U Programs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>%</td>
</tr>
<tr>
<td>Total</td>
<td>4,419</td>
<td>100</td>
</tr>
<tr>
<td>Goal</td>
<td>558</td>
<td>12.6</td>
</tr>
<tr>
<td>Aim</td>
<td>752</td>
<td>17.0</td>
</tr>
<tr>
<td>Components</td>
<td>1,415</td>
<td>32.0</td>
</tr>
<tr>
<td>Activities</td>
<td>1,694</td>
<td>38.3</td>
</tr>
</tbody>
</table>

Source: Developed by authors with information from the Portal Aplicativo de la Secretaría de Hacienda (PASH).

Inter-institutional measures have also been taken into account to include efforts at the three levels of government. Cascade matrixes have been designed to integrate programs operated through the states and the federal public administration offices in order to link results obtained at the sector level with the National Development Plan. To that end, administrative records and measurements of results have been enhanced thereby providing important information for decision-making.

A common practice in several states, such as Chiapas, Puebla, Jalisco and Guerrero, has been to generate a matrix of indicators that can be reproduced for all programs. Nevertheless, there have also been cases in which program planning has been aligned with policy strategies integrating results indicators.

b) Evaluation

During the 90s, very few evaluations took place in Mexico, despite some efforts of building an evaluation system at the end of the 80s. Perhaps the most important one was the well known evaluation of Progresa-Oportunidades. This evaluation was conducted by very professional policy makers at the Executive. We can say that this evaluation has been a permanent example for the rest of the programs. However, during the 90s, especially at the beginning of this period, almost no impact evaluations were done.
As we mentioned before, on the year 2000 Congress demanded to have external evaluations for all federal programs. Despite the institutional importance of this event, the evaluations that were carried out suffered from several problems: the terms of reference were both very general and ambitious, they requested to measure impact for all programs every year; the external evaluators (usually researchers from universities and some consultants with relatively low experience in evaluation) were not prepared to handle so many evaluations; the programs had the freedom to choose any evaluator they wanted at almost any price, and therefore there was a flaw on the objectivity and quality of the evaluations. Although these evaluations were handed in to Congress and the Ministry of Budget (SHCP), there was not a clear mechanism to change policy with the information generated by the evaluations.

For this reason, the General Law on Social Development assigned to CONEVAL responsibility for standardizing and coordinating evaluations of social programs operated by federal offices. To fulfill this responsibility, CONEVAL, in coordination with the SHCP and SFP, issued standards for regulating the evaluations such as the General Evaluation Guidelines which we already mentioned.

The Guidelines point out that CONEVAL along with SHCP and SFP should issue every year an Annual Evaluation Plan (PAE). Similar to what happens in Chile, the idea of this PAE is to organize systematically the evaluation needs for the Federal Government every year (see Figure 1).

In order to establish the PAE, the following steps are taken every year. First of all CONEVAL, SHCP and SFP establish the priority areas and programs of public policy which should be evaluated in a certain year. At the same time, a consultation with the main ministries is carried out in order to understand their evaluation needs. The proposed programs and policies to be evaluated are then sent to Congress along with the Federal Budget Plan for next year. Congress then approves both the budget and the PAE.

The idea of the evaluation scheme is to have different types of evaluations, since we believe that we cannot understand the programs’ performance with only one type of evaluations. For this purpose, CONEVAL issues the terms of reference for different types of evaluations, giving some degree of freedom for the impact evaluations and for those evaluations undertaken by other institutions (complementary evaluations).
So far there have been five Annual Evaluation Plans: 2007, 2008, 2009 and 2010 which can be seen on www.coneval.gob.mx. As Figure 1 can show, the main evaluation instruments are the following:

i) Consistency and Results Evaluations.

The main question this evaluation answers is: To what extent social programs are oriented towards results, based on their own documentation and processes? This instrument is a diagnosis about the institutional and management capacity of the programs to achieve results. The idea was to have a homogeneous instrument, with measurable answers, in order to compare performance among programs. This evaluation was like a baseline for the programs’ achievements, which help the new Government to see what type of tools they had at the beginning of their term.

The Terms of reference were elaborated by CONEVAL (www.coneval.gob.mx). The evaluation has 100 questions about the following topics: Design, Strategic Planning, Target Population and Coverage, Operation, Perception of Beneficiaries and Final Results. The Design topic is mainly based on the Logical Framework. Of those questions, 80 have a binary answer (YES/NO). The evaluations were done by external institutions (universities, consultants, experts), hired by the ministries, but coordinated by CONEVAL.

The evaluation took place between 2007 and 2008. In March 2008, 106 evaluations were delivered to Congress and the Executive, just as the Annual Evaluation Plan had stated. The next evaluation will take place in 2011-2010 to see how programs do in terms of working toward results. The evaluations can also be found at www.coneval.gob.mx.

ii) Design Evaluations.

Design evaluations were intended for new programs. They seek to evaluate the relevance and the overall internal logic of the program. The idea was to help decision makers and, especially operators, to see if new programs had been implemented in the way they were supposed to. A special topic here is to ask if the program has clearly identified what type of
problem it is aimed to solve and if it can achieve its goals with that specific design. Reference Terms for these evaluations are also available at the CONEVAL page. The results of the evaluations for the first 13 programs were delivered to the Congress on October 2007.

iii) Impact Evaluations

As we know, this type of evaluation is perhaps the best way to measure the contribution of a certain program to the solution of a socioeconomic problem. Rigorous Impact evaluations can measure causality between the intervention (program) and a certain socioeconomic problem.

CONEVAL understands the importance of impact evaluations, but it also understands that a) For budget reasons we cannot have impact evaluations for all programs in Mexico (there are around 130 social programs and almost 550 programs in total in the Federal Government), b) For technical reasons sometimes is difficult to do impact evaluations (lack of information, lack of qualified evaluators and suitable methodologies, and c) Other evaluation instruments are needed in order to complement impact evaluations.

Thus CONEVAL is not only doing or coordinating impact evaluations; the idea, as expressed in Figure 1, is that impact evaluations should be part of the country’s Results Based Monitoring and Evaluation System. We believe that we can have more results in terms of accountability and improving public policy if we insert impact evaluations within an M&E system, than if we only have isolated impact evaluations.

Having a single impact evaluation, as was the case of the evaluation of Progresa-Oportunidades in the 90s, is good news. But having a whole M&E system, we believe is even a better objective in terms of evaluation efforts for a country in which just recently, transparency and democracy were elusive.

Of course the rigorous impact evaluation of the program Progresa-Oportunidades made a remarkable difference on the country’s evaluation culture since 1997. Fortunately nowadays there are evaluations with similar rigor for other social programs.

The enthusiasm for the evaluation of Progresa – Oportunidades in 2000, but perhaps more importantly the institutional changes made in the evaluation culture since 2000, has
increased the demand for impact evaluations in Mexico. Almost 41 impact evaluations with a high degree of rigor (by CONEVAL standards) have been done for the social policy in the past 8 years (10 of them were done for Progresa – Oportunidades). Annex 1 shows a summary of various impact evaluations carried out recently, in which we can see the method used, who financed the evaluation, the main challenges faced during the evaluation, the main results and the main use of the information so far. We will talk about this last issue further down.

iv) Strategic Evaluations.

Analysis of isolated programs only shows one part of the reality of public policy. Thus, CONEVAL has also conducted evaluations of concrete policies that go beyond the evaluation of programs. In 2009, CONEVAL finished the Evaluation of the Nutrition and Food Supply Policy, which analyses if the Federal Government’s current tools (programs, regulations and actions) are adequate to meet the new challenges the country faces in terms of nutrition (for example: under-nutrition has declined on average in the past 20 years, but the levels of under-nutrition for indigenous children is more than double than the country levels; obesity has gained space in Mexico for the past 20 years; in 2006 more than 30 per cent of women between 24 and 49 years old were considered obese)

A similar evaluation was conducted to analyze the Social Funds delivered to states and municipalities (Fondos del Ramo 33)\(^5\). The evaluation, which was completed in 2010, analyses the results and processes of the funds the Federal Government assigns every year to states and municipalities to improve social policy. Both evaluations can be found in www.coneval.gob.mx

v) Complementary Evaluations

CONEVAL, SHCP and SFP are coordinating the evaluation process in Mexico. However, it is also true that sometimes programs and institutions are willing to carry out evaluations in order to answer important questions about their own performance. With this in mind, we classified as Complementary Evaluations the evaluations done directly by programs and

\(^5\) Ramo 33 is a Federal Budget Fund, which stands for Federal Contributions for Federal States and Municipalities
institution, but they have to be part of the M&E System. In some cases, especially when these studies are impact evaluations, CONEVAL offers technical assistance.

c) Using the Information from M&E

It is not straightforward that as soon as the evaluations are finished, they are used to improve public policy or to support accountability. For this reason it is important to have a specific mechanism for the use of information provided by the M&E system.

In 2008, CONEVAL designed a mechanism to incentive the use of the information generated by evaluations. This mechanism implied that all programs have a public document – Work Plan - with actions and deadlines, in which they state how they are going to solve the main challenges reported by the evaluations.

For example, almost 80 percent of programs have defined aspects that are susceptible to improvement and have established a work program to address these aspects. There are complete diagnostic assessments for over 100 programs in the areas of design, planning, coverage, focalization, operation, perception and results.

In its commitment to transparency and accountability, CONEVAL has worked to ensure that all evaluated programs must present the evaluations on their web pages, with data from the hiring processes. Civil society organizations have used these inputs in their independent monitoring of government actions.

In addition to the recommendations made to specific federal programs, CONEVAL has published an important document in which we assess the social development policy as a whole and make general recommendations to various institutions. Some of these document's recommendations come from the general recommendations that are analyzed in the mechanism outlined before:

The Evaluation Report of Social Development Policy of 2008 (a similar documents is ready for 2009) summarizes the findings that CONEVAL has realized over various aspects of
Social Development Policy. The Report is the first update diagnostic on the status of the Social Development Policy. The report provides a macro analysis of social variables, the evolution of poverty, and the results of federal social programs. It seeks to analyze the results and evolution of social development in various dimensions, in order to contribute to the transparency and accountability of the federal government and to make recommendations for the improvement of the social development policy.

In order to improve the decision-making process, all evaluations and reports are delivered to main public policy actors:

- Congress
- The National Commission for Social Development (social development ministries of all federal states, representatives from Municipalities, from Congress and from the federal government)
- The Inter-Secretarial Commission for Social Development (represented by 10 Federal Ministries related to social topics)
- The President’s Office
- The Ministry of Finance (SHCP)
- Federal offices and the directors of their programs

d) Performance Evaluations (Annual Report)

The information from the components of Figure 1 (Planning, Evaluations and the Mechanism to follow up on recommendations) is summarized every year in a Report named Performance Evaluation. This is a 6 page report for each program, with the purpose of monitoring the program’s progress in reaching goals and objectives, and as a summarizing instrument—bringing together all the information coming from the M&E system, including the main results from all kinds of evaluations, indicators, goals, coverage and the degree of improvement the program has had on its Working Plan of previous years.
These reports are handed to policy makers at Congress and at the Executive and are placed, as all evaluation instruments, in CONEVAL’s web page. This is a very useful instrument because otherwise it would be rather difficult that policy makers can have the time to read and analyze the information generated from the M&E System (every program has a logical framework with almost 25 indicators, some of them are reported on a quarterly basis; there are many evaluations for each program with almost 300 hundred pages in most of them, etc.).

This short Performance Evaluation links directly to the impact evaluations done for the program in previous years in the following way: the first page of this report summarizes the main results coming from the program’s impact evaluations. If the program does not have any (rigorous) impact evaluation, the report says so (“This program does not have yet an impact evaluation”). Of course by stressing the importance of impact evaluations, and because these reports are public, there has been a great demand for impact evaluations. However, as we said before, it is sometimes difficult for technical or budget reasons to have impact evaluation in all programs. Because of this, and in order to be fair with most programs, the lower part of this first page also summarizes the evolution of the program’s results indicators.

III. Results from the evaluation process.

For the past three years CONEVAL, together with the newly created SED has provided information for decision makers in both the Executive and Congress. Here we present a summary of the use of the information from Mexico’s M&E System and its disseminating process.
a) Policy changes

We mentioned before that the follow up mechanism on the evaluations’ recommendations was an important tool to actually use evaluations for changing public policy. These changes not only take place within programs, but the evidence is also used to improve other programs or practices.

The following cases are examples of how public policy has been modified when the information from impact evaluations was used:

**Progresa – Oportunidades Program.** The impact of this intervention is measured by improvements in nutrition and health condition, and school enrollment. In response to the recommendations, the program modified the food supplement formula in its nutritional component, expanded coverage to urban areas, and expanded coverage to include senior high schools in its educational component. Due to recent evaluations showing that the impact in urban areas is significantly lower than in rural areas, the program changed the parameters for urban areas in 2010.

**Food Program:** Results were monitored in growth, diet quality, and household consumption for the target population. The evaluation measured impact with two interventions: in cash or in kind. With the recommendations, it was possible to improve the basic food basket consumed, and to promote monetary transfers in poorer areas. The transfer with cash was relatively cheaper than the in kind version.

**Social Milk Supply Program:** Results were reported in the incidence of anemia in children from birth through two years of age showing a 25 percent reduction. Following these results, this feature has been promoted in other programs. Now the Rural Supply Program, which has low-price shops in poor rural areas, is selling this type of milk. Furthermore, the fortified milk implemented by this program is sold now in private shops with a higher price; this may benefit all children, not only the poor ones, and helps the program to increase its own resources. No other milk has this type of iron which children can easily absorb to reduce anemia.

---

6 Annex 1 shows a summary of the most important characteristics of various impact evaluations, including their use for public policy.
**Piso Firme Program**: This program has shown progress in a number of areas. Because results are monitored, household illnesses, such as parasites, anemia and diarrhea, have declined, and children’s cognitive capacities have improved. In response to these results, federal and local governments have channeled more resources to this type of program. In the past 5 years the budget for these type of actions has increased around 80%.

**Micro-regions Program**: The evaluations presented before had one common factor: they were all demanded by program managers. The evaluation for the Micro-regions program was not the case. Here the new rules of the game in terms of evaluations demanded to have an external impact evaluation of this program, whose main objective was to improve the living standards in remote rural areas. The idea was to identify micro-regions in the country (that is, poor regions with common characteristics), and then coordinate all ministries to improve infrastructure for basic needs (water, electricity, etc.) within this regions.

Managers did not want to have this evaluation done; however, it was performed with the help from international evaluators. The evaluation estimated no impact of this program on the access to basic needs, except for the use of computing rooms (with internet access) in poor rural areas. Due to this information the program stopped in 2007 and another program is now in place with a different approach. The new idea is to reach only the poorest 125 municipalities in the country.

Apart from the changes experienced in programs due to their evaluations, a bigger step in terms of using the information from the M&E System has come recently from the Ministry of Finance (SHCP). Starting in 2009, SHCP authorities have been using evidence from the System to make strategic decisions for the following year’s budget. This step has not been easy neither in Mexico, nor in all countries that have been engaged in this sort of decisions. Although the way of constructing the largest part of the budget is still old fashion (according to political negotiations and using budget indicators) the new institutional changes along with the existence of evaluation instruments imply that gradually these decisions will be based on results.
b) Dissemination

Figure 2 shows that one of the main purposes of an M&E system is to have accountability in public policy. Past evaluations have not had great publicity. Nowadays guidelines have brought great dissemination for the evaluations on the government agencies’ websites. However, the dissemination of evaluations and monitoring tools is not only for accountability purposes. If we want to have impact on decision makers, dissemination is essential.

Now, all evaluations coordinated by CONEVAL are on its web page, as well as in the programs’ and institutions’ web pages. CONEVAL’s website has been translated to English as well. We hope the dissemination strategy can reach many readers: citizens, policy makers, researchers and all those interested in evaluation tools.

c) Greater interest of Congress on Evaluation topics.

Congress has been interested in evaluations for several years, although their use has been limited. Today the relationship between CONEVAL and Congress is closer. Workshops about Logical Framework were conducted to the Federal Audit (ASF), an institution from Congress. The instrument is important because CONEVAL is on the process of approving the programs’ indicators and the ASF together with the SHCP must make comments on them. The communication between CONEVAL, ASF and the Congress is easier since the indicators are generated through the Logical Framework instrument. Still an important challenge is that Congress should use more the M&E System in order to make decisions.

d) Monitoring and Evaluation statewide.

The Law of Budget and Fiscal Responsibility of 2007 requested local governments –states and municipalities- to realize evaluations and to significantly improve the indicators of the funds that they receive from the federal government. In order to strengthen the practice of evaluations in the country, CONEVAL has had contact with all federal states. Almost 10
states (out of 32) are now using the Logical Framework for better planning, and at least 5 states have now external evaluations.

There is some enthusiasm with the idea of creating new results indicators for these funds. However, taking evaluations and transparency to local authorities is still an enormous challenge in Mexico.

Challenges

As it can be observed, during the last years several institutional changes have been put in practice in Mexico in terms of evaluation and transparency. These mandates require agencies and entities to develop indicators, to carry out evaluations, to implement tools that support results-based budgeting, to create a Performance Evaluation System, and to incorporate the States into this new approach to public administration. Even though these measures have helped to create a suitable environment for the improvement of public policies, the implementation of such elements brings medium and long term challenges that are necessary to institutionalize the creation of a monitoring and evaluation system.

a) Institutional Coordination

Among the main coordination challenges of the institutionalization of monitoring and evaluation schemes is to establish a stronger link between policy-making agents. In order to achieve this, leaders with enough authority, with means to generate interconnections between high level decision making and operational performance, and with ability to recognize incentives should be identified. These chains can be related at different levels: global institutions (the Ministry of Finance (SHCP), the Ministry of Audit (SFP) and CONEVAL; governmental dependencies (budget, evaluation, planning and operative departments as well as entities related to the reference sector); budget institutions (Congress, SHCP and budgetary departments in each dependency), and finally, institutions for auditing or control (Federal Superior Audit Office (ASF), SFP, Internal Control Organs and Public Commissioners).
Another major challenge is to promote a more active participation of Congress, through the various committees and research centers, as well as the President’s Office, in the use of results driven out from programs and policies evaluations. These relationships should encourage the agreement of “shared results” and should promote greater coordination in programs in which more than one institution or sector is involved.

b) Strategic Planning

Strategic planning is a key factor in building a monitoring and evaluation system. A closer link between national, sector-based, and institutional planning is an essential element in the evaluation of public programs and policies since it helps to clarify the variables that are being measured.

This component suggests that the institutions’ and programs’ objectives should be adjusted in order to agree with the National Development Plan goals. This has never been done before because development plans tend to be static.

It is also needed to link planning and budgeting in order to strengthen the impact of public programs by taking into consideration all budgetary adjustments that might be needed such as coverage extensions, improvement of formulas or food baskets, modifications to the financing schemes, etc.

An element, so far forgotten, at different planning levels, is the relationship between governmental dependencies and entities, since there is often a leap from sector-based programs to budgetary programs. Improving this chain will help to clarify programs objectives and expected results, construct better indicators to measure performance, and identify realistic goals based on financial, material and human resources.

c) Monitoring and Evaluation

Monitoring and evaluating the performance of public policies and programs, implies learning from experience and making adjustments to the design, implementation and allocation of resources of them. This can be quite difficult since it involves political
decisions such as merging, removing or redesigning programs along with the costs they carry.

In spite of all the improvements that have been carried out in the field, important challenges are still on the radar. Those include identifying key users of the results from these evaluations; increasing its use for planning, designing, and budgeting; enhancing its dissemination and socialization within the citizens; integrating program evaluations with states’ guidelines; and reviewing laws that are related to evaluators’ contracts with the aim of ensuring greater flexibility, clarity and efficiency in the process.

It is also necessary to identify which departments are responsible of keeping track and controlling the implementation of recommendations from both the monitoring and the evaluation of public programs and policies.

Also, it is very important to be very careful when information is generated to assure that relevant variables are reliable. For this reason, technical qualifications of program operators and of planning and evaluation departments’ staff should be strengthen; internal capabilities to analyze fundamental (and not only budgetary) information of the programs should be improved; administrative records that support the monitoring and external evaluation should be enhanced and standardized; and synergies between statistical committees and the National Institute of Statics and Geography (INEGI) should be exploited.

**Conclusions**

Evaluating private investments is an every-day activity. It should be this way, since private investors are constantly asking about their economic returns. Although the evaluation of public programs is as important (or more) as evaluating private investments, M&E systems haven’t been running for many years. In the case of Mexico, it took various decades to build an M&E system. We are still working on it.

Using better technical instruments, but especially by changing the rules of the game and institutionalizing evaluations, Mexico has made great progress in taking certain public decisions based on results. In relatively few years, the Federal Government has created
an evaluation system for the social development policy. This evaluation system, which includes impact evaluations, has increased the knowledge on how to improve the public performance based on evidence, and has improved the decision making process.

There are still some important challenges to face: improving the coordination between all actors involved in the evaluation; improving the strategic planning of the Federal Government; strengthening the external evaluators’ market; improving the technical capacity of those in charge of the evaluation within the institutions. Perhaps the biggest challenge is making mandatory the evaluation in local government by changing their own rules of the game.

By changing the rules of the game in terms of evaluation and transparency, at the end of the day we have been able to evaluate almost 140 programs in the past 5 years. Some of their manager may still refuse to be evaluated, but the new institutional settings favors cooperation. Our technical skills on evaluation are finally paying off.
ANNEX 1

SUMMARY OF SOME IMPACT EVALUATIONS IN MEXICO

Progresa – Oportunidades Program

- **Method**: Randomization
- **Evaluator**: IFPRI
- **Key elements**: Program and evaluation directed by the Ministry of Finance; Academics behind the program; Centralized program; Evaluation designed along with the program;
- **Challenges**: Political pressure to include control localities in advance; now it’s difficult to find control groups
- **Demand**: From those designing the program (help from IDB)
- **Financial support**: From the Ministry of Finance
- **Results**: Increase in nutrition, school enrollment and health, especially in rural areas. Lower impact on urban areas. The iron used in food supplements was not absorbed initially by children.
- **Use**: The program is still alive; the program expanded to urban areas and to high school; the formulae of the food supplement changed due to the evaluation results. The program’s parameters for urban areas changed in 2010.

Milk Program

- **Intervention**: Milk with iron, Zinc and vitamin C offered to poor families. Normal milk was offered to poor families since 1945.
- **Method**: Matching. Fortified milk was not offered at the same time in all States
- **Evaluator**: INSP (Mexico)
- **Key elements**: The director of the program favoured the evaluation; Academics at Sedesol decided to have an impact evaluation; Centralized program;
- **Challenges**: None
- **Demand**: From the director of the program; from Sedesol authorities; “we want to show that the program is working”
- **Financial support**: From Sedesol and from the program
- **Results**: 25% reduction of anemia in 0-2 years old children
- **Use**: Now the fortified milk is used in other programs;
Food Program

- **Intervention:** Food support to poor families in rural areas (poorer families than Progresa families)
- **Method:** Randomization of localities: controls, food basket and nutritional training, food basket without training, cash transfers
- **Evaluator:** INSP (Mexico)
- **Key elements:** Evaluation decided by Sedesol authorities, before the arrival of program operators; Academics behind the program; Centralized program; Evaluation designed along with the program;
- **Challenges:** Political pressure to include control localities; difficult, but not impossible, to convince program operators afterwards
- **Demand:** From those designing the program (Sedesol authorities)
- **Financial support:** From Sedesol
- **Results:** positive impact on growth, on dietary quality, and on household consumption; dietary quality was most improved in the Food Basket groups; costs were lower using cash.
- **Use:** Cash transfers in remote localities; food baskets improved (with fortified milk)

Cement Floor Program

- **Intervention:** The State of Coahuila replaced dirt floors with cement in almost all poor families’ houses.
- **Method:** Regression discontinuity and matching methods. The urban area of La Laguna is shared between the State of Coahuila and the State of Durango (almost a natural experiment)
- **Evaluator:** Berkeley University
- **Key elements:** The State Governor wanted to show that his main program was a success; Sedesol was doing another evaluation in the same areas;
- **Challenges:** Finding good data from both States
- **Demand:** Directly from the State Governor; from Sedesol authorities;
- **Financial support:** From the State and partially by Sedesol
- **Results:** On child health: reduced parasites, diarrhea, anemia, increased cognitive ability; on mothers: increased satisfaction with quality of life;
- **Use:** Federal government and other states are increasing the budget for this type of programs
Micro-regions Program

- **Intervention**: Basic infrastructure in poor municipalities
- **Method**: Regression discontinuity.
- **Evaluator**: Tecnológico de Monterrey (Mexico)
- **Key elements**: Support from good academics from abroad (Mathematica/Boston) and from Mexico
- **Challenges**: Finding good controls; Directors and operators of the program did not want the evaluation
- **Demand**: Sedesol (Undersecretary for Planning and Evaluation) wanted to know if this important program had any impact;
- **Financial support**: Sedesol and Conacyt
- **Results**: Almost no impact
- **Use**: The program is over, replaced now by another territorial approach

Housing Program (*Tu Casa*)

- **Intervention**: New houses or resources for housing improvement.
- **Method**: Comparison between beneficiaries with different exposure to the program.
- **Evaluator**: Berkeley University
- **Key elements**: Evaluators had the ability to change methodology when the first attempt was not possible; the administrative records were always a problem.
- **Challenges**: Finding a suitable method; attrition; the main authorities of the program changed during the evaluation; less support from them.
- **Demand**: Sedesol (Undersecretary for Planning and Evaluation) wanted to know if this important program had any impact;
- **Financial support**: Sedesol and Conacyt, IADB
- **Results**: Problems with final estimation, but an important finding was that many new houses were empty due to lack of resources in new locations.
- **Use**: Almost none
References


