Mexico
Analysis of the Information, Monitoring, and Evaluation Guidelines of the Programs in the Public Federal Administration

April 2009

Poverty and Gender Unit
Poverty Reduction and Economic Management
Latin America and the Caribbean Region
### ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>APF</td>
<td>Federal Public Administration</td>
</tr>
<tr>
<td>CEPAL</td>
<td>Economic Commission for Latin America</td>
</tr>
<tr>
<td>CONAPO</td>
<td>National Council for Population</td>
</tr>
<tr>
<td>CONEVAL</td>
<td>National Evaluation Council of the Policy for Social Development</td>
</tr>
<tr>
<td>DGPOP</td>
<td>General Management of Programming, Organization and Budget</td>
</tr>
<tr>
<td>Diconsa</td>
<td>Commercial Supplier and Driver – Conasupo</td>
</tr>
<tr>
<td>DOF</td>
<td>Official Gazette of the Federation</td>
</tr>
<tr>
<td>DQAF</td>
<td>Data Quality Analysis Framework</td>
</tr>
<tr>
<td>ECR</td>
<td>Consistency and Results Assessment</td>
</tr>
<tr>
<td>HGM</td>
<td>Tools for the Management of Microdata</td>
</tr>
<tr>
<td>ILPES</td>
<td>Latin-American Institute for Economic and Social Planning</td>
</tr>
<tr>
<td>IMSS</td>
<td>Social Security National Institute</td>
</tr>
<tr>
<td>INEGI</td>
<td>Geographical Statistics and Information Technology National Institute</td>
</tr>
<tr>
<td>INMUJERES</td>
<td>National Institute for Women</td>
</tr>
<tr>
<td>MI</td>
<td>Indicator Matrix</td>
</tr>
<tr>
<td>MIDO</td>
<td>Integral Model of Performance of Surveillance and Control Entities</td>
</tr>
<tr>
<td>M&amp;E</td>
<td>Monitoring and Evaluation</td>
</tr>
<tr>
<td>MML</td>
<td>Logical Framework Matrix</td>
</tr>
<tr>
<td>OMB</td>
<td>Office of Management and Budget</td>
</tr>
<tr>
<td>PART</td>
<td>Program Assessment Rating Tool</td>
</tr>
<tr>
<td>PBR</td>
<td>Budget based on Results</td>
</tr>
<tr>
<td>PEF</td>
<td>Expenditures Budget of the Federation</td>
</tr>
<tr>
<td>PMG</td>
<td>Modernisation Management Program</td>
</tr>
<tr>
<td>PND</td>
<td>National Development Plan</td>
</tr>
<tr>
<td>Pp</td>
<td>Budget Schedule</td>
</tr>
<tr>
<td>ROFP</td>
<td>Operating Rules for Federal Programs</td>
</tr>
<tr>
<td>SAGARPA</td>
<td>Ministry of Agriculture, Livestock, Rural Development, Fisheries and Food</td>
</tr>
<tr>
<td>SED</td>
<td>Performance Evaluation System</td>
</tr>
<tr>
<td>SEDESOL</td>
<td>Ministry for Social Development</td>
</tr>
<tr>
<td>SEMARNAT</td>
<td>Ministry for the Environment and Natural Resources</td>
</tr>
<tr>
<td>SEMBR</td>
<td>Evaluation and Monitoring System Based on Results</td>
</tr>
<tr>
<td>SFP</td>
<td>Ministry for Public Function</td>
</tr>
<tr>
<td>SHCP</td>
<td>Ministry of the Treasury and Public Credit</td>
</tr>
<tr>
<td>SIGOB</td>
<td>Programming and Management System</td>
</tr>
<tr>
<td>SIMER</td>
<td>Results Monitoring and Evaluation System</td>
</tr>
<tr>
<td>SMyE</td>
<td>Monitoring and Evaluation System</td>
</tr>
<tr>
<td>SNEIG</td>
<td>Statistics and Geographical Information National Systems</td>
</tr>
<tr>
<td>TDR</td>
<td>Terms of Reference</td>
</tr>
<tr>
<td>UR</td>
<td>Operations Program Responsible Unit</td>
</tr>
</tbody>
</table>

**Vice President:** Pamela Cox  
**Country Director:** Axel van Trotsenburg  
**Sector Director:** Marcelo Giugale  
**Manager:** Jaime Staevsda  
**Lead Economist:** David Rosenblatt  
**Task Leader:** Gladys López-Acevedo
ACKNOWLEDGEMENTS

This study was carried out by a team led by Gladys López-Acevedo. The team was integrated by Manuel Fernando Castro Quiroz, Ray Rist, Alejandro Medina Giopp, Gilberto Moncada Vigo, Gita Beker Busjeet, Lucy Bravo, Jonna Lundwall, Irene Clavijo and Rosa María Hernández-Fernández.

This task was carried out with the valuable information and comments from several Government Officials of CONEVAL: Gonzalo Hernández, Executive Secretary; Edgar Martínez, Coordination General Director; Gabriela Pérez, General Director of Evaluation; and Shakirah Cossens, Social Programs Monitoring Director.

We also express our thanks for the participation of the advisers of CONEVAL, and other persons and teams of the Ministries interviewed, specially the Ministry for Social Development (SEDESOL), the Ministry for Public Function (SFP), the Ministry of the Treasury and Public Credit (SHCP), the Ministry for Public Education (SEP), the Ministry of Agriculture, Livestock, Rural Development, Fisheries and Food (SAGARPA), and the Ministry for the Environment and Natural Resources (SEMARNAT).

We finally thank the comments and suggestions of the external advisers who were consulted during the preparation of this document.
INTRODUCTION

This collection of works represents the results of a “Fee for Services” contract with the National Evaluation Council of the Policy for Social Development in Mexico (CONEVAL), enforced between 2007 and 2008. The goal of the job was to provide support to CONEVAL in the strengthening of the General Guidelines of Evaluation of the Federal Programmers of the Federal Public Administration (Guidelines) that were published in March 2007 by CONEVAL itself, the Ministry of the Treasury and Public Credit and the Public Function. It was for such a purpose that the team of the World Bank worked on four different reports that looked at providing an integral analysis of the Guidelines. The four reports include: an executive summary, an institutional analysis of the Guidelines, a management tools analysis of the Guidelines and an analysis of the basic information within the context of the Guidelines.

The Guidelines were published to consolidate the vision of the Mexican government to create a 'menu of evaluations' aimed at providing different key users within the public administration with performance information that would serve their different needs. It is within this context that the analysis had as its general objective to analyze the results of the first year of the implementation of the 2007 Guidelines and also to reflect on the content of the Guidelines based on international experience. The CONEVAL requested in-depth information on some specific areas, including:

- **To analyze the Guidelines within a legal context:** Mexico has a long history of initiatives in the scope of performance evaluation and orientation of the public administration, hence one of the key questions of this job was to analyze whether the Guidelines were complementary to the current legislation or if there were duplications and conflicts. This analysis can be seen in Report 2 of this Volume.

- **To place the Guidelines in an international context:** CONEVAL was interested in knowing whether the technical tools included in the Guidelines reflected a high level of quality taking into consideration the international experience on the measurement of performance. This analysis can be observed in Report 3 of this Volume.

- **To get to know the perception of the Ministries and the programs on the usefulness of the Guidelines and the implementation process:** A series of interviews were held with programs and entities subject to the Guidelines, so as to know, from their perspectives, the successes and challenges found during 2007. This analysis can be observed in all the four reports submitted.

- **A reflection on the infrastructure of the basic information to insert the requirements into the Guidelines:** To carry out an M&E initiative, it is necessary to have an infrastructure of basic information to provide feedback both to the indicators and to the evaluations. A quick analysis on the status of the information for results in Mexico can be observed in Report 4.
# TABLE OF CONTENTS

**Introduction** ................................................................. 1
National Monitoring and Evaluation Guidelines

**REPORT 1** ............................................................................. 6

**REPORT 2** ............................................................................. 17
*Manuel Fernando Castro Quiroz*

**REPORT 3** ............................................................................. 37
*Alejandro Medina Giopp*

**REPORT 4** ............................................................................. 65
*Gilberto Moncada Vigo*

**References** ........................................................................... 74
NATIONAL MONITORING AND EVALUATION GUIDELINES

INTRODUCTION

The Guidelines as articulated are ambitious in reach, endeavoring to be essentially a two pronged tool, containing well defined operational instruments for M&E but also containing strategic guidance on the coordination of the M&E process. This report through the analysis of the design, content and implementation process of the Guidelines, aims to provide insights which will inform the revision process of the Guidelines, scheduled to be continued in 2008. In 2007 the first wave of the implementation of the Guidelines concentrated on the elaboration of Indicator Matrices (MIs) using the logic framework and consistency evaluations, the second wave in 2008 will concentrate on implementing a selection of evaluations based on the Annual Evaluation Agenda.

THE CONTEXT OF RESULTS ORIENTATION IN MEXICO

The enthusiasm for results is palpable among stakeholders in the Mexican public administration. This enthusiasm has been fueled by important legislation as well as a number of extremely influential impact evaluations of key social programs in Mexico. The national and international recognition which the impact evaluations of Progreso/Oportunidades have garnered have played an important role in putting evaluation on the map in Mexico. The Guidelines come at a time when many important initiatives driving a results orientation of the public administration are developing, each emphasizing different usages of M&E. The National System of Results (SED) which is chiefly a collaboration between SHCP and SFP has the dual objectives of creating a linkage between the budget and results (lead by SCHP), and improvement of management (lead by SFP).

There are also important efforts driven by the Office of the President (Presidencia) to link sectoral objectives to the National Development Plan (NDP). The Guidelines add an important perspective to the spectrum of initiatives driving a results orientation in Mexico: the usage of M&E tools for the evaluation of the design and implementation of social development programs with the objective of improved social policy results. This use of M&E tools has the potential to contribute to reduced poverty.

KEY CONCLUSIONS: SUCCESSES AND CHALLENGES

The diagnosis has found that the Guidelines as currently elaborated provide many opportunities for deepening and expansion. Key successes of the Guidelines are, the creation of a common language among the public administration in terms of results orientation, initiating a process of learning among public servants in M&E and providing ministries and programs with a potent menu of M&E tools for the evaluation of the design and implementation of government programs with the objective of improved policy results. These are laudable achievements and there exists no doubt that the
Guidelines are an important step which has contributed significantly to the new culture of performance which Mexico seeks to create in its public administration. The analysis has found however that important challenges exist, these need to be overcome in order to maximize the utility of the Guidelines.

Key challenges include delineating clearly roles and responsibilities of institutional stakeholders for the implementation of the Guidelines, tightening the connection with the NDP, Sector and Program objectives and strengthening the monitoring and information component of the Guidelines so as to ensure the sustainability of the evaluation agenda laid out in the guidelines. In terms of the role for CONEVAL as Mexico moves forward in its result orientation the Team sees many opportunities for the Agency to strengthen its role through the Guidelines, setting quality standard in the tools of M&E, providing technical assistance to units in Ministries and Agencies involved with implementing the Guidelines and other M&E initiatives such as the SED, and leading initiatives which will expand the supply of good evaluators to respond the demand the Guidelines have generated.

**Defining Roles and Responsibilities**

The analysis has shown opportunity to enhance the Guidelines to provide more clarity and direction in terms of the roles, responsibility’s and mechanisms of coordination between different actors. Indeed a general conclusion that this analysis has reached is that the Guidelines can move far beyond what is currently included in the norm in terms of clarifying and deepening its content. There is room in the Guidelines to show not only WHAT is to be achieved but WHO exactly is responsible for different steps in the process. A sample of these key processes which remain unclear in the current Guidelines include a) consolidation of all the Indicator Matrices (MI) produced by the programs across the public administration b) ensuring a quality standard of the exercises across programs and ministries c) providing technical support to programs to complete exercises d) validating the results of exercises such as the MI e) main users of the information generated by the Guidelines. Identifying clearly in the Guidelines these issues will strengthen not only the quality of the exercises but also contribute to their sustainability.

**Linkage with Strategic Planning at the National and Sector Level**

It is recommended that the Guidelines incorporate a permanent instrument/incentive which allows for this logical chain, linking the basic unit of programs, to sector objectives, to the National Development Plan (part is contained in the logic frameworks). The Guidelines are the right forum to provide this methodological tool and operationalize the vision of results based planning. Finally bringing Presidencia to the table makes sense due to the fact that it also has an important role as a user of the results produced by the M&E system, a high profile user of M&E results helps to strengthen the social accountability function of an M&E system in an important way. In Colombia for example, the president’s use of information from the Colombia national M&E system
(SINERGIA) in his weekly ‘town hall’ meetings has brought an enormous amount of value and visibility on the information put in the system.

The Evaluator Market

The Guidelines increase the demand for evaluators in the country significantly; strengthening the market for evaluators is an important part of ensuring the success of the Guidelines and the ambitious evaluation menu its put forward. Though not in the strict remit of this analysis, it is recommended based on the diagnosis of the limited pool of evaluators currently available in Mexico that measures be taken to alleviate this shortage and expand the market. Strategies for this could include a revision of the existing laws to make hiring non public institutions easier, innovative contracting partnerships between evaluation and sector specialists and certification initiatives.

Matrix of Indicators (MI)

The MI has a number of uses. For example, in the Rules of Operation and informing the results based budgeting initiatives of SCHP. It is important to be aware of these uses and to for example in the case of the Rules of Operation avoid duplications and make use of opportunities for complementarities. An important consideration to take into account is limiting excessive reporting burden on programs and achieving clarity. With respect to the MIs four key recommendations can be made: a) to complement and extend these to link with sector and national goals as discussed above, and b) to ensure a unified standards and training across the public administration c) to capacitate Units in ministries to lend technical assistance in the elaboration of the MIs and finally, d) to establish agreed clear and feasible guidelines in relation to when they must be delivered, to whom they must be delivered and for what purpose.

Evaluations

Preliminary diagnosis of the menu of evaluations shows that the 5 types cover well the menu of evaluations which are considered useful for programs according to international best practices and offer new innovations such as indicator evaluation and process evaluations. The consistency and design evaluations completed this year as part of the first wave of implementation were seen by evaluators to be a good instrument for analysis, furthermore programs themselves perceived evaluations to be a useful tool. Included in the report of managerial instruments is a detailed diagnosis of the two evaluation tools. In both cases it is important to note that in particular the standardized ToRs given to programs were seen as very useful by programs, giving them clear guidance as to what they should seek in evaluators. As mentioned before there exists frustration regarding large amounts of money spent on sub-par evaluations, due to a lack of capacity in Ministries to contract with adequate knowledge of mechanism to ensure evaluation quality. The ToRs are a clear move forward in this respect, reported by both the programs and evaluators to have been very useful.

The main challenge however that the Team found in its analysis of the evaluations was the need to be very clear and strategic in the number and timing of the evaluations that
are implemented (graphical example on the use of the instruments is provided below). The menu is large and could if implemented too quickly and unselectively impose unnecessary stress on programs. It is as yet unclear for all of the evaluations when and in what circumstance they will be applied, the evaluation of design does establish circumstances; this should be replicated for all evaluations to the extent possible in the Guidelines. Taking into account the lack of supply of evaluators and the cost of evaluations, it is advisable that the evaluations to be completed be carefully considered, the Consistency of Results Evaluation has within it the potential to be utilized as a tool to identify which programs need to have specific evaluations included in the Guidelines Menu. These considerations should be taken into account in the evaluation agenda which will be provided by the National Evaluation Plan.

**Figure 1: Mexican Menu of Evaluations Based on the Logic of a Program**

![Diagram](image)

*Source: Produced by World Bank staff, 2008*

**LEGAL ANALYSIS AND THE FISCAL REFORM**

Mexico has successfully built a broad based and solid legal foundation stressing the significance of results in the public administration. This is extremely important for the sustainability of the system in the longer term and is a clear strength for Mexico as it moves forward. The main laws, which have pushed the results agenda forward in Mexico, are a) the Law of Social Development (DOF 20-01-2004), b) the Fiscal Responsibility Law (DOF 30--2006) c) the Federal Law of Transparency and Access to Information (DOF 06-06-2006) and d) the Annual Budget Laws of 2006, 2007 and 2008. These laws have provided coordination to what was a limited and disorganized legal framework for results orientation. Deeper analysis of these legislative documents is contained in the report of institutional issues, important themes elaborated there include the need for CONEVAL to fully take advantage of the opportunities which lie in this rich normative framework.

The predominant conclusion which is reached by this analysis is that the Guidelines provide a complement to and deepening of the existing norms. Indeed, one of the key strands in the normative development of results orientation in Mexico, has been the so-called ‘Fiscal Reform’, this term represents a collection of legislative initiatives, including for example the Fiscal Responsibility Law (DOF 30--2006) which have had the
general objective of promoting the modernization, transparency, efficiency and management of public income and expenditures.

The Guidelines have played an important role in adding to this 'financial administration' concentration, the perspective of improving the quality and impact of public allocation. In sum, from a legal perspective the Guidelines have filled a gap by providing the Mexican public administration with a set of methodological tools and norms allowing programs to demonstrate and improve the results of their programs. The legal analysis has shown that if any recommendations are to be given in terms of concrete actions, there is room for the development of existing mandates through complementary policy notes, clarifying existing directives.

**LOOKING TO THE FUTURE: CONEVAL’S ROLE**

Given the discussion above the Team considered CONEVAL’s positioning within the institutional framework for M&E. There are a number of strategic needs created by the move towards results orientation which CONEVAL could play a leading role in fulfilling. CONEVAL is an institution which contains a high concentration of technical knowledge regarding social policy evaluation. But it is also a small institution with limited capacity in the sense of the volume of work it can process and perhaps more importantly in terms of the institutional weight it has.

Areas in which it can leverage its comparative advantages can be grouped into a set of activities surrounding setting standards for quality evaluations, providing technical assistance in evaluation tools, providing guidance in the managing of these tools, establishing the evaluation agenda and conducting a limited number of strategic impact evaluations. In addition, as discussed, a monitoring strategy is lacking at the moment. However as this is articulated in the future, CONEVAL could also play an important role as an establisher of standards and lender of technical assistance in this area. This is advisable as the two are interdependent.
REPORT 1: EXECUTIVE SUMMARY OF ANALYSIS OF GENERAL GUIDELINES FOR THE EVALUATION OF THE FEDERAL PROGRAMS OF FEDERAL PUBLIC ADMINISTRATION

INTRODUCTION

This document presents the main findings and recommendations of the World Bank work team hired by the National Council for the Evaluation of Social Development Policy in Mexico (CONEVAL), to analyze the Evaluation Guidelines published on March 30, 2007 in the Official State Gazette. These Guidelines were issued jointly by the Finance and Public Credit Secretariat (SHCP), CONEVAL and the Civil Service Secretariat (SFP) to regulate the evaluation of federal programs, the drawing up of an Indicator Matrix (MI) and follow-up systems and the strategic objectives of Federal Public Administration (APF) organizations.

The guidelines established are ambitious in scope, as well as being an essential tool with two main features: well-defined monitoring and evaluation instruments (M&E) and strategic guides for coordinating their implementation.

Through the analysis of the design, contents and implementation of the guidelines, this report is designed to provide information that will contribute to their implementation in the short term and their revision in the long term. In 2007, the first stage of implementation of the Guidelines was carried out, which focused on drawing up the Indicator Matrix (MI) and the Consistency and Results Evaluations (C&R). The second stage, to be carried out in 2008, will focus on undertaking a selection of evaluations, on the basis of the 2008 Annual Evaluation Plan.

The Context of Results-Orientation in Mexico

The importance of reinforcing results-based management is obvious in Mexican public administration. This importance has been encouraged by current legislation, as well as a series of evaluations with an enormous impact on key social programs in Mexico. The national and international recognition achieved by impact evaluations such as the one carried out on the Progresa-Oportunidades program has played a significant role in disseminating the importance of evaluation in Mexico. The Guidelines were produced in a context in which several major results-oriented initiatives are being developed in APF. Each of these initiatives stresses the various applications of M&E. On the one hand, the Performance Evaluation System (SED), an SHCP initiative, is designed to create a link between budget and results, and on the other, the Program to Improve Administrative Management (PMG), directed by SFP, emphasizes institutional improvement on the basis of performance incentives.

The President's Office (Presidencia) is also making a significant effort to link sectoral objectives to the National Development Plan. The Guidelines were drawn up in accordance with these initiatives, adding a new perspective to results-orientation in Mexico through the use of M&E tools to evaluate program design and implementation.
They are intended to achieve better results in federal policy and thereby contribute to poverty reduction strategies.

**Main Conclusions: Achievements and Challenges**

This diagnosis has found that the Guidelines provide opportunities for their consolidation as well as areas of expansion. In this respect, the main achievements regarding the Guidelines can be said to revolve around three main points: a) the creation of a common language within APF in terms of results-orientation; b) the start of a learning process among government officials on the subject of M&E; and the c) the provision of a standardized menu of M&E tools for both the Secretariats and the programs, in order to achieve better public policy results.

This constitutes significant progress in the effort to create a new culture of performance in Mexican public administration. Nevertheless, significant challenges have been identified that must be dealt with to maximize the Guidelines' usefulness.

These challenges include: a) reinforcing the functions and responsibilities of the institutions involved in implementing the guidelines; b) strengthening the link between the PND and the objectives of the Sectors and Programs and c) increasing the use of the evaluation results as well as reinforcing the information strategy for results. This last point is crucial, since providing quality information for evaluations is essential to ensuring the sustainability of the evaluation agenda drawn up in the Guidelines.

This document summarizes the main findings of the diagnosis and research missions undertaken in Mexico, as well as the extensive review of the key documents that served to produce the report drawn up by the Team on the analysis of the Guidelines.
MAIN FINDINGS

Legal Analysis and Tax Reform

A review of the legal framework in which the Guidelines operate showed that Mexico has successfully created a broad, solid base underlining the central positions of management by results within public administration. This suggests that the normative framework regarding results in Mexico is one of the most advanced in Latin America. The main regulations that have driven the results agenda are: a) The Social Development Law (DOF 20-01-2004), b) the Federal Law of the Budget and Budgetary Responsibility (DOF 30-03-2006), c) the Federal Law of Transparency and Access to Information (DOF 06-06-2006) and, d) the Annual Federal Expenditure Budget (2006, 2007 and 2008). These laws provide greater coordination regarding results-orientation and raise the challenge of how to make the best use of existing laws, such as ensuring that the Federal Law of Transparency and Access to Information reinforces the role of evaluation in Mexico.

The predominant conclusion of this analysis is that the Guidelines provide a complement to existing norms, in addition to consolidating them. Indeed, one of the key aspects in the normative development of results-orientation in Mexico has been Fiscal Reform. This is designed to ensure that laws such as the Budgetary Responsibility Law (DOF 30-03-2006) promote modernization, transparency and efficiency in public income and expenditure management.

The Guidelines have played an important role by adding the perspective of improving the quality and impact of public budget assignment to "financial management". In short, the Guidelines have bridged a normative divide, providing public administration in Mexico with a set of methodological tools and laws that will enable programs to evaluate and improve their results. In terms of legal analysis, this consultancy suggests a series of specific actions concerning the development of existing mandates, through the creation of complementary notes to policies that clarify existing guidelines.

Institutional Agreements: definition of functions and responsibilities

The inter-institutional cooperation achieved for the development and publication of the guidelines was an unusual, successful initiative, led by three institutions with different mandates in APF. An analysis of the implementation process, however, has shown that there is room for improvement as regards institutionalizing monitoring and results-based analysis in APF. This includes: a) the consolidation of all the Indicator Matrices (MI) produced for the programs, b) ensuring quality standards for both the Indicator Matrices and the evaluations, c) providing technical support for programs in order to carry out the exercises d) validating results and e) using the information provided by the Guidelines.

It is therefore suggested that an inter-institutional committee be established to reinforce coordination of the implementation of the Guidelines. This committee would be a forum through which CONEVAL, SHCP and SFP could achieve consensuses and exchange and express ideas and strategies. Another suggested strategy is the use of pilot tests rather
than a “big bang” process such as the one chosen last year. Pilot tests enable institutions to identify the difficulties in implementation and the establishment of successful strategies for achieving a broader impact on the part of the initiative. Pilot tests have proved successful in several countries including the United States, Chile, Egypt and Sri Lanka.

The Role of Information Users

Another strategic decision suggested to reinforce the Guidelines is to increase the use of evaluations. This analysis showed that, as a first step, it is essential to identify the main users of the information produced by the tool menu. It is therefore suggested that an instrument be developed that would use a logical chain, linking the basic unit of programs to sectoral objectives and the National Development Plan (part of which is found within the logical framework). This would enable public policy decision-makers to be included at all levels of this logical chain. A multi-annual PND evaluation could also be included in the evaluations menu. Both activities could secure the role of key users of the Guidelines’ results. A strategy such as this helps significantly strengthen the function of social responsibility in an M&E system.

In Colombia, for example, the use of information produced by the National M&E System (SINERGIA) has provided enormous value, prestige and publicity for this information.

Liasing between the Three Levels of Performance: Programs, Sectors and National

The analysis revealed that in general, the Guidelines stress the importance of reinforcing the link between national planning and the budget on the basis of M&E tools. But the M&E approach is still largely centered on the programs as units of analysis. The MI and the various evaluation methods established in the Guidelines concentrate on the program and fail to consider other important evaluation units such as the National Plan and the sectoral plans. This aspect not only limits the scope of the evaluation in promoting the alignment between the programs, plans, budget and its results but also weakens the role of planning as a reference framework for management and evaluation, contrary to the aim established in SED.

In addition to the above, the Team sees opportunities for strengthening the link between the three levels of performance of the Guidelines. For example, it suggests examining how the evaluation of C&R, in addition to measuring the alignment of the programs with the sectoral plans, could examine the consistency of Sectoral Plans with the National Plan and that of these two instruments with the budget.

There is also a lack of instruments for measuring results at the intermediate or upper strategic level (as an evaluation of PND).

M&E at the Organization and Agency Level

This document has discussed the diagnosis in terms of institutional leadership at the level of the most comprehensive ones (CONEVAL, SHCP and SFP). However, coordination at the secretariat level is equally important. The Guidelines demand the establishment of an
Administrative Unit (Unit) within the existing structure of the secretariats with the responsibility of coordinating the contracting and ‘supervising the evaluators.’

The diagnosis of the current situation of these Units revealed several challenges. First, there is a lack of clarity about who is in charge of contracting the evaluations. To date, both the programs and the evaluation units have been involved in the hiring. One consequence of the lack of coordination observed was that in some of the organizations, evaluators had very little time to complete the evaluations due to the lengthy contracting process. Delivering the evaluations is required by law, meaning that they are forced to follow a very strict schedule. This creates contradictory incentives and does not encourage quality evaluations.

The diagnosis also showed that at the moment, there is an enormous difference between the capacities of the Units in the secretariats. For example, there are evaluation units with an advanced technical capacity due to their years of experience with the System of Evaluation and Monitoring by Results (SEMBR), whereas other secretariats have very few personnel, with very little experience.

As a result of this analysis, it is suggested that an agenda be promoted to strengthen the monitoring and evaluation systems at the secretariat level. Although this study focuses on the implications of the implementation of the Guidelines, it is worth noting that this affects the implementation of the entire results agenda. This agenda, which has been established in Mexico by involving the SED and PMG initiatives and the Guidelines, has significant implications for the secretariats’ capacity to manage the monitoring and evaluation systems.

It would therefore be important to coordinate initiatives, rationalize efforts, and take advantage of the synergies and natural skills of the various institutions. The aim is to create systems and teams at the secretariat level with the capacity to implement management by results. This process of learning and construction implies inter-institutional coordination. Cooperation opportunities between CONEVAL, SHCP and SFP should be explored to coordinate the design of a work plan that will promote capacity building in the issue in question over the next few years.

**Indicator Matrix (MI)**

The results of the 2007 implementation process observed by the Team underline the need to improve MI Coordination. On the one hand, the initial message given when the Guidelines were published was that MIs focused on planning and evaluation. It was also understood that the programs would have some time to complete the MIs, assimilate training and receive feedback on the preliminary versions from both CONEVAL and SHCP before delivering the final versions.

However, the request to provide a selection of indicators for use in the annual budget (PEF) for 2008 created anxiety among the programs, since the indicators would be used to determine their budgets.
The time given to complete this request, approximately two weeks earlier than originally planned, placed a significant amount of pressure on those carrying out the exercise, most of whom were doing so for the first time.

A second major problem was the confusion regarding the concepts involved and the differences in the messages sent by the trainers.

Given that the MI represented the first stage of implementation of the Guidelines in 2007, this exercise has been responsible for introducing the concepts and the basic knowledge of a results-oriented culture into Public Administration in Mexico.

The challenges associated with this tool are concerned, as mentioned earlier, with the technical capacity within the secretariats to help programs draw up MI and with their training to do so. A review of the MI samples showed that there is a wide range of quality in the exercises completed. In some cases, there still appears to be confusion, for example, about whether the MI should reflect what the program is doing at the moment or what it seeks to do over the next few years as well as the difference that exists between beneficiaries and intermediaries.

Another challenge identified in the diagnosis was the difficulty faced by the programs in drawing up the Goal and Purpose indicators. This is due to the shortage of sources of information used to feed the indicators, the absence of clarity regarding sectoral objectives and on occasions, the lack of clarity in the processes of technical and political validation.

Moreover, it was found that an MI was sometimes produced for very broad programs containing several projects, which complicates the process of drawing up and understanding program objectives.

As a result, four key recommendations can be made: a) encourage the complementarity and breadth of MI to link them to sectoral and national objectives, as pointed out earlier in this document, b) ensure standards and unified training through APF, c) train the Units in the secretariats to provide technical assistance in drawing up MI and lastly, agree on clear, feasible guidelines regarding when they should be delivered, to whom and for what purpose.
Evaluations

In recent years, there has been an international tendency towards the use of a broad menu of evaluation methodologies to measure specific aspects that affect a program's success. The diagnosis of the menu of evaluations contained in the Guidelines shows that the five types of evaluations specified are regarded as best international practices (such as impact evaluations). Moreover, the Guidelines and Terms of Reference (TDR) for every evaluation carried out by CONEVAL introduce the standardization of innovative evaluations, such as the C&R, Indicator and Process evaluations. On the whole, the types of evaluations included in the Guidelines are of a high standard and continue the international trend to use a broad menu of methodologies to measure performance.

However, the main challenge identified by this consultancy firm is to be very clear and strategic about the number of evaluations to be carried out and the time in which they should be implemented. There is a broad range of options and if they are implemented too quickly and/or on a non-selective basis, they may cause unnecessary tension in the programs. For example, the design evaluation establishes specific circumstances in which it should be carried out. However, this must be reproduced for the rest of the evaluations, as far as possible. The Evaluation of Consistency and Results is also considered to have the potential to be used as a tool for identifying what programs need to complete other specific evaluations within the evaluation options. These considerations must be borne in mind in the Annual Evaluation Program.

Graph A: Logic of Combining Tools

![Graph A: Logic of Combining Tools](source: Drawn up by The World Bank, 2007.)

In 2007, the design and C&R evaluations were undertaken as part of the first stage of implementing the Guidelines. The diagnosis shows that the evaluations were regarded by the evaluators as a good instrument of analysis and that the programs evaluated perceived the evaluations as a useful management tool.

It is important to note that the secretariats experienced great difficulty in hiring suitable evaluators that could guarantee the quality of the evaluations. The TDR constitute a step forward in strengthening this aspect. However, one of the challenges is to clearly identify the use and uses of the evaluation results. For example, the evaluation results could provide key input for the SED and PMG.
It is also suggested that "improvement plans" be drawn up on the basis of the recommendations accompanying each C&R evaluation. Improvement plans are an integral part of the Program Assessment Rating Tool (PART) in the United States and Government Program Evaluation (EPG) in Chile, since they provide incentives for the programs to use the evaluations. Both proposals required minor adjustments in the evaluation itself and in the case of the improvement plans, the establishment of arrangements to follow them up. However, since international experience has shown that the survival of evaluation initiatives largely depends on the use of the results of the latter, it is thought that the effort would be cost-effective.

**The Evaluators’ Market**

Since the Guidelines significantly increased the demand for evaluators in the country, strengthening the market of evaluators constitutes an important element in ensuring the success of the Guidelines that consider an ambitious menu of evaluations. Although strictly speaking, it is beyond the scope of this analysis, it is suggested that specific measures be taken to boost the increase in the number of qualified evaluators in the country and thereby expand the evaluation market in Mexico. Some of the strategies for achieving this could include a review of existing laws on the issue to facilitate hiring private institutions and innovative contracting societies and sectorial specialists and specialists in evaluation and certification initiatives.

**A Function for Monitoring and Information**

Although the Guidelines clearly propose a function for monitoring at the same time as evaluation, there is no articulated tool to make monitoring operative within APF, which is obvious in the case of evaluation and MI. There is a broad understanding among the parties involved in Mexico regarding the importance of providing quality information. Good evaluations depend on the good information given to the evaluators which in turn feeds the program’s MI. As such, the development of quality systems for monitoring, as well as the existence of formal information gathering sources are an important requirement for the sustainability of M&E systems.

It is suggested that an articulated vision be promoted for monitoring public administration that will feed and complement the evaluation infrastructure currently outlined in the Guidelines. This vision must include procedures, standards, functions and responsibilities, internal and external coordination mechanisms and rules for dissemination.

The diagnosis of the state of the information systems existing today in APF concludes that there is little confidence in being able to cover the demand for quality information created by the ambitious evaluation agenda established in the Guidelines. The reasons behind this concern include the finding that there is limited availability of basic information (mainly as regards administrative records) which is unreliable, too general and unsuitable for calculating indicators. Moreover, the programs, which are responsible for data collection, have focused mainly on collecting information on processes.
Lastly, there is a lack of universal quality standards used by public administration for information gathering, as a result of which there is a wide variation in the quality of this information.

MAIN CHALLENGES

- Establish an inter-institutional committee to promote planned, periodic interaction and communication between the parties involved with the Guidelines.

- Improve the training and technical assistance strategy for both indicator matrices and issues related to various types of evaluation.

- Strategically consider the uses of evaluation tools included in the guidelines. In the short time, opportunities should be examined for using C&R Evaluation as a generator of information for SED and PMG. It would be important to review this tool to make it more executive and useful for key actors. Efforts should also be made to examine how to use the tool to diagnose what programs need other specific evaluations such as indicator and process evaluations (evaluation agenda). These efforts would help transform the C&R evaluation into an evaluation instrument with specific uses and users.

- Balance Incentives: A clearer explanation to the target audience on the uses of the information produced by the tools in the Guidelines will ensure that the programs not only see this as “extra work” but also recognize their value, which will make it possible to create more balanced incentives.

- In the short term, the possibility of creating a sort of “management contract” that would support the implementation of evaluation recommendations at the program level should be examined. In the C&R evaluation, evaluators should submit recommendations on the basis of which a work plan could be created.

- Encourage a debate among those involved on the function of monitoring and its role in the Guidelines. In the short term, efforts should be made to create a consensus on the monitoring model desired by the key parties involved. A clear definition should also be provided of: a) the objectives of the tools; b) the methodologies, c) the uses of the tools; d) the processes and e) institutional agreements.

- The laws regulating external contracts should be reviewed to ensure greater clarity in the process as regards: a) who can be hired; b) who the suppliers can be; c) what forms of contract there are and d) what the conflict resolution mechanisms are. Other initiatives that may be submitted for consideration are: a) creating an evaluators' register and b) introducing innovative ways of increasing the group of evaluators, for example, through “partnership contracts.”

- In the long term, efforts must be made to form alliances with universities to promote education initiatives in the area of M&E programs. For example, Canada and Australia implemented an evaluation program in universities to train a sizeable group
of evaluators for both government and the private sector. Another long-term initiative involves establishing a program to facilitate the transmission of knowledge to boost the technical capacity of the evaluations.

Information by Results Agenda

Within this analysis of the Guidelines, it is essential to include a preliminary diagnosis of the state of information in Mexico. Providing quality information is a key issue for ensuring that the results of the MI evaluations and indicators are reliable. Below are some of the aspects that could help reinforce an agenda of information by results.

- Build the technical capacity of the personnel on the program to produce and use basic information. A short-term activity could be to: a) develop standard criteria to manage information in the programs and b) design and implement a capacity building strategy for the program's operating personnel on information management.

- Improve administrative records so that they can be used more actively as a source of basic information. A short-term activity could be to: a) diagnose the production of information in the programs and b) design and implement a strategic program to improve administrative records.

- Standardize the databases used by the M&E system. One activity could be to focus on carrying out diagnoses of the databases currently used in the programs. On the basis of the medium- and long-term diagnoses, one can: implement a strategic plan to reinforce the databases, b) develop rules and protocols to capture the programs' information flow and c) design and establish a capacity building strategy for database management.

- Establish strategic alliances with INEGI in order to be able to take advantage of its technical capacity and statistical infrastructure. Some activities could include establishing agreements to: a) reinforce training in information management programs and b) ensure rapid, continuous access to their information. It is also necessary to draw up an inventory of useful information for the programs. In the long term, technical committees should be established to administer the information in the program and incorporate it into INEGI's working network.
REPORT 2: INSTITUTIONAL ANALYSIS OF GENERAL GUIDELINES FOR EVALUATION OF FEDERAL PROGRAMS OF FEDERAL PUBLIC ADMINISTRATION

Main author: Manuel Fernando Castro Quiroz

INTRODUCTION

This document presents the preliminary results of the “Supporting Social Policy Evaluation-CONEVAL” consultancy. The document is the result of the review of secondary information and interviews with the persons responsible for the functions of planning, budgets, follow-up and evaluation of the national and sectoral levels in Mexico.

Taking into account international and national experience, on the basis of the analysis and consultative review of the current institutional and legal framework, existing planning, budget, monitoring and evaluation (M&E) instruments and available information systems and institutional capacities, it proposes elements to reinforce the evaluation guidelines drawn up in 2007 by the National Council of Evaluation of Social Development Policy (CONEVAL), the Secretariat of Finance and Public Credit (SHCP) and the Civil Service Secretariat (SFP). It also includes recommendations to facilitate its instrumentation in the organizations and offices at the central level in Mexico.

INSTITUTIONAL AND LEGAL CONTEXT OF GENERAL GUIDELINES FOR EVALUATION

Over the past ten years, Mexico has taken a major step towards reinforcing M&E functions in Federal Public Administration (APF). The most significant progress occurred during a period characterized by great enthusiasm for results-based management in central administration, particularly at the level of the Social Development Secretariat, SEDESOL, the Finance and Public Credit Secretariat (SHCP) and the Civil Service Secretariat (SFP). This has made it possible to institutionalize a series of new functions and instruments through major normative provisions, and the introduction of new management practices in the secretariats. Issuing General Evaluation Guidelines for Federal Programs constitutes one of the main landmarks in this process.

The principal changes introduced have formed part of a dual, concomitant process of transformation in the Mexican public sector that includes social policy reforms on the one hand and fiscal reform on the other. Fiscal reform, underway at least since 2003, includes actions aimed at modernizing and making public finance management more efficient and transparent on the side of income (mainly through tax system and customs reforms) and on the side of expenditure (such as the reforms of the budgetary and acquisition systems within Public Financial Administration).

Under the fiscal reform approach, the evaluation of results has played a key role in improving financial administration. Its importance lies in the fact that all the efforts of fiscal reform to modernize and increase the state's income base (the main goal on the income side) will be of little value in its management unless the quality and impact of allocations are improved (ultimate goal on the expenditure side).
Within this context, there are several provisions governing evaluation tasks and other related tasks in both general and specific activities in Mexican administration. The main laws that have driven this process of institutional change and which define the framework for developing CONEVAL's M&E functions are: i) the General Law of Social Development, ii) the Federal Law of the Budget and Financial Responsibilities, iii) the Federal Law of Transparency and Access to Public Government Information and the annual laws of the Federal Expenditure Budget 2006, 2007 and 2008 (see Table 1).

The Law of Social Development paved the way towards the institutionalization of this function by making social development policy evaluation and civic participation and control of the planning and implementation of the latter mandatory in the state. Through the creation of the National Council of Evaluation of Social Development Policy (CONEVAL), this law established the institutional framework for making evaluation a permanent practice in the social sector on the basis of technical guidelines and criteria, while ensuring the minimal availability of human and financial resources to promote their development.

For its part, the Federal Law of Budget and Financial Responsibility, passed in 2006, updated and modernized the budgetary systems and introduced the bases for orienting public financial management and in particular, shifting federal expenditure planning towards management by results. As the main framework of reference, in order to achieve this goal, it created the Performance Evaluation System (SED) and established the coordination of the latter by the Finance Secretariat. The scope and definition of specific institutional roles and arrangements for the implementation of SED is still underway and according to this same law, must be determined by March 2008 at the latest.

---

2 In this respect, the following laws also count: The Law of Planning (DOF 13-06 2003), Advanced Supervision (DOF 30-08 2007) Promotion of Activities of Civil Society Organizations (DOF 09-02 2004).

3 The norms have been developed and complemented by a series of normative provision including: i) The Organic Status of the National Council of Evaluation of Social Development Policy and its Regulatory Decree (DOF 24-08-2005); ii) the General Guidelines for the Evaluation of Federal Programs; iii) the General Guidelines for the 2008 Programming and Budgetary Process and iv) the Decree of APF Austerity and Discipline in Spending.
### TABLE 1: Principle Laws and Regulations Related to CONEVAL Functions

<table>
<thead>
<tr>
<th>Law</th>
<th>Aspect regulated</th>
<th>Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Law of Social Development (DOF 20-01-2004)</td>
<td>- Program and social policy evaluation and follow-up mechanisms.</td>
<td>- Evaluation-planning-budget link</td>
</tr>
<tr>
<td></td>
<td>- Creation of CONEVAL.</td>
<td>- Organization and functions of CONEVAL and other bodies.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Defines Evaluation procedures as part of social policy.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Participation and social comptrollership mechanisms.</td>
</tr>
<tr>
<td>Federal Law of Budget and Financial Responsibility (DOF 30-03-2006)</td>
<td>- Planning, assignment, use and evaluation of federal income and expenditure.</td>
<td>- Defines principles, procedures and persons responsible for ensuring</td>
</tr>
<tr>
<td></td>
<td>- Assignment of resources for evaluating programs and structure of CONEVAL.</td>
<td>- Defines reference framework for evaluation of expenditure.</td>
</tr>
<tr>
<td></td>
<td>- Dissemination and updating of performance goals and indicators and budgetary results.</td>
<td>- Introduces figure of annual evaluation program for federal programs.</td>
</tr>
<tr>
<td></td>
<td>- Creation of Federal Institute of Access to Information.</td>
<td></td>
</tr>
<tr>
<td>CONEVAL Organic Statute (DOF 24-08-2005)</td>
<td>Organization and functions.</td>
<td>- Principles and general objective.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Assets.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Organic structure.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Evaluation typology.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Coordination mechanisms.</td>
</tr>
<tr>
<td>General Guidelines for Programming and Budget Process 2008</td>
<td>- Rules, sphere of application and Results-based Budget Approach (PBR) and Performance Evaluation System (SED)</td>
<td>- Conceptual and methodological definition of PBR and SED.</td>
</tr>
<tr>
<td></td>
<td>- Planning-budget-evaluation integration.</td>
<td>- Programmatical Classifiers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Guidelines for advancing towards PBR and SED.</td>
</tr>
</tbody>
</table>

Source: Drawn up by author on the basis of the Diarios Oficiales de la Federación cited in the table.

On the basis of the above, the main advances implemented as a result of these laws and regulations include the following:

- Compulsory nature of drawing up the budget on the basis of quantifiable objectives and parameters by incorporating indicators and goals to measure performance in the
objectives and strategies in the National Development Plan and its programs (Art. 16);

- Need to take into account evaluation information on the performance of the previous fiscal year in the annual programming and budget;

- Incorporation of programmatical classifiers to facilitate the evaluation of expenditure results in terms of policy objectives, programs, projects, activities and their goals (Art. 28);

- Obligation to make information on expenditure results public in view of the provision of the Federal Law of Transparency and Access to Public Government Information (Art. 106);

- Institutionalization of the evaluation of expenditure results as a recurrent practice of budget management within the framework of SED (Arts. 110 and 111).

These aspects were recorded and reinforced by the subsequent laws of the Federal Expenditure Budgets for 2006, 2007 and 2008, during which evaluation became mandatory after the assignment of resources was defined together with the establishment of the Annual Evaluation Program of federal programs.

But one of the most comprehensive laws, with the greatest influence for the future development of the evaluation functions in the Mexican States is known as “General Guidelines for Programming and Budgeting for the 2008 Tax Year,” issued by SHCP in July 2007. In addition to including evaluation guidelines in its contents, this law is the first conceptual and operative development of the results-based management and budget model and the Performance Evaluation System-SED. It establishes the mechanisms for achieving greater harmonization between the National Development Plan, Sectoral Plans and the Annual Budget and defines an integral scheme and the procedures for its adoption.

An important additional feature is the establishment of the results-based Evaluation and Monitoring System (SEMBR), designed to serve as a technological tool for providing feedback on planning and the budget using information from the performance indicators and goals.

The Federal Law of Transparency and Access to Public Information constituted another significant normative step in the development of M&E in Mexico. On the one hand, this law reinforces the dual nature of the evaluation as an instrument that generates and demands public information, within parameters of quality, timeliness and free access. On the other hand, it establishes a broader institutional framework in which evaluation is crucial to the purposes of an efficient, transparent, accountable public administration. At present, this normative instrument and the institutional bodies comprising it such as the Federal Institute of Access to Information (IFAI) have yet to be incorporated into CONEVAL’s agenda of activities.
Establishing alliances with this type of actors could be used to promote evaluation and lend more weight to the function as well as increasing CONEVAL's positioning within the framework of the national public information agenda. According to the information obtained, there is still no agenda of joint work between CONEVAL and IFAI, which is why a significant potential has been identified in this field.

Together with these provisions, the new administration is promoting two complementary initiatives that have an enormous effect on the process of reinforcing results-based management. The first, implemented by the Civil Service Secretariat, is the Management Improvement Program (PMG). This program seeks to ensure common standards and develop and institutionalize certification methods and management by excellence. The second, led by the President's Office, seeks to reinforce strategic planning in government management through the harmonization and integration of the National Development Plan with sectoral Strategic Plans and the budget. The aim of this integration is to achieve greater alignment of sectoral programs with government priorities, facilitate their evaluation and have more information for assigning resources, bearing in mind the performance goals and indicators to be monitored and evaluated within the framework of SED.

These aspects define the context within which, in March 2007, CONEVAL, in conjunction with SHCP and SFP, drew up the General Guidelines for the evaluation of Federal Programs. These guidelines introduced the first formal technical criteria for the development of the evaluation and the preparation of strategic objectives, the Indicator Matrix and the monitoring and evaluation systems. The Guidelines thereby contributed to the materialization of these laws for all the programs operating with budgetary resources from Public Federal Administration.

In conclusion, nowadays Mexico obviously has a broad, sufficient normative framework to promote the development and consolidation of the evaluation function, some of whose main components are still in a process of regulation and definition. This framework identifies at least three organizations whose functions overlap with those of central administration as regards M&E (SHCP, SFP and the President's Office) and one that partially overlaps with it in the social sector (CONEVAL).

**ELEMENTS OF INTERNATIONAL EXPERIENCE**

Not many developing countries have advanced M&E systems. In particular, the main experiences in this respect are found in Latin America, and from various points of view, Mexico can be considered one of them, together with Chile, Colombia and Brazil. Given that the advances of each of these experiences are specific, it is possible to identify important lessons in different areas (see Table 1).
The following lessons can be drawn from features of the monitoring and evaluation systems in the Latin American region:

- The existence of a modulator of M&E functions with sufficient hierarchy and technical capacity to orient the evaluation and monitoring systems, consolidate and control the quality of performance information and ensure its use in decision-making often helps reinforce the system.

- Although the use of laws and regulations may be important in countries with a certain normative tradition, it is not a necessary condition, as the case of Chile shows. Nor is it essential to have formal policy guidelines, although in some cases, they have proved useful, as in Colombia from 2002 onwards and Mexico from 2007 onwards.

- Links with the budget require an active role in the system and significant technical capacity on the part of the budgetary authority, as shown by the cases of Chile and Brazil.

- The definition of a global aggregation unit such as PND, the budget or presidential goals may facilitate the development of monitoring functions from a central unit.

- The evaluation typologies used by countries are usually similar, since they include common techniques for evaluating impact, processes and the logical framework, with
slight innovations\textsuperscript{4}. However, what appears to be more important is the technical rigor, credibility of results and above all, the use of recommendations in the decision-making process.

- Accountability to Congress is a central element for materializing the formal political dimension of the system. However, the incorporation of the civic dimension may contribute to reinforcing the role of M&E as an instrument of transparency and encouraging participation and social control, thereby creating a broader level of validation and greater interest on the part of the President's office and citizens.

- Orienting the system towards a broader, results-based management approach is often a factor that promotes the integration of the M&E function into other processes such as the budget and human resource hiring or management among others.

- Systems may experience progress or regression associated with numerous factors, whether technical or political. The course should be therefore be maintained through alliances with numerous actors in the state who create a demand for information on performance (such as Congress, academia and civil society organizations, etc.)

- The President's Office may serve as a catalyst and provide armor-plating from internal spheres that may be resisting the process of change, although this is not necessarily a condition.

Generally speaking, one of the main lessons from international experience is that the development of a system with these characteristics is measured by the level of use of the monitoring and evaluation information in at least two objectives: i) decision-making related to planning, budgeting and/or the processes of drawing up programs and ii) permanent accountability, whether formally to Congress and control mechanisms or informally, to citizens.

In view of this, international experience also shows that the development of an M&E system is a long-term process requiring a sustained effort to turn it into instrument that will drive processes of "change" and modernization in public administration.

This change has been observed to be based on proper incentives, building capacities to produce timely, reliable information and the introduction of appropriate institutional arrangements. These aspects are indispensable for ensuring that M&E contributes to integrating all stages of government management as well as the transparency of the latter. In order to achieve this, the common element of success identified in all cases is the capacity of the system to respond to both the technical and political demands it may face.

**MAIN FINDINGS AND RECOMMENDATIONS OF GENERAL EVALUATION GUIDELINES**

Bearing in mind the normative and institutional framework described earlier, as well as the opinion of the various actors involved and the review of available instruments, the

\textsuperscript{4} The Consistency and Results Methodologies in Mexico or the Executive Evaluations in Colombia are included within the category of innovative methodologies.
analysis showed that evaluation guidelines constitute a significant advance in APF. Generally speaking, they can be said to have helped fill a policy gap regarding the function of evaluation and the need to advance towards their integration into national and sectoral planning as well as the public budget.

In particular, the General Guidelines have established a uniform framework in all programs for:

- Defining strategic objectives and indicators;
- Introducing and using common tools for facilitating planning and evaluation (Indicator Matrix);
- Defining the evaluation typologies to be used as well as scheme for following up the results of evaluations and disseminating the latter;
- Examining coordination mechanisms and defining the terms of reference for the evaluations and
- Taking the first steps towards the design and implementation of an integral results-based monitoring and evaluation system in federal programs.

Despite this, the study managed to identify a series of aspects that could be reviewed and adjusted as regards both contents and the procedures and arrangements for their implementation. In these aspects, a high level of consensus was detected between the persons responsible for the various sectors and programs who were interviewed.

Below is an analysis of the findings and the principal recommendations regarding the contents of the guidelines and the implementation of the latter, according to the scheme presented in Table 3.
### TABLE 3. Summary of Findings regarding General Evaluation Guidelines

<table>
<thead>
<tr>
<th>Aspect analyzed</th>
<th>Topics of findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contents of Guidelines</td>
<td>Strengthening link between M&amp;E and strategic planning</td>
</tr>
<tr>
<td></td>
<td>Accuracy of institutional arrangements</td>
</tr>
<tr>
<td></td>
<td>Inspiration from guidelines of Chilean model</td>
</tr>
<tr>
<td></td>
<td>Reinforcement of monitoring function</td>
</tr>
<tr>
<td></td>
<td>Elements that could be summarized or eliminated from guidelines document</td>
</tr>
<tr>
<td>Implementation of guidelines</td>
<td>Adjustment of procedures and times for preparing Indicator Matrix</td>
</tr>
<tr>
<td></td>
<td>Unification of training contents for drawing up Indicator Matrix</td>
</tr>
<tr>
<td></td>
<td>Reinforcing the evaluation market approach</td>
</tr>
<tr>
<td></td>
<td>Defining of role of CONEVAL</td>
</tr>
</tbody>
</table>

*Source: Drawn up by The World Bank, 2008.*

### Strengthening Link between M&E and Strategic Planning

The guidelines define the strengthening of the link between M&E and strategic planning at the national and sectoral level. The guidelines’ effort to reinforce evaluation within Mexican administration focuses strongly and mainly on the budget since it defines the program as the basis, exclusive unit of evaluation. As a result, consideration of other units of analysis for the evaluation of management, such as the National Development Plan, the sectoral plan and the organization is still low. Although the guidelines establish that the strategic objectives of the organizations and departments must be aligned with the priorities and strategies of the plans (both national and sectoral), they do not consider evaluations of these aggregated units of management. The evaluation focuses on programs, proposing various methods for them.

This perspective limits the scope of the evaluation, particularly its role as a validator of the quality of the planning exercises by measuring their results. The lack of systematic evaluation in planning exercises is one of the factors that explains the still weak nature of this function in several of the secretariats and federal administration, which also limits the introduction of the results-based management model promoted by SED and PMG.

It is recommended that the guidelines include the PND, sector and/or department in the guidelines as M&E units, to facilitate the aggregation of management units around sectoral and/or institutional objectives and plans. This would help lend order to the internal management of the sectors and departments regarding the plan and budget, while facilitating their coordination with the National Development Plan. This is an objective included in the Guidelines (Art.) 4) but whose contents are not developed due to the heavy emphasis on programs. A propos of this, evaluation methodologies such as the one for Consistency and Results, in addition to measuring the programs' alignment with sectoral plans could examine the consistency of sectoral plans with the National Plan or that of these two instruments with the budget exercises.

Achieving coordination between the National Development Plan, Sectoral Plans and the budget is one of the goals of SED and part of the efforts currently made by the President’s office. That is why CONEVAL could add a significant amount of value to the SED process by using the evaluation of consistency and results for this purpose, among others. Nevertheless, facilitating this task would require developing what some countries such as
Colombia, Brazil and Costa Rica have called Policy Options (a sort of sectoral, logical framework) which organizes management at the sector level by combining all management units, including programs, which facilitates their evaluation.

In this respect, one way of filling this gap would be to incorporate an indicative plan or sectoral logical framework matrix and an orientation in this respect into the guidelines that would encourage the Secretariats to establish an aggregation sequence in the management. This sequence could begin in the project and/or program as basic budgeting, including both the objectives of the sectoral strategic and its harmonization with the national development plan. This could become an important evaluation guideline that at the same time would help improve planning exercises, since the absence of this harmonization of objectives and goals makes it impossible to evaluate plans.

In this respect, evaluation appears as the counterpart of planning, since without any measurement of results, plans lose validity. This approach would also facilitate the swifter assimilation of the evaluation function in the sectors and departments insofar as it is compatible with other management units currently used by sectors and departments.

Having a framework of this nature would facilitate the development of the evaluation function and lend greater validity to the planning exercises. At the same time, this would make it possible to advance progressively, as a medium-term task, in the incorporation of a unified pragrammatic approach in planning and the budget.

**Accuracy of institutional arrangements**

As for the precision of institutional arrangements (roles, responsibilities and coordination mechanisms) presented in the guidelines, there is room to clarify the procedures for: i) the definition and validation of indicators, ii) the application of instruments (Indicator Matrix) and monitoring tools, iii) the coordination of sectoral plans with the National Development Plan and the budget and iv) explicitly defining the use they will be given by the Finance Secretariat, the Civil Service Secretariat and CONEVAL. It would also be possible to establish a specific role for the President's Office, which has undertaken tasks designed to reinforce planning exercises.

Although in this respect, the Second Heading of the guidelines is clear about its purpose (WHAT it wishes to achieve), it is not explicit about HOW to achieve it and above all, about establishing WHO will be responsible for consolidating the exercise, ensuring technical standards for all sectors and departments and controlling the quality of the exercise to define goals and indicators and the Indicator Matrix. As stated, the guidelines stipulate that there will be coordination between the SFP, SHCP and CONEVAL in these respects, but they do not define in what direction or which tasks each of them will carry out.

---

5 Validation of indicators refers to the acceptance and appropriation of the indicators and goals of the managers of the organizations and sectors (directors and secretaries) and their commitment to the achievement of results, so that the evaluation does not become merely a technical procedure but also produces policy consequences for programs and sectors.
They also state that the sectors must define their matrices and indicators and coordinate their plans with the PND. However, they do not specify which of them will be responsible for coordinating these tasks and technically support these programs or who or at what level they should technically and politically validate the indicators and goals. It would appear that the guidelines assume that the capacities for undertaking these tasks already exist and that they are the same throughout all sectors, as well as having the necessary technical and political validation before being uploaded onto the department's websites (Art. 5).

Due to the lack of a more precise definition of these aspects, confusion was observed among both the evaluators and the departments\(^6\). In some cases, there was even distrust regarding possible administrative and budgetary reprisals from the SFP and SHCP in the event of poor performance. An enormous heterogeneity was also observed in the indicator matrices, which proves the dissimilar capacities existing in the institutions.

In response to these factors, it is essential to define the institutional roles and responsibilities of the Civil Service and Finance Secretariats and CONEVAL (for the social sector) at the external level, and for the secretariats’ planning and evaluation guidelines at the internal levels, as well as of the Secretaries and public managers in the political validation of the exercise at the highest level of the departments. This involves a clearer, more detailed definition of:

- Procedures for the technical formulation of the indicator matrix and its approval (both internal and external);
- The users of the evaluations (SHCP, SFP, CONEVAL, the sectors themselves, Congress, the President, other secretariats, etc.);
- A modulating entity (both internal and external) responsible for consolidating the exercise of defining goals and indicators, defining a standard and controlling the quality of the definition made by the sectors or programs\(^7\) and,
- The use to be made of the evaluations, especially by SFP and SHCP.

The guidelines should also the management purposes of the evaluation, in other words, its role in encouraging the continuous improvement of the programs and sectors. There should also be a reduction in the links (such as the one suggested in Art. 7 of the Evaluation Guidelines) with an administrative responsibilities approach that can be demanded within the civil service sphere. This may require establishing a clearer

---

\(^6\) This was borne out by the high degree of confusion and frustration expressed by the employees of the various Public Education Secretariat programs regarding the definition of the matrix of indicators in the discussion workshop carried out by the Secretariat with the presence of World Bank and CONEVAL officials on January 15 2008. Similar observations were made by employees of the other Secretariats interviewed.

\(^7\) International experience is full of examples showing that a trained, experienced modulator is needed to orient the M&E systems. In Latin America, the cases of Brazil, Chile and Colombia are illustrative (see Table 2).
separation between the Civil Service Secretariat's performance evaluation and auditing function at both the conceptual and operative level.

The auditing function, indispensable in administration, corresponds to a different approach from results-based management. This approach often prevents departments from setting ambitious goals and relevant indicators.

**Inspiration from Guidelines in Chilean model**

The Chilean model is an excellent example of good practices in management by results and fits in very well with the purpose of the guidelines. However, a better adaptation of this model to the Mexican context could be achieved. In particular, aspects of Mexican institutionality such as the existence of a National Development Plan and the role played by the President's Office in this process, the federal dimension and structure of the public sector, especially the technical and institutional capacities existing in the sectoral departments, particularly in the Finance and Public Credit Secretariat, could be better reflected in their contents.

The capacity for response of the Finance Secretariat and sectorial Secretariats to the demands of the results-based management model proposed in the Budget law should be reinforced. This is because, unlike Chile, in Mexico, Congress plays a more important role in the budgetary process while the Budget Office does not have the same technical capacity to advance towards the follow-up and evaluation of expenditure. This could require the creation of a technical monitoring and evaluation unit at the Head Budget Office of the Finance Secretariat (like the one that exists in Chile) and the development of a plan of action that would technically strengthen the Head Offices of Planning and Evaluation of the Secretariats as modulators of the M&E functions within the sectors and as the main technical interlocutors of CONEVAL and the Finance Secretariat in this issue. Once again, this would involve defining the role and scope of CONEVAL, the Finance Secretariat and the Head Offices of Sectoral Planning in relation to the M&E system more precisely.

Since CONEVAL cannot legally operate as an evaluation regulator in APF, it should work mainly on the precise definition of its role in the social sector and on the division of labor with the Secretariats. According to the Secretariats, it is still unclear what role CONEVAL will play in the evaluation and monitoring processes in the sector and what specific tasks they should assign to it.

According to Federal Administration, CONEVAL should act as a leader in the provision of technical support and advice to SHCP in the processes of developing SED and its positioning will determine the extent to which this Secretariat and the other sectoral

---

8 The auditing function refers to controlling compliance with the administrative and financial procedures in the norms from which the legal processes and administrative responsibilities of the officials responsible to the respective authorities are or can be derived. The performance evaluation function within the parameters of New Public Management, for example, refer to the need to incorporate managerial practices into the public sector which, based on “transparent” incentives rather than legal or administrative provisions enable the expected results to be achieved with the resources assigned in public interventions.
bodies will regard their contribution as having added value. The specialized knowledge of the CONEVAL advisors and employees on the issues of evaluation and their independence place CONEVAL in an exceptional position to act as a technical voice authorized in this issue.

However CONEVAL should strategically identify the role it wishes to play. A propos of this, since SED will require a technical organization to regulate, coordinate and orient the tasks for generating information on the performance of all the Secretariats, and given that the model implemented by CONEVAL is innovative and functional for this purpose, efforts should be made to persuade the Finance Secretariat and the President's office to expand its mandate.

At present, neither the Secretariats nor the Federal Authorities have definitions of an office in charge of coordinating the M&E tasks. At the level of guidelines, the choices for filling this gap are the Head Planning Offices, for the internal level of the sectors and entities and at the external level, the Finance Secretariat (as in the case of Chile) or the President's Office (as in the case of countries that have Presidential Goal Systems).

Unsuccessful attempts were made to implement this last option in Mexico and the Finance Secretariat does not as yet have sufficient technical capacity in this area. For this reason, a suggested alternative would be for this function to be carried out by CONEVAL, through a technical assistance agreement with the Finance Secretariat. This alternative also has the advantage of enabling CONEVAL to assume the task of the technical orientation of M&E at the federal level ad hoc.

As for the specific monitoring function, CONEVAL obviously has a mandate to act as a governing organization and as such, is called to regulate the function, for which it will establish the guidelines, technical procedures and standards to be followed by the Secretariats. However, assuming the operative function of monitoring the results of all the programs will entail certain risks that should be considered. Firstly, given that the monitoring corresponds to daily activities related to the management of the executive branch's programs, this might constitute a high operative burden in response to which it might be necessary to expand CONEVAL's technical team (given the high number of programs).

Another aspect to be taken into account is that in countries where monitoring is carried out centrally, this task is associated with units aggregated around the budget, when they are carried out by the Finance Secretariats (as in Chile) or around the National Plan, when it is carried out by Planning Secretariats, (as in Colombia, Costa Rica and Indonesia). This is obviously not a requirement and CONEVAL could assume this task for the social sector, but it would require a high level of integration with the planning and budget functions of the respective secretariats.

Given that the sectors have limited information for monitoring sectoral indicators and since these are internal tasks for the departments for the permanent generation of information for handling their programs, this work could correspond more to the M&E programs and offices of the various Secretariats, which must assume responsibility for
the veracity of the data. If CONEVAL assumed some degree of responsibility for the quality of the data, this could affect its credibility and independence. This is because this information is continuously subjected to different kinds of political questioning both inside and outside government.

Nevertheless, if the Council considers that it is important to participate actively in this field, an intermediate option would be the consolidation of the information on the monitoring from the Secretariats, on the basis of which periodic reports on results could be prepared on the basis of the monitoring information. Another important consideration is that every effort must be made to prevent the duplication of efforts and tasks with the Secretariats which, in any case, would have to monitor their programs to comply with SED requirements.

Within the model currently being implemented, it would be advisable to clarify the role as regulator and provider of technical support that CONEVAL should play as regards monitoring and the operative role to be played by SHCP and the other Secretariats. CONEVAL could promote the development of monitoring systems in the various sectors, as in the case of SEDESOL, accompany the Secretariats in their technical tasks, provide training and evaluate and certify the processes of the generation, administration and use of the monitoring information, among other aspects that will encourage the proper development of this practice and the incorporation of suitable tools. To this end, CONEVAL could have an Information System for consolidating data from the monitoring of the sectors, on the basis of which it could supervise their quality. This, however, could require tasks currently not considered by CONEVAL.

Reinforcement of Monitoring Function

Due to their recent incorporation, the guidelines emphasize the evaluation function and the still low development of the monitoring function, meaning that there is still an imbalance in the efforts to promote a results-oriented management approach. In practice, due to the still limited regulations, monitoring is carried out by sectors and programs with a restricted definition of standards and specific technical procedures for their development (Fifth Section).

One of the objectives of a performance evaluation system is to manage to strike a balance and achieve a strict complementarity between evaluation and monitoring. Follow-up is intended to provide continuous, timely information for improving the programs, programming the budget and ensuring permanent accountability by government. It constitutes a managerial need for the programs and the modernization of the public sector.

In view of this, it would be advisable to develop monitoring guidelines (Chapter 1, Fifth Heading) to establish a more balanced, harmonious framework between them and the evaluation, planning and budget functions. This implies incorporating procedures, standards, persons responsible for the process and internal and external coordination mechanisms for carrying out this function, as well as express rules for the public presentation of results.
CONEVAL can play an important role in providing technical support for SHCP, the President’s Office and the Secretariats in defining these aspects. At the same time, the guidelines should specify the responsibilities of the programs and the sectoral Head Offices of Planning in this issue, as well as the coordination mechanisms with the governing body defined for this function and with CONEVAL for the purposes of evaluation in the social sector. Defining a structured guide to complement the guidelines in this respect is contemplated as a second phase in the process of supporting CONEVAL.

Elements that could be simplified

The guidelines establish a greater scope for the evaluation function and introduce a series of innovations associated with their practice. One of the most important of these is the link established between evaluation and the budget and planning functions, an aspect that would require the incorporation of new elements such as the definition of coordination bodies, new expense classifiers in the sectors for M&E purposes, schedules for producing reports and greater development in some of its contents.

Although the analysis shows that in general, all the elements they contain are of value for the departments and help to arrange the development of the evaluation, if the aim is to simplify and increase the effectiveness of the guidelines, some of its contents could be handled separately, such as the rulings or technical appendices issued exclusively by CONEVAL.

This is only recommended in the event that the modifications make the document extremely long, meaning that it might be necessary, among other things, to expand some of the chapters (such as the one on monitoring) and/or incorporate new issues (such as integration with SED and SIMER). In response to this situation, we suggest concentrating the contents of the guidelines on basic definitions, policy aspects, institutional arrangements and the procedures that strictly imply for the departments a mandate that requires a higher level of hierarchy than that of CONEVAL.

In this respect, since the methodological definitions of Chapters I to V of the Third Heading do not entail any of the previous elements, they could be replaced by a general article indicating that the evaluation typologies defined in Article 15 should be carried out in accordance with the technical criteria defined by CONEVAL.

Adjustment of Procedures and Times for Preparing Logical Framework

As for the implementation of the guidelines, as shown by international experience, the incorporation of institutional changes of this nature forms part of a process on learning along the way, which takes time. It is therefore common to find that these processes contain both good and bad decisions, meaning that they usually entail frustration and tiredness in the states.

In this respect, it is advisable to take states’ perception into account regarding the process carried out to date in order to ensure that the introduction of new tools and practices is useful, particularly for the states and sectors. Unless this happens, it might be difficult to
achieve its appropriation and commitment, which will affect the institutionalization and consolidation of the model.

In this respect, it is useful to consider that the programs' demand to draw up the indicator matrix in a relatively short period of time entailed enormous pressure for certain responsible teams (given the unequal capacity and low knowledge of certain sectors regarding the procedures for drawing up the logical framework). Moreover, the change in rules initially established, due to the acceleration of the process by the SHCP, in addition to the message that the indicators would be used in the 2008 process of budgetary planning, raised enormous expectations.

Two lessons can be drawn from this process: first of all, it is important to bear in mind the technical capacities of the programs and their capacity for response to novel technical exercises such as these. In particular, they are designed to prevent these exercises from being perceived as an exaggerated burden in relation to their usefulness. Secondly, perverse incentives that send different messages on different fronts should be avoided, since they can cause uncertainty and lack of credibility in the departments.

In the implementation of the guidelines, in the opinion of the departments and programs interviewed, the request to put forward the matrix exercise in some cases such as Sedesol interrupted a gradual process of drawing up the Logical Framework that was being carried out on the basis of a reasonable schedule. In other cases, it implied a new attitude for which many departments were not yet prepared. This forced the departments to comply with the exercise but in many cases, made them underestimate their goals in order to avoid commitments that would be difficult to achieve. They also did this to avoid budgetary or administrative punishments that could be incurred by potential failure to comply. In certain cases, this lowered the standards of the logical framework exercise.

The expansion of the framework of the questions in the instrument to comply with the new information requirements for the SCHP sometimes meant that the programs ended up with several matrices in the logical framework. At the same time, the unclear use of the exercise for the goals set appears to have created frustration and a lack of credibility regarding the instrument among the departments. Thus, an excellent initiative in practice, implemented too quickly, led to consequences in terms of credibility, confusion and frustration in the departments' view.

In response to this context, it would be advisable to act strategically to reverse the processes of loss of credibility, confusion and frustration caused to date. This could be achieved by:

- Reaching an agreement with SHCP and SFP on the purpose, a precise procedure and one set of rules for the processes of adjusting or redefining the MML and publishing them as soon as possible by issuing new instructions to the departments⁹;

---

⁹ This proved effective in Colombia following the confusion originally generated by the President's Office through the introduction of SIGOB for example.
Defining a reasonable period of time (six months or a year) for improving the indicator matrix, bearing in mind that this is a dynamic process in which the matrices must be modified as the programs adjust.

Revising and adjusting the indicator matrix instrument and complementing it with an instrument for the addition of objectives and goals at the sector and department level so that it could be easily assimilated into the department's everyday practice and

Clarifying the use SHCP and CONEVAL are making of the logical framework matrices drawn up by the departments so that they do not feel that their efforts were in vain.

Unification of Training Contents to Draw Up Logical Frameworks

The quality and relevance of the department's training process obviously largely determines the success of the implementation of the evaluation guidelines and their instruments. This process requires a single conceptual framework (a single discourse) and a standardized method within the learning-by-doing approach. The school method in sectors with widely differing, generally low capacities has proved ineffective in practice in achieving the quality required.\(^\text{10}\)

Although training may be provided by several agents, it is essential for the contents to have been previously defined and unified by a single governing body. Nevertheless, if possible, although expensive, it would be ideal if a single trainer were responsible for the process, to ensure quality. This is an important aspect to be considered since the quality of the definition of the MML and the indicators will determine the quality of the M&E exercises and their credibility.

In practice, the process of providing training in the methodology of the logical framework assigned to various departments revealed the need to have a single method with the same concepts to avoid confusion in the departments.

The training scheme should be reinforced with a "learning-by-doing" approach, which implies providing training by drawing up or improving the final logical framework with the departments. This scheme has a high cost-benefit ratio, since the quality of the logical framework will determine the results of the management and evaluation.

It is also necessary to define a single body to be responsible for the methodology and validating the contents, so that training departments could be hired to spread this method rather than any other, however good it may be. Although this approach has already been used in some of the training processes, it would be useful to extend it, particularly to the process of definition or adjustment of the Indicator Matrix.

\(^{10}\) This is the case of Costa Rica, Indonesia and Colombia in the implementation of follow-up tools.
Reinforcing the evaluation market approach

The institutionalization of the evaluation function in Federal Public Administration requires promoting a proper balance between supply and demand in evaluation. On the demand side, when the department’s technical capacity for evaluation is low, they are often unclear about its use, place little importance on it in management and are not in a condition to carry out a proper intervention with the evaluators.

On the supply side, Secretariats such as SGARPA and SEDESOL for example and some of the evaluators interviewed say that there are still a low number of qualified evaluators, which in their view, limits the development of the function and at the same time, produces poor evaluations in many cases. This aspect is common to other Latin American countries, since certain evaluation techniques require specialized knowledge and the dissemination of the latter within the academic community, which takes time to develop. Both aspects produce evaluations of an uneven quality, in which the results often have little influence on the definition of programs and the assignment of resources. As one can infer from this, these aspects contribute to reducing the quality of the expenditure on evaluation.11

On the demand side, auditing of the evaluations should be carried out by the Head Offices of Planning in each sector and department so that in each case, there will be a department for which evaluation and the quality of the latter are a priority, in addition to having the necessary capacity for this. Since programs are permanently involved with executive aspects, which often exceed their capacity, even though evaluation is important, it is the last priority on their agenda. This may require reinforcing the task of providing evaluation training through courses, workshops and stays targeting strategically selected sectors and technical officials.

On the supply side, incentives should be incorporated into the hiring process to build capacities in existing evaluators and increase the number of the latter. This could make it necessary, during the contracting process, to explicitly reward the creation of temporary alliances or unions between national and international organizations or professionals from different disciplines, to encourage the transfer of skills in evaluation and expand the range of potential evaluators. These aspects can be included in the evaluation guidelines issued by CONEVAL once they have been harmonized with the Law of Public Hiring.

One could add a focused plan of action regarding training, preferably coordinated by a governing body in the area (CONEVAL in the social sector and SCHP in the rest, while ensuring coordination between them) which should include an annual agenda of courses, workshops and seminars designed to build the capacities of both supply and demand. Once again, CONEVAL could assume these tasks for all of federal administration, within a cooperation agreement with SHCP.

The training function could become one of CONEVAL’s main tasks as the governing body of evaluation and as a promoter of the development of the skills required for the

11 SAGARPA declares that it assigned approximately $60 million USD.
in institutionalization of this function within quality parameters. In this field, CONEVAL has a significant comparative advantage and could add a high value to the sectors and the Finance Secretariat given its capacities and its role as an independent organization, thereby defining specific areas within its sphere of competence.

Defining of role of CONEVAL

As mentioned earlier, given that nowadays the scope of the general evaluation guidelines exceeds CONEVAL's sphere of competence, which is restricted to the social sector, it is essential for CONEVAL to identify specific areas within this sphere. It is also necessary to determine the functions of the other key actors (SHCP and SFP) and incorporate the President's Office as one of the main users of the results of the evaluations in order to report on the decision-making processes to the highest level of government.

Although due to its technical knowledge, CONEVAL could play a key role in promoting a coordination solution that would make it possible to define a single frame of reference for the action of all sectors in M&E, this aspect exceeds both its scope and current capacities.

CONEVAL's tasks within the sphere of its competence within the framework of the Law of Social Development should be made more specific. For example, as regards the orientation of the management by results model in its sector or in relation to the impulse and technical support of the M&E systems in the secretariats under its authority. These exercises would make it possible to consolidate the sector as a pilot and model for federal administration as a whole. During this process, CONEVAL could provide the necessary technical accompaniment for states in the sector so that they could adequately meet the requirements of the President's Office and the Finance and Civil Service Secretariats.

This would also involve maintaining continuous communication with these departments in order to ensure the harmonization of the sectoral model with the management by results framework of central administration, which may prove beneficial for the latter by using this means to take advantage of CONEVAL's knowledge of the matter.

Areas Requiring Further Analysis

Given that the indicator matrix constitutes a new instrument for planning programs and seeks to become the main frame of reference for their evaluation, it is essential to determine the extent to which the contents and purpose of this new tool could be duplicating or complementing the Programs' Operating Rules by establishing the extent to which they could be harmonized. This task, however, requires greater analysis and forms part of the tasks to be undertaken by the department or within the framework of a process of support that would require more time.
INTRODUCTION

The purpose of this report is to analyze the implementation of the General Guidelines for the Evaluation of the Federal Programs in Federal Public Administration (APF), (henceforth referred to as: Guidelines) issued by the National Council of Evaluation of Social Policy (CONEVAL), the Civil Service Secretariat (SFP) and the Finance and Public Credit Secretariat (SHCP) published in March 2007. It provides an analysis of its implementation, with emphasis on the management tools contained in the Guidelines, within the implementation framework of the Performance Evaluation System (SED).

The document is divided into three sections. The first contains a general reflection on the process of implementing the Guidelines in relation to their scope, process and results; the second undertakes a diagnosis of the management tools implemented while the third provides recommendations for making adjustments to the Guidelines. Lastly, the appendices are presented.

It is not possible to draw inferences from the findings since the sample of programs was not the result of a random selection or independent from the programs, and is therefore not representative.

GENERAL GUIDELINES FOR THE EVALUATION OF THE FEDERAL PROGRAMS OF FEDERAL PUBLIC ADMINISTRATION

The Performance Evaluation System (SED) currently underway includes performance evaluation at three levels: programs, institutions and sectoral policy. The program level is organized through the General Program Evaluation Guidelines, while the institutional level (or that of public agencies) is mainly organized through the Management Improvement Program and the sectoral and policy program is in the process of being defined within the consolidation of SED.
### TABLE 1: Performance Levels Evaluated by SED

<table>
<thead>
<tr>
<th>Performance levels</th>
<th>Instruments</th>
<th>Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sectoral policy</td>
<td>Some mentions in the General Evaluation Guidelines and Annual Evaluation Program led by CONEVAL. Indicators of Federation Expenditure Budget (PEF) at the sectoral level.</td>
<td>Defining coordination of the sectoral level and the National Development Plan.</td>
</tr>
<tr>
<td>Institution level and Public Federal Administration offices (APF)</td>
<td>Management Improvement Program led by Civil Services Secretariat and Implementation Guidelines. PEF institutional indicators.</td>
<td>Led by SFP, by agreement with SHCP.</td>
</tr>
<tr>
<td>Program</td>
<td>General Evaluation Guidelines for APF Federal Programs (CONĖVAl regulates their application in the field of social programs).</td>
<td>Led by CONEVAL, by agreement with SHCP and SFP.</td>
</tr>
</tbody>
</table>

Source: Drawn up by The World Bank, 2008.

It is important to advance in the implementation of performance evaluation mechanisms at the three levels to have elements that will permit an integral performance analysis and establish links between the various levels to create synergies and better input for supporting the decision-making process.

However, it is unusual to find integral coverage of the three levels regarding performance evaluation in national Latin American and Caribbean evaluation systems. For example, the Argentinean model focuses on the program level through the operation of Social Program Monitoring and Evaluation Indicator Systems (SIEMPRO) and its Physical/Financial Follow-Up System (PFMS). In Uruguay, the systems cover the institutional level through the Results-Based Evaluation System (SIEV). Colombia uses the National System of Evaluation of Public Management (SINERGIA) to deal with the sectoral policy and program levels while in Costa Rica, the National Evaluation System (SINE) covers the institutional and program level.

### TABLE 2: Performance Levels Evaluated by Monitoring and Evaluation Systems in Certain Latin American Countries

<table>
<thead>
<tr>
<th>Level</th>
<th>Argentina</th>
<th>Chile</th>
<th>Colombia</th>
<th>Costa Rica</th>
<th>Uruguay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sectoral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Institutional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Program</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


---

12 This could imply expand CONEVAL's current mandate and functions.
The Program Level

The purpose of the Guidelines, according to the Official Federation Gazette (DOF) is to regulate the evaluation of federal programs, draw up an indicator matrix and monitoring systems and establish the strategic goals of Federal Public Administration offices and departments\(^{14}\).

The Guidelines are therefore located at the federal program level. CONEVAL coordinates and regulates their implementation in the social sector, which is why the challenge is the total coverage of federal programs and their integration at the institutional and sector levels.

Overall, the implementation of the guidelines achieved a large number of evaluations of social programs by hiring external evaluators. APF organizations and offices also drew up indicator matrices. This dynamic also permitted the drawing up or adjustment of the offices’ strategic objectives. Nevertheless, monitoring systems were barely drawn up or developed.

The Federation Expenditure Budget (PEF) approved a total of 330 indicators corresponding to 208 budgetary programs\(^ {15}\). Each budgetary program (Pp) has an Indicator Matrix (MI) from a federal program of those subject to operating rules and other subsidy programs obliged to present one.

Program evaluation is carried out through the application of various types of pre-established evaluations in the Guidelines and according to the general orientation of the Annual Evaluation Program (PAE) led by CONEVAL. However, for several types of evaluations, the management tools to be applied and the subsequent Terms of Reference (TDR) have yet to be developed. Nor do the Guidelines specifically describe the links or sequence of application of the evaluations. It is therefore necessary to determine, either in the Guidelines and/or the Annual Evaluation Program: how many, which, in what order and with what frequency, how and for what purpose the various evaluations in federal programs are applied as well as their usefulness and expected impact on improving the programs' performance and impact.

At the program level, two dimensions must be evaluated to determine the degree of success achieved: i) the program’s performance and ii) the impact achieved by the program. At the policy level, the Guidelines contemplate carrying out the evaluations by analyzing the contribution of a program or set of programs to a strategy, policies and institutions.


Program Performance

According to the Guidelines, federal program performance is evaluated through the study of six specific dimensions: design, strategic planning, coverage and focusing, operation, perception of target population and results, which partly reflect a program’s cycle of management, from the identification of a problem or need to deal with it to the identification of its net results in solving them.

Performance evaluation at the program level is carried out through the administrative hiring processes of external evaluators, which may be universities, research centers or consultancy firms through the implementation of the Terms of Reference (TDR) corresponding to the evaluation of Consistency and Results (ECR) which specifically determines the variables and aspects to be evaluated for each of the dimensions pointed out earlier that constitute its performance.

The Evaluation of Consistency and Results becomes the managerial tool for measuring the performance of federal social programs. It is on the basis of the findings of the latter that another type of specific evaluations could be activated to undertake a more detailed investigation of the low results of a certain dimension or find out more about the conditions in which very good results were obtained in order to identify best practices that could be repeated by other programs in the sector or for any program.

According to this logic, ECR findings could reveal the need to investigate, for example, aspects concerning program design and operation as well as the quality of their information. It is through these findings that the implementation of design, process and indicator evaluations should be carried out.

In the case of the United States, there is an instrument that achieves the same goal as the ECR of evaluating federal programs' performance. It is known as the Program Assessment Rating Tool (PART) and is administered by the Office of Budget Management (OBM), the difference being, in the implementation process, that PART is a more agile instrument in terms of having a quarter of the questions of those in ECR. Moreover, it is not undertaken by external evaluators but through a self-evaluation of the offices themselves with the support of PART technicians. Once performance has been evaluated and the areas requiring greater investigation on the basis of the findings have been identified, then specific evaluations will be carried out by resorting to external evaluators although not for all the programs or simultaneously.

For its part, the program impact evaluation will be responsible for using “rigorous methodology to evaluate the change in the indicators at the results level that can be attributed to the implementation of the federal program.”

---

16 These Terms of Reference can be consulted on the CONEVAL Website: www.coneval.gob.mx
The Annual Evaluation Program should determine all the strategic and impact evaluations that will be carried out annually in specific federal programs.

Table 3 shows the scope of the various types of evaluation contained in the Guidelines to be implemented at the program level, while Appendix 1 contains a graph showing the existing link between the various types of evaluation.

**TABLE 3: Types of Evaluation of Guidelines**

<table>
<thead>
<tr>
<th>Type of Evaluation</th>
<th>Who measures and who is it applied to?</th>
<th>How often?</th>
<th>Who does so?</th>
<th>How do they do it?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design evaluation</td>
<td>New federal programs or those with a poor design performance. It includes the identification of the problem, the program’s contribution to the department’s strategic objectives, etc.</td>
<td>For new programs that emerge during the application of the guidelines or programs with poor performance, according to the ECR results.</td>
<td>External evaluators.</td>
<td>Through the Terms of Reference of the TDR1 Design Evaluation. Office work.</td>
</tr>
<tr>
<td>Evaluation of Consistency and Results (ECR)</td>
<td>All federal programs with operating rules and other selected programs. It is the main evaluation for measuring performance and includes the study of: design, strategic planning, coverage and focusing, operation, perception of target population and results.</td>
<td>According to what is established by the National Evaluation Program (there is now a baseline due to the evaluations carried out to date).</td>
<td>External evaluators.</td>
<td>Through the Terms of Reference of the TDR2 Design Evaluation. Office work.</td>
</tr>
<tr>
<td>Evaluation of processes</td>
<td>Since its sphere is the production and delivery of goods and/or services in the program, it measures the efficiency and efficacy of their operating processes. The operating aspect of ECR evaluation should be applied to programs that have achieved a low performance in the operating sphere.</td>
<td>To be defined in the 2008 Annual Evaluation Program.</td>
<td>External evaluators.</td>
<td>Through the TDR3 Terms of Reference of Process Evaluation (there is a proposal that will be implemented in a pilot experience). Office and field work.</td>
</tr>
<tr>
<td>Evaluation of indicators</td>
<td>Its sphere is the production of information on goods and services indicators and results.</td>
<td>To be defined in the 2008 Annual Evaluation Program.</td>
<td>External evaluators.</td>
<td>Through the Terms of Reference of the TDR4 Indicators Evaluation. (There is a proposal that will be applied in a pilot experience).</td>
</tr>
</tbody>
</table>

---

18 There are Annual Evaluation Programs led by CONEVAL that determine the various evaluations to be carried out every year, what programs they should be applied to, in what sequence and in most cases, through what methodology. However, since impact evaluations are multiannual, a long-term Strategic Evaluation Plan must be drawn up, rather than an annual one.

19 Details of the contents of this evaluation can be found in Diario Oficial de la Federación, (2007), “Lineamientos generales de evaluación de programas federales”. Marzo 30, título tercero, título tercero, capítulo segundo, artículo decimocuarto.
The Guidelines establish the process through which the administrative hiring of the programs' external evaluators will be organized. However, during the implementation of the Guidelines, various situations emerged that revealed the need to reinforce regulation within the sphere of administrative hiring.

**DIAGNOSIS OF MANAGERIAL TOOLS CONTAINED IN GUIDELINES**

From the managerial point of view, the application of the Guidelines includes various methodological approaches and implies a minimal capacity for monitoring and evaluation on the part of APF entities and offices. The methodological approaches within which their application can be framed can be grouped into four groups:

- Indicator Matrix (MI);
- Evaluations studies: design, consistency and results, processes, indicators, impact and strategy;
- Federal Programs' Operating Rules;
- Administrative hiring mechanisms

The Guidelines require a certain capacity for monitoring and evaluating the programs in APF offices and units to ensure their proper implementation. The Guidelines suggest the existence of an area responsible for this purpose.

The areas responsible for evaluation could, together with the directors of the offices and organizations, suggest to CONEVAL the programs that could be incorporated into the annual evaluation program, whose importance warrants an impact evaluation. They could also supervise the contracting of external evaluators, the results of their work and promote monitoring activities in the programs and the state or office in which they are located.
The relevant diagnosis and findings, together with specific recommendations for adjusting the Guidelines, are given below.

**Indicator Matrix**

The Indicator Matrix (MI) not only forms part of the General Evaluation Guidelines for Federal Programs but plays an important role in the Performance Evaluation System (SED) and the operating scheme of the Results-Based Budget. It is also an instrument that encourages the drafting or adjustment of Federal Programs' Operating Rules (ROPF).

The MI is a crucial element of the Guidelines, since it provides the main inputs for drawing up a preliminary plan for the Federation’s Expenditure Budget (PEF) and establishes the bases for determining the level of success achieved by a program.

The MI is a planning and programming tool based on the methodological approach of the EML logical framework, which some decades ago became an international standard for preparing projects.  

---

There are slight variations in the method according to the agency that applies it. Training in the use of the logical framework led by CONEVAL is based on the proposal of CEPAL’s Latin American Institute of Economic and Social Planning (ILPES). This can be consulted in ILPES, (2004). *Metodología del Marco Lógico*. Santiago de Chile: ILPES, CEPAL. October 2004
Process of Drafting Indicator Matrix

The logical framework approach is a method known to most top-level government employees in APF organizations and offices and federal program directors, especially those linked to cooperation and welfare in multilateral organizations. However, since it has not become regular practice or a planning method, knowledge and skills training is required for its preparation. Logical framework training was carried out through the Latin American Institute of Economic and Social Planning (ILPES) to establish the conditions to draw up a suitable Indicator Matrix.

On the basis of their experience and the training received, the APF offices and organizations began drafting their MI in 2007. During this process, in order to include performance indicators in the Federal Expenditure Budget Project (PEF 2008), the Finance and Public Credit Secretariat demanded the delivery of Indicators Matrices with a deadline for delivery and specific demands, which created a certain amount of confusion and a variety of expectations among some of the federal program operators interviewed.

It is worth noting that the Indicator Matrix served as the basis for delivering information to the SCHP, which must establish the equivalence between the program’s goals and objectives and the specific allocation of resources or what are known as budgetary programs. The establishment of new budgetary classifications based on the unit of programs is one of the greatest obstacles found internationally.

In the case of Mexico, the objective of the results-based budget is precisely that MI is directly linked to the Pp budgetary programs. In other words, Pp is a category and programmatic element directly linked to the MI of federal programs. This is due to the recent changes in the programmatic classification.

---

21 In the case of SEDESOL, in 2004, its programs participating in the development of the System of Evaluation and Monitoring Based on Results (SEMBR) were given training in the methodology of the logical framework through CEPAL’s ILPES.

22 A meeting was held with the federal programs subject to the operating rules of the Public Education Secretariat. The meeting was led by the Head Office of Planning and Evaluation, during which several persons responsible for operating the programs expressed confusion and direct expectations about the process and the consequences of drawing up the MI. In other programs such as the DICNSA Food Support Program, a sectorized SEDESOL organization, employees also said that the process and objective of drawing up MI was not very clear.

GRAPH 1: Link between MI and budget

<table>
<thead>
<tr>
<th>Categoría y Elementos programático</th>
</tr>
</thead>
<tbody>
<tr>
<td>Función</td>
</tr>
<tr>
<td>Subfunción</td>
</tr>
<tr>
<td>Programa</td>
</tr>
<tr>
<td>Actividad Institucional</td>
</tr>
<tr>
<td>Programa Presupuestario</td>
</tr>
<tr>
<td>Indicador</td>
</tr>
</tbody>
</table>

Matriz de Indicadores

Source: Drawn up by The World Bank staff, 2008.

In order to comply with their obligations derived from the General Evaluation Guidelines oriented towards the MI, APF federal programs must follow the two Guidelines described in Table 4

TABLE 4: Indicator Matrix and Budget

<table>
<thead>
<tr>
<th>Methodologies</th>
<th>Guidelines</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methodology for drawing up indicator matrices</td>
<td>General evaluation guidelines for APF social programs</td>
<td>Indicator Matrix</td>
</tr>
<tr>
<td>Methodology for drawing up indicators of APF budgetary programs.</td>
<td>General Guidelines for Programming and Budget Process for 2008 tax year.</td>
<td>Budgetary program indicators</td>
</tr>
</tbody>
</table>

Source: Drawn up by The World Bank, 2008.

Diagnosis of Implementation of Indicator Matrix

Positive aspects of the process of drawing up MI include the following:

- The establishment of a basic managerial language linked to the concepts and tools of evaluation among civil servants linked to federal programs;

- A good general knowledge of the usefulness of the logical framework as a planning instrument, even though there are difficulties in applying the method correctly;

- Program directors focus on results-based management and the awareness that a social program justifies its existence, not only because of the delivery of goods and/or services it provides but also because of the results it achieves in improving the conditions of well-being of its target population.

- A discussion focusing on federal program operators, the identification of precise indicators and the information requirements to enable them to exist. Program operators are beginning to be aware of the costs of achieving new information;

- The programs’ reflection on their contribution to sectoral goals and sometimes to showing whether or not these goals have been defined and validated;
• An initial exercise of connection between the Indicator Matrix and the drafting of budgetary program indicators, which establishes a link between programming and budgeting.

The diagnosis also made it possible to identify various opportunities for improvement.

The exercise of drawing up the 2007 MI by the program operators was abruptly interrupted by SHCP’s request to deliver Budgetary Program Indicators to incorporate them into the 2008 PEF program. This caused conflict in some of those responsible for operating the programs for the following reasons: i) the belief that the indicators delivered would be used to increase or reduce the budget; ii) fear that the indicators under pressure of time and in the required format did not faithfully reflect either the reality of the program or what one would wish to achieve the following year.

Some of those responsible for the programs confused the meaning of the recommendations regarding their indicators matrices made by CONEVAL and other organizations. They felt that they questioned the program’s reason for existence and were concerned about its possible elimination. They failed to perceive the recommendations as methodological observations aimed at ensuring the correct use of the planning and programming tool.

This created the existence of various versions of indicator matrices. The first was the MI the programs began to draw up as from March 2007 (MI1) and the second was the one delivered to SHCP; which corresponds to the budgetary program indicators (IPP2) requested in August 2007 and the MI completed after having delivered what was required to SHCP (MI3), which provides a better reflection of the program for the following year. Moreover, the program’s current operating rules do not correspond to any of the matrices, since they must adjust to the final versions of the MI.

Some federal programs experienced difficulties in drawing up goal and even purpose indicators, for a variety of reasons: i) they lack suitable sources of information for these indicators (surveys, administrative records, etc.); ii) they lack total clarity about the sectoral goals to which the goal should contributed and iii) in some cases, the authorities of the states and offices on which the program dependes failed to carry out an effective process of technical and political validation. It is important to note that some of these situations are not suprising, since given the nature of the MML approach, the cost, time, technical complexity and skills required for data collection rises with vertical level of the indicators (Binnendij, 2001).

The process of validation of both the MI and the budgetary program indicators must be clearly established. The SHCP for example, proposes the following scheme, which, according to its objectives, enables the Head Offices of Programming and Planning to play a key role. However, the Head Offices of Monitoring and Evaluation should also play an equally important role, meaning that the final responsibility should be shared.

---

TABLE 5: Validation Process of MI and PEF Indicators

<table>
<thead>
<tr>
<th>Actor</th>
<th>Responsibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit Responsible for Operating the Program, Head Office of Programming, Organization and Budget (DGPOP)</td>
<td>Initial register and validation.</td>
</tr>
<tr>
<td>Competent authorities (SHCP, SFP, CONEVAL, INMUJERES, etc.)</td>
<td>Review and issuing of recommendations.</td>
</tr>
<tr>
<td>Head Office of Programming, Organization and Budget</td>
<td>Makes final adjustment</td>
</tr>
<tr>
<td>SHCP</td>
<td>Incorporation of indicators</td>
</tr>
</tbody>
</table>


The identification of goal and purpose indicators in federal programs that are not in the social sector is complex, given the nature of the rest of the APF programs, for example, in regulatory programs or public property such as anti-monopoly, cultural promotion or national defense programs.

Another aspect identified in the process of drafting MI is that some programs do not know whether the MI should reflect what the program does nowadays (fully reflecting its operating rules) or what the program seeks to achieve in the future, in other words, in some cases, it is still not acknowledged as a planning tool. In some cases, there is confusion about who the beneficiaries and intermediaries are.

Likewise, the MI of programs of various APF organizations and offices were reviewed and analyzed, yielding results that varied widely in quality. They range from one extreme of extremely well-written indicator matrices, on which one can make specific, insignificant observations on how to improve by specifying the assumptions identified or the attributes of the information sources required (geographical, thematic coverage) and the variables or modules of sources of information to MI with serious problems of consistency in their narrative that invalidate the entire instrument. These MI include unclearly defined goals and purposes, poor writing, components that are not linked to the purpose and activities disconnected from the components, among others. Table 6 contains a summary of the main problems found and their frequency.


SAGARPA, National Water Commission, SEDESOL-sectorized and unsectorized entities–Education Secretariat, Health Secretariat, Economy Secretariat and SEMARNAT.

The review of the comments and suggestions of CONEVAL’s technical team on MML confirms the findings on the weaknesses of the use of the method.
### TABLE 6: Evaluation of Drafting of Indicator Matrices

<table>
<thead>
<tr>
<th>Aspects</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problems Detected in Drafting the Indicator Matrix</td>
<td></td>
</tr>
<tr>
<td><strong>Type of problem</strong></td>
<td><strong>Aspects</strong></td>
</tr>
<tr>
<td>General problems</td>
<td><strong>Wrong use of conjugations</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Ambiguous, confused writing.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Poor use of numbering; for example of activities in relation to components (and/or lack of numbering).</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Extremely heterogeneous formats (Excel, scanned documents, type of font, length, extra column added, etc.)</strong></td>
</tr>
<tr>
<td>Problems with vertical logic</td>
<td><strong>No consistency between goal and purpose.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Program’s target population not precisely defined.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>No consistency between purpose and components.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Components with a very broad scope that need to be divided</strong></td>
</tr>
<tr>
<td></td>
<td><strong>No consistency between activities and components.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Confusion between activities and components.</strong></td>
</tr>
<tr>
<td>Horizontal logic</td>
<td><strong>Aspects</strong></td>
</tr>
<tr>
<td>Indicators</td>
<td><strong>Failure to specify any of the following elements: time, quality, quantity, location.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Lack of formula.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Lack of definition of goal.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Lack of definition of baseline.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>No grouping or identification of indicators related to criteria of efficacy, efficiency and economy established by SHCP</strong></td>
</tr>
<tr>
<td>Means of verification</td>
<td><strong>Aspects</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Not identified.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Over general identification.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Ambiguous identification, such as “Information will be provided by...”</strong></td>
</tr>
<tr>
<td>Assumptions</td>
<td><strong>Aspects</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Non-external conditions identified.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Unimportant conditions identified.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Identification of conditions that are unlikely or very likely to occur.</strong></td>
</tr>
</tbody>
</table>

*Source: Information from SAGARPA, SEDESOL, Education and Health Secretariats, SEMARNAT and Economy Secretariat.*

It is important to note that the matrices of federal social programs of a group of offices and institutions visited during the World Bank mission to the country and agreed on in conjunction with CONEVAL were analyzed. Some matrices were consulted on line with certain social secretariats, such as the external evaluations of the design of five SEDESOL social programs.

A sample design with a random, independent selection of the programs was not drawn up, the sample is not representative and it is not possible to draw inferences from the results found, since they only explain the set of matrices analyzed.
It is important to note that although the MI approach is often used for projects in an organization, it has been applied to enormous programs (which include various projects) and involve various institutions, making the inter-organizational complexity and the specificity of certain projects more difficult to reflect.

The issue of scale is not considered in the application of the instrument. The evidence from the MI reviewed indicates that some programs are extremely large. They contain lines of action that could be considered projects while other programs are simpler and smaller.

**TABLE 7: Main Advantages of Logical Framework**

<table>
<thead>
<tr>
<th>Advantages of Logical Framework</th>
<th>Practice observed in applications of Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ensures that decision-makers formulate fundamental questions and analyze assumptions and risks.</td>
<td>One could conclude that the MI has elicited a reflection among program operators and certain key decision makers about crucial questions in the programs, their assumptions and risks.</td>
</tr>
<tr>
<td>It involves important actors in the planning and monitoring of the process.</td>
<td>On the basis of the interviews, top-level decision-makers (important actors) are not always involved in validating MI. General speaking, operating levels draw up the matrices and when they reach the purpose and goal levels, they require the participation of actors with more decision-making power, which did not occur in every case. This should be rectified. On the other hand, the process does not usually involve the beneficiaries.</td>
</tr>
<tr>
<td>When used dynamically, it is an effective managerial tool for guiding implementation, monitoring and evaluation.</td>
<td>Although there is no definitive MI for the programs, it is essential to improve the monitoring capacity within the APF programs, organizations and offices.</td>
</tr>
</tbody>
</table>


Although this does not in any way require the modification of the norm, possible adjustments in the application of the methodology should be indicated. In fact, in its application, the MML distinguishes the aspects of programs and projects. In other words, the MML of a program implies drawing up a general MI and another one for each of the projects it includes, establishing a logical sequence in a cascade.

Table 7 prevents an evaluation of the achievements of the application of the logical framework in the drafting of the indicator matrix, according to the expected benefits of its application.

---

28 The MI does not require carrying out all the steps in the logical framework approach that begin with the identification of the problem and its causal development and the analysis of those involved, which gives a key role to the potential or actual beneficiaries of the programs or projects. However, the process of drawing up the Indicator Matrix could include focus groups or interviews with beneficiaries who provide their views on the pertinence, timeliness and quality that should characterize the infrastructure, goods and/or services provided by the program and the scope of its purpose. The ILPES guide on writing the Logical Framework places a great deal of importance on beneficiaries at the design (or redesign) stage of the programs.
The Various Types of Evaluation Contained in the Guidelines

The Annual Evaluation Program (PAE) established by CONEVAL for the 2007 fiscal year for APF federal programs is designed to determine the types of evaluation that will be applied to federal programs, through an “integral, gradual and useful” program for resource allocation decisions within the framework of the budgetary process for the 2008 fiscal year.

Thus, PAE becomes the instrument that establishes the most suitable combination of types of evaluations for applying to a set of programs and specific policies. As mentioned earlier, the two basic dimensions of evaluation of a program are performance and impact. The former is best evaluated by ECR and the latter by impact evaluation.

Evaluation of Consistency and Results (ECR) includes the analysis of the following dimensions: design, strategic planning, coverage and focusing, operation, perception of target population and results of federal programs. The aim is to achieve an overall evaluation of the program's performance. Its application is carried out through TDR that show external evaluators how to conduct the evaluation study in such a way as to achieve a standardized methodology that will facilitate comparison between the programs. For each dimension, TDR defines its division into subcriteria and the latter into specific questions, making a total of 100 specific questions.

Positive aspects of ECR include:

- The evaluators participating in the external evaluations and the civil servants in federal social programs interviewed regard it as a good instrument of analysis that makes it possible to evaluate the performance level of a social program in an integral fashion.

- The existence of a standard document of terms of reference for facilitating the hiring of consistency and results studies and achieving a level of standardization is regarded as extremely positive. This evaluation not only presents important information on the overall performance of a program but also makes it possible to identify the performance aspects that should be analyzed in more depth. This would also make it possible to determine the importance of applying other types of evaluations, such as evaluations of: design, process and indicators contained in the Guidelines.

- The ECR evaluation also makes it possible to establish a baseline for a program's performance at a specific time. This makes it easier to establish series through the accumulation of the results of studies over time, which in turn makes it possible to determine the evolution of the performance of a program, both at the global level and as regards the performance of specific aspects and also for establishing performance
comparisons between programs. It is therefore necessary to establish performance indices of each dimension and overall indices\textsuperscript{29}.

It is also possible to identify opportunities for improvement related to ECR:

- It is important to clarify the strategic use of the tool and its results, since this will determine its adjustments and the dynamics of its application. The C&R evaluation results received by CONEVAL are distributed to Congress, social program operators and SCHP and SFP and are published on the Secretariats’ on-line portals. They are rarely used and there are no significant consequences as regards accountability, transparency, adjustment of federal social programs or adjustments in budgetary assignment\textsuperscript{30}.

- As for frequency, given the current demands and the time required to carry them out, C&R evaluation does not permit an evaluation of federal social programs nor do they constitute a means for regularly informing budgetary offices on the progress of the program’s performance. It permits an annual evaluation. However, if the goal were to support the monitoring of the programs, it would have to be more frequent. This would imply an adjustment to have a more agile instrument and would require making the process of hiring external evaluators more flexible too\textsuperscript{31}.

- It is easy to adjust the tool to prevent duplication in certain questions, and define criteria and concepts that will permit a more agile instrument\textsuperscript{32}.

- It is possible to develop methods to weight the criteria, sub-criteria and questions and design a system of evaluation and consistency for the answers that will permit the validation of certain issues and the instrument as a whole\textsuperscript{33}.

- It is necessary to establish direct connections from the results of the studies to the budgetary process.

- Having external evaluators undertake annual C&R evaluations in the current format is extremely costly.

\textsuperscript{29} The usefulness of this measurement dynamic in the evaluation of a program’s performance through the ECR instrument can be seen in Hernandez Licona, G., (2007). Monitoring and Evaluation System in Mexico. CONEVAL, Mexico City. CONEVAL is currently drawing up a performance index.

\textsuperscript{30} On the basis of meetings and interviews with office employees and program operators at SEDESOL, Public Education Secretariat, SEMARNAT, among others, as well as employees of the Civil Service Secretariat, and the Secretariat of Finance and Public Credit. Also on the basis of the interview with Dr. Myriam Cardozo Brum, a researcher at the Universidad Autónoma Metropolitana UAM-X an expert on evaluating social programs and an advisor to Congress on this issue.

\textsuperscript{31} If the objective were to support the redesign of programs on the basis of rigorous research, it could be applied biannually and perhaps not to all programs, but selectively.

\textsuperscript{32} For details on the specific recommendations for C&R TDR, see another World Bank document on the Presentation of Comments on Consistency and Results Terms of Reference.

\textsuperscript{33} Another World Bank document included in this technical assistance provides various alternatives for weighting the C&R evaluation structure.
Design evaluation tools include the logical framework approach and the verification of coherence with a sectoral and national objective, through the analysis of a set of questions designed to evaluate this coherence. It is used in two situations, when there are new federal programs, which, since they have no operating or results evidence as yet, need to verify the consistency of their design. The second situation suggested in this report is when the ECR reveals that this dimension has a low result.

In general, the existence of a standard terms of reference document is very highly rated while this type of evaluation is regarded as important. However, there is some difficulty in identifying indicators at the purpose and goal level for reasons explained earlier and the connection to sectoral policy levels. The design evaluations analyzed are usually well written and comply with the aspects requested in the TDR. However, the recommendations are extremely general.

According to the Guidelines, process evaluation is designed to use fieldwork to analyze whether the program carries out its operating processes in an effective, efficient manner and whether it contributes to improving management. This evaluation identifies the “bottlenecks” in the way programs operate. Fieldwork is carried out for their analysis and specific improvements are proposed to solve them.

This evaluation can have very positive effects on the redefinition and adjustment of federal programs' operating rules, since the redefinition and adjustments of the processes to enhance their efficiency and contribution to impact should be reflected in the ROPF.

Based on the Guidelines, the indicators evaluation is designed to use fieldwork to analyze the relevance and scope of a federal program's indicators for achieving results.

The design of the impact evaluation (experimental or quasi-experimental) of key programs is carried out on the basis of the quality and availability of the information in the program and the budget. They are designed to determine the scope of the change in the conditions of well-being of the population benefitting from the goods and/or services of the programs evaluated.

Impact evaluations generally have broad credibility given the methodological rigor with which they are applied. However, some evaluations have caused controversy in certain aspects such as the size of the samples for the surveys applied, which casts doubt on the inferences that could be drawn from them.

The National Evaluation Programs drawn up by CONEVAL annually establish which evaluation impacts are pertinent and timely. However, given the multiannual nature of the impact evaluations (whose results occasionally involve the establishment of a base line, meaning that subsequent evaluations will serve as a line of comparison), it is essential to

---

34 Another World Bank document included in this technical assistance includes of TDR proposal for processes.
35 The TDR proposal for indicators also forms part of the World Bank’s technical assistance.
36 This is the case, for example, of the impact evaluation of DICONSA’s Food Support Program reported on by external evaluators.
have an evaluation strategy that will include the long and medium term, such as a National Evaluation Plan from which annual operating plans are derived.

Specific evaluations are also carried out that are not contained in the Guidelines but determined by the Secretariat of Finance and Public Credit, the Civil Service Secretariat and the National Council of Evaluation of Social Development Policy as needed.

The strategic evaluations of a program or set of programs are undertaken on the basis of strategies, public policies and institutions. Although CONEVAL’s mandate focuses on federal programs, strategic evaluations are oriented towards establishing the contribution of one or several programs to national policies and strategies. It is essential for the National Evaluation Plan to determine which Strategies and Policies are to be evaluated. Given the experience and technical specialization of CONEVAL, its mandate could be expanded to deal with the evaluation of general public policies, and not just federal programs for social policies.

Operating rules

Although federal programs' operating rules do not form part of the Guidelines, they should be considered since they are closely linked to the drafting of their indicator matrices.

The ROPF are a set of provisions for operating programs whose structure includes: program objectives; general guidelines; coverage; target population: specific operating guidelines: institutional coordination; programmatic-budgetary reports; follow-up and monitoring and complaints and denunciations. They are designed to determine the modus operandi of a program on the basis of the principles of efficiency, equity and transparency.

The operation of federal programs is subject to ROPF. The Guidelines' application leads to the drafting of MI matrices. These matrices sometimes establish modifications and adjustments to indicators and also suggest adjustments to the way the program operates. Once agreement has been reached on the adjustments to the matrices, it is necessary to undertake the procedure for adjusting the corresponding ROPF, given their normative importance.

External contracting

The general provisions of the Guidelines state that “their purpose is to regulate the evaluation of federal programs.” 37 The background to these provisions can be found in the “Minimum Requirements Agreement” issued in 2002 by SHCP and the Secretariat of Comptrollership and Administrative Development (now SFF) which establish the minimum requirements that APF organizations and offices should consider to select the academic and research institutions or specialized national or international organizations.

38 Diario Oficial de la Federación April 30 2002.
interested in carrying out evaluations of the federal programs subject to operating rules. Although an initial scheme for regulating the evaluation reports was presented, they concentrated more on adhering to the operating rules than on measuring the programs' performance or impact.

The regulation must determine precisely who can hire, who can be providers, what modalities are used to undertake the process and what mechanisms and rules exist to coordinate the supply and demand of evaluation studies.

The Guidelines also establish the offices’ and organizations’ overall obligation to designate units outside the program’s operations to hire evaluators, the requirements for being an evaluator and the terms of reference of the evaluations. However, in the implementation there have been various situations regarding the forms of contracting, contracting units, types of evaluators, supervision of evaluation, etc., that have caused confusion among some of the actors involved, as one can see from the interviews with federal program evaluators and operators.

In general terms, for some years now, there has been a national market of evaluation studies with certain weaknesses, such as: an insufficient number of providers with sufficient technical capacity, the units that hire evaluators should consider the participation of public agencies, which establishes variations in the hiring mechanisms with private agencies and adhering to the Law of Public Sector Acquisition, Leases and Services. Sometimes conflicts of interest arise when the operating unit of the program to be evaluated hires the evaluation, or there is a lack of or ambiguity in the mechanisms for resolving controversies between the hirer and the evaluator, the difficulty of establishing the right price, an insufficient supply of academic training in the field of monitoring and evaluation, a demand for studies that is either general or ambiguous among other weakness which mean that the market does not operate entirely efficiently.

39 Contracting a national research center is carried out by an Agreement, which is much quicker than using administrative contracting to hire a private center, as a result of which many persons responsible for operations opt directly for agreements.

40 This situation was explained during the meeting with program operators at the Education Secretariat, who explained that although the organizations and units to which the programs are attached sign contracts with external evaluators, it is the operators who express their non-objection to the reports and interact with the consultants, meaning that the contracting unit simply pays. This issue also arose with external evaluators, who in some cases, stressed the complexity of dealing directly with the person being evaluated.

41 During the meeting with a group of external evaluators held at the World Bank Office during the mission, they indicated the lack of clarity in the mechanisms and procedures for solving controversies with the contracting entity.

42 There is still not an academic supply of broad training in the field of monitoring and evaluation of social policies and programs.

43 On the basis of interviews with the employees of APF offices such as SEDESOL, SEMARNAT and the Education Secretariat who indicated that there is a wide variation in the quality of the results of studies by external evaluators and that there are not many organizations (whether public or private) with experience in evaluating social programs. This means that the same evaluator is sometimes hired for several years and that sometimes, although not often, some biddings for public contracts are declared void.
In specific terms, a major weakness is the lack of clarity regarding those that contract the evaluations. The guidelines state that, “Without modifying their structure, the offices and organizations should designate an administrative unit outside the operation of federal programs to coordinate the hiring, operation and supervision of the evaluators.” In fact, teams are formed that tend to include those responsible for running the programs (since they are the ones most familiar with the operation who can guide and supervise the hiring process). Sometimes, one organization pays but another one follows up the contract.44

At the same time, the Guidelines establish the participation of public institutions (research centers) which establish a different contracting mechanism, since two public organizations are involved, and the establishment of agreements, which eliminates stages and times and is preferred to private suppliers. According to the research project currently being undertaken at the Universidad Autónoma Metropolitana, by Myriam Cardoza who was interviewed during the November 2007 mission, in 2006 and 2007, more direct assignments (agreements) were signed than in previous years.

Another problem is the lack of synchrony between contracting times (which includes the information, reception of proposals, analysis, selection and allocation, and drawing up the contract), the time for carrying out the evaluations and delivering results on the basis of the budgetary cycle times. The time for undertaking the adjudication of a consultancy project may vary according to its modality and the efficiency of its management. If it is long or inefficient, this takes time away from the evaluation itself and one may lose the opportunity to accompany the budget cycle. The evaluators sometimes have a very short time for performing the evaluation due to the lengthy process of adjudicating contracts.

As with other aspects of the Guidelines, other positive aspects have been found in relation to external contracting, such as:

- Federal program operators rate the standard documents (TDR1 and TDR2) very highly, which enables them to contract consultancy services for consistency, results and design studies. Drawing up terms of standard reference or models for the evaluations of indicators and processes will also make it possible to carry out these evaluations.

- The TDR of the Consistency and Results Evaluation are rated by external evaluators and federal program operators as an extremely good instrument for evaluating the performance of federal programs.

- In some cases, the evaluation has been participatory and involved those responsible for operating the programs, which has enabled them to learn about monitoring and evaluation.

44 At the meeting with the Education Secretariat, the Head Office of Planning and Evaluation suggested that the program operators contract the evaluators, since they have extensive knowledge of the program, but this is impossible, since it would involve a conflict of interest.
Organizational Requirement

The effective application of General Evaluation Guidelines for Federal Programs requires increasing the management capacity of the APF offices and organizations that coordinate the activities included in the Guidelines. The Head Offices of Planning and Evaluation could be destined to play that role. Their main functions, several of which are already carried out, would be:

- Contribute to the drafting of CONEVAL's annual evaluation program by suggesting priority programs for the organization;
- Supervise and validate the process of drawing up the MI of the programs that include the institution and office;
- Participate in the validation of the identification of indicators for budgetary programs;
- Supervise the contracting, development and results of the external evaluations to be contracted;
- Publish the results of the evaluations according to the Guidelines;
- Promote monitoring mechanisms within the operation of each program and office;
- Develop more frequent, timely performance evaluation tools than ECR (or its substitute).

During the 2007, sharp disparities were observed between the Head Offices of Planning and Evaluation of the various offices. Some had highly-qualified staff, their own staff and a modern system of monitoring and evaluation that was already being implemented. Others had only two people to provide the technical accompaniment for all the programs while still other organizations had yet to define a unit to undertake these functions. There is still no clear definition of the minimal functions and organizational arrangements these units should have.

It would be useful to determine the monitoring and evaluation functions of the areas chosen to be responsible for applying the general evaluation guidelines in the APF organizations and offices, which should also help CONEVAL implement the Guidelines. Pilot experiences could be carried out with certain Secretariats to design and test specific managerial tools, such as management commitments.

It is also necessary to coordinate the initiatives established by PMG for improving institutional management.

RECOMMENDATIONS FOR ADJUSTING GUIDELINES

The following section contains recommendations for adjusting the General Guidelines for the Evaluation of the Federal Programs of Federal Public Administration.
Strategic Orientation of Guidelines

The strategic orientation of the Guidelines regarding the Budget by Results should be clarified. Well-defined objectives are available as regards:

- Regulating the evaluation of federal programs
- Drawing up an indicator matrix and monitoring systems;
- Drawing up the strategic objectives of the offices and organizations of Federal Public Administration.

The use of evaluation studies, however, has not been clarified. They may focus on providing information on the budget, transparency or accountability, supporting the redesign of programs or a combination of all of these. The strategic orientation defined has consequences regarding the use of tools and their dynamics of application.
Main Audience for Guidelines and Implementation of Tools

During the process of drafting the new Guidelines\textsuperscript{45}, one should bear in mind that their main audience are the actors directly involved in their implementation, particularly the program operators and external operators. It is therefore necessary to clarify the concepts, methodologies, procedures and schedules for achieving each of the aspects contained in their objectives: conducting evaluation studies, drawing up indicator matrices, etc.

In this respect, it should be very clear what products the program must deliver, to whom, within what time limit, in what sequence and with what characteristics and who will intervene in their validation, as well as the roles played by the main actors involved. For example:

Regarding MI, the following should be clarified:

- Minimum characteristics MI should comply with;
- Methodology for drawing up MI;
- Process and actors that intervene in its validation;
- Synthesis of Methodology of Identification of Budgetary Program Indicators;
- Process and actors that intervene in its validation;
- Delivery times;
- Usefulness to the program;
- Reports;
- Publication and dissemination of information (on Websites for example).

Regarding the types of evaluations, the following should be clarified:

- Final objective and usefulness;
- What they measure;
- Usefulness for the program;

\textsuperscript{45} It is suggested that the following be eliminated: In Heading V, from the integral system of monitoring and evaluation based on results of federal programs; single chapter, articles 34 to 6. These articles refer to the characteristics of this system. However, since neither the organizational nor the institutional aspects nor the informatic tools of this system have been designed, it should either be eliminated or else a brief description should be provided in the Guidelines objectives. If its remains in its current state, it could give the impression of having to deliver more information to other organizations. Regarding the types of evaluations, it could simply say that CONEVAL will determine the evaluations to be undertaken in keeping with the National Evaluation Program.
- What programs are implemented and with what frequency;
- How they are coordinated;
- Who the external evaluators could be;
- How the external evaluators are selected;
- Who contracts, who supervises and who pays the external evaluators;
- Processes of certifying evaluators.

Although CONEVAL's Website has the guidelines and the main methodologies and forms, it would be useful to include or index other necessary ones from SHCP and SFP. It might be possible to draw up a SED portal in which all the documentation on concepts and methodologies, dates and forms, etc., were in a single site.

**External Evaluations**

In the medium term, performance evaluations could be carried out through a rapid evaluation carried out by APF with the technical support of CONEVAL. These findings could provide the basis for undertaking external evaluations in those programs that require investigation.

In other words, the current external evaluation of programs would be replaced by controls of a different nature in which the program would have a certain level of internal evaluation that would prove data, corroborate information and concentrate on impact evaluations and evaluations of specific dimensions of performance and institutional reinforcement.

The following specific actions are suggested for the long term:

- The strategic orientation of the Guidelines within the framework of the Budget by Results should be clarified.
- Consider the advances at the performance evaluation level at the institutional and sectoral level to achieve an integrated approach.
- Pilot test the Terms of Reference for the Evaluations of Indicators and Processes that will enable these studies to be included in the National Evaluation Program and permit their detailed mention in the Guidelines.
- Define specific criteria of use for the various types of evaluation, their sequence and coordination: Clarify their final objective, use for the program and determine how and

---

46 CONEVAL has begun to compile a register of providers of evaluation services for social programs. It would be useful to generate a process of certification of the evaluators through an analysis of work already carried out, evaluations, etc.
when they are to be applied, etc. The Guidelines could make a brief mention, while the in-depth development of this point could be included in CONEVAL’s Annual Evaluation Program.

- Clearly define the process of validating the MI; harmonize times with SFP and SHCP:
- Update CONEVAL’s Website with the archives required for implementing the Guidelines and other sites of interest.
- Clearly define the monitoring and evaluation functions that should be performed by the areas created for this purpose in the APF units and offices.
- Clarify as much as possible the process of organizing the administrative contracting of external evaluators.

The following specific actions are recommended for the medium and long term:

- Exploring the possibility of quick performance evaluations based on the self-evaluation of their programs by units and organizations and only using external evaluation for validation and verification.
- Expanding CONEVAL’s regulatory function to other sectors and the program level.
- Having a pool of certified external evaluators.

Appendix 2 also contains the Guidelines’ recommendations regarding suggested practices for supporting the role of CONEVAL in managing SED.
APPENDIX 1: Logic of Intervention of a Federal Program

Once the program has been designed and its intervention and results obtained, it can be analyzed through the chain of inputs, activities, products, effects and impact, which are interlinked.

The design includes the identification of the problem, the analysis of the contribution of the program objectives to those of the institution and the verification of the logical framework which culminate in detailed programming.

The design will be used to establish the input requirements, which are the financial, material and human resources required by the program. They support the undertaking of the activities contemplated in the design of the program that permit the production and delivery of previously defined products. A program's products are the goods and services it provides according to the components defined and its target population.

The products delivered are expected to produce effects on the beneficiaries' behavior and conditions according to the nature of the components established and lastly, the impact is the measurement of the changes observed in the beneficiaries' conditions of well-being attributable to the implementation of the program.

The following graph shows the scope of the various types of evaluation established in the general Evaluation Guidelines.

Graph A: Logic of Combining Tools

Source: Drawn up by The World Bank staff, 2008.
APPENDIX 2: Suggested Practices for Supporting CONEVAL's role in SED Management

Draw up a proposal for a global index and performance dimension index for consistency and results

CONEVAL has suggested establishing a global performance index of federal social programs on the database derived from the application of Consistency and Results Terms of Reference that will show the evolution of a program's performance over time (and in its various dimensions) and permit a comparison of performance between programs.

The Consistency and Results Terms of Reference constitute a useful managerial approach for evaluating the performance of federal social programs. Its structure, based on dimensions, sub-criteria and questions, enable the programs to be analyzed through their management cycle, covering every stage from design to the achievement of results.

The TDR1 identify the overall performance level of a program at a set time and make it possible to determine the behavior of certain specific performance dimensions such as design, results, etc.

CONEVAL has made the first efforts to draw up and apply a performance index of federal social programs on the basis of the information obtained from the TDR1 on a regular basis.

Another World Bank document that includes this technical assistance makes comments on the index and presents alternatives for weighting its structure of dimensions, sub-criteria and questions.

Standardize Training regarding the Logical Framework Approach

The interviews held during the mission with the program operators and external evaluators, together with the findings on the quality of the indicator matrices reviewed revealed the need to standardize materials and the approaches to training regarding the production of indicator matrices. The aspects that revealed the greatest problems in the diagnosis should also be emphasized.

Reinforcing Monitoring at the Level of APF Programs, Offices and Organizations

In view of the scant use given to evaluations, it is necessary to consolidate a mechanism which, through the program, will oblige organizations to incorporate improvements and implement relevant recommendations, in a sort of “management commitment.”

It is necessary to establish the actors responsible and the right incentives to carry out and incorporate them into the Annual Operating Programs. It is also important to consider the type of management commitments that will be carried out with PMG in order not to duplicate efforts or create contradictions. The MIDO system and the new PMG tools oriented towards achieving management agreements should be regarded as a framework.
for their establishment. There are other actions designed to reinforce monitoring and evaluation, which should be carried out in the short term, which include:

- Establishing the functions and role of the Head Offices of: planning, monitoring and evaluation, programming and budget or similar areas regarding MI validation; contracting; application and use of evaluation studies and development of an institutional monitoring and evaluation system;

- Establishing alternatives and showing better practices for organizing SM&E at the institutional level.

- Establishing liaison and coordination mechanisms between programs, offices and organizations and those implemented by SED, CONEVAL, SFP and SHC (such as a commission in which various head offices participate);

- Designing management commitments that will promote the implementation of the studies’ recommendations at the program level:

- Developing more frequent, timely performance evaluation tools than ECR.

In the long term, the following actions are suggested:

- Support the design and/or reinforcement of existing SM&E;

- Implement program management commitments;

- Design tools for diagnosing the quality of data production on the basis of the organization's basic statistics.

**Effective Regulation of the Evaluation Studies Market**

The following actions are recommended for strengthening the evaluation studies market and achieving an appropriate, cost-effective ratio for the evaluation exercise:

**Short-term actions:**

- Finish drafting the TDR of types of evaluations of indicators and processes;

- Improve the writing of the requirements, contracts and follow-up mechanisms of the studies to be hired;

- Advance in the registration of evaluators and the design of a scheme for their certification;

- More precision regarding the definition of who can contract, who can be providers; forms of contracting and the resolution of controversies;
- Synchronize contracting times with the times of the evaluation study and budgetary times;

- Administrative contracting training program for offices and programs (by training trainers);

- Forge alliances with universities to promote actions to provide training in program monitoring and evaluation.

- Establish a knowledge transfer program coordinated by CONEVAL and supported by organizations such as CONACYT to build the technical capacities of evaluation teams.

Take the following medium and long-term actions:

- Contract on the basis of the certification of external evaluators;

- Train organizations and programs in administrative contracting;

- Design the scheme for certifying evaluators

Quality of Information

It would be useful to have a diagnosis of the quality of data production through basis statistical sources: Surveys, censuses and administrative registers in Federal Public Administration offices and organizations, for which an abbreviated version of the data quality analysis focus (DQAF) drawn up by the International Monetary Fund could be used. The analysis consists of five dimensions: prerequisites of data quality; data integrity; methodological validity; precision and reliability; usefulness and accessibility. Its aim is to diagnose the quality of the organizations' statistical data production which largely feed the programs' indicator matrix.

This approach could be supported by the APF offices and organizations through the committees for statistical support established by the National Institute of Statistics, Geography and Informatics at various levels of public administration.

In the short and medium term, the following actions should be considered:

- Validation of information diagnosis;

- Drafting of methodological proposal;

- Application of pilot test.
REPORT 4: ANALYSIS OF BASIC INFORMATION IN THE CONTEXT OF
THE GENERAL GUIDELINES FOR THE EVALUATION OF THE FEDERAL
PROGRAMS OF FEDERAL PUBLIC ADMINISTRATION

Main author: Gilberto Moncada Vigo

INTRODUCTION

This report provides a brief, general evaluation of the state of the basic information required for implementing the results-based monitoring and evaluation process of Mexico’s social programs, in keeping with the provisions of the General Guidelines for the Evaluation of the Federal Programs of Public Administration.

The findings are based on interviews with those responsible for certain social programs, technicians in the Social Development Secretariat (SEDESOL), the Civil Service Secretariat (SFP), the Finance and Public Credit Secretariat (SHCP) and a group of external evaluators of social programs. They are also based on talks by the National Council of Evaluation of Social Development Policy (CONEVAL), and the review of external evaluations and complementary information related to the monitoring and evaluation initiatives undertaken by the Mexican government.

The availability of quality, timely and relevant statistical information is one of the pillars of the M&E system of federal programs at the Mexican public administration level. This is one of the main challenges facing the M&E structure of social programs and the Performance Evaluation System.

GENERAL GUIDELINES FOR THE EVALUATION OF FEDERAL PROGRAMS
AND BASIC INFORMATION

The provisions of the General Guidelines for the Evaluation of Federal Programs in Public Administration, issued by SHCP, SFP and CONEVAL include regulating the evaluation of federal programs, drawing up an indicator matrix (MI) and monitoring systems. The availability of basic timely, good quality, statistical information is crucial for supporting this process.

The evaluation activity implies undertaking a systematic, objective analysis of federal programs’ performance to determine the pertinence and achievement of their objectives and goals as well as their efficiency, effectiveness, quality, results, impact and sustainability. In order to support the various types of program evaluation, whether they involve consistency and results, indicators, processes, impact or specific aspects considered in the guidelines, information is a crucial input.

At the same time, monitoring implies periodically gathering information on the various activities carried out by federal programs, thereby becoming a basic administration tool for those responsible for management. Monitoring must consider data and information

---

47 Basic information is information produced on the basis of administrative records, household surveys, censuses and other sources, which serve as input for estimating performance indicators.
collection on key activities, as well as mechanisms for systematizing, analyzing and using information.

The General Guidelines also determine that the offices and entities should produce the MI for each program, on the basis of Logical Framework Methodology (MML). This implies that the MI should consider a set of indicators that make it possible to measure the achievements and/or changes linked to the actions of federal programs and monitor and evaluate results. It also means that the MI must identify the means of verification that enable it to make the calculations and measure the indicators correctly. The means of verification may include statistics, survey information, administrative records, etc.

As one can see, within the context of evaluating programs in Federal Public Administration, the availability of basic information constitutes a central element in monitoring and evaluating programs, in revealing the results of management and facilitating the formulation and adoption of crucial decisions.

In this respect, in keeping with the General Guidelines, there is the challenge of boosting national capacities to produce and systematize basic information that is relevant and useful for the M&E process of federal programs.

Below is a brief diagnosis of the information status of the programs, together with recommendations to improve their availability and quality.

**CURRENT STATUS OF BASIC INFORMATION FOR M&E**

**Low priority of information issue on current agenda**

Nevertheless, the respective actions to promote their production and access to them in a coordinated, integrated fashion, on line with the needs of social policy, have yet to be established. Social programs' current concern and urgency are focused on defining MI in keeping with the guidelines established by SHCP, SFP and CONEVAL: However, the issue of information will acquire preponderance when the stage of formulating MI and the problems associated with it have been overcome and one proceeds to the rigorous calculation of indicators, particularly of results and impact.

Within this framework, it is worth noting the effort made by SEDESOL to build its institutional capacity to improve the basic information that will permit a systematic estimate of its indicators. As part of this initiative, SEDESOL, with the support of World Bank, has implemented a pilot plan for a System of Monitoring and Evaluation Based on Results (SEMBR), designed, among other things, to facilitate the evaluation of social programs to improve their efficacy, efficiency and impact on beneficiary populations.

**Limited Availability and Poor Quality of Information**

The availability of basic information is limited, unreliable, highly aggregated or incomplete for calculating the purpose and impact indicators defined in the indicator matrices of social programs. This is true of the HABITAT Program, which lacks a system of statistical information that would make it easier to find out about the complexity of the
components of the urban environment and estimate and indicator describing the urban improvement of a selected territory.

The information produced by the programs themselves is insufficient while the main sources of basic data of the National Systems of Statistics and Geographical Information (SNEIG) which includes the National Institute of Statistics, Geography and Informatics (INEGI) and the other public institutions often lacks the geographical and/or thematic coverage of specific social programs or is not available. SNEIG's regular statistical operations are often not related to the frequency with which indicators are required to be presented. For example, the SEDESOL State Incentives Program lacks up to date information from the National Population Council (CONAPO) and INEGI which would make it possible to define the target population.

Likewise, the implementation of ad hoc statistical operations is restricted by the lack of budgetary resources for annual evaluations. This is the situation in the National Scholarship and Financing Program of the Public Education Secretariat, which is unable to construct reliable information systems or carry out a proper follow-up of the program.

Data capture is contemplated more in operating terms than from a monitoring and evaluation perspective. Generally speaking, there is very little information to estimate purpose and impact indicators while INEGI constitutes a compulsory source of consultation although in some cases, the information it possesses is not entirely compatible with the program.

There is a need to establish common criteria and quality standards for basic information

Basic information must be subject to certain common quality standards to guarantee its use in the M&E process. This is a crucial point, particularly to achieve the credibility of the process results. To date, it has been observed that social programs are independently responsible for producing and organizing available information and establishing their own standards in a particular fashion. It is important to identify an organization that will supervise and regulate the standards for the quality of information in the monitoring process in a systematic fashion.

For example, the Rural Supplies Program run by Diconsa requires the standardization of a methodology for determining the target population, in other words, one that will determine when a population lacks a consistent supply system. The Potable Water, Sewage and Drainage Program in Urban Zones of the Secretariat of the Environment and Natural Resources needs to have a format for measuring the number of beneficiaries in order to unify counting criteria. The 2006 External evaluation for this program found that beneficiaries are miscounted, since they found a greater number of beneficiaries than that of the local population.

Weakness of Databases and Lists of Beneficiaries

Information on social programs has severe limitations, due, among other things, to the lack of homologation of the variables in the data bases (such as different definitions for
the same variable) or incomplete characteristics of variables, which prevents their use. The data bases are not interconnected, and in some cases, access to these bases is not easy. Likewise there are limitations on the technical capacity of human resources for handling information.

For example, SEDESOL’s 3 x 1 Migrants’ Program needs to improve its information system to have timely databases that are as complete as possible. Likewise, the Farm Day Workers’ Strategy Program of the Employment Support Program of the Secretariat of Labor and Social Welfare requires a detailed review of the data base to improve the follow-up of the actions. The Ecotourism-Alternative Tourism Program in Indigenous Zones run by the National Commission for the Development of Indigenous Peoples (CDI) requires an electronic data base to homogenize and systematize information on the projects.

The registers of many social programs either lack proper information or are incomplete or restricted to intermediaries rather than the end beneficiaries, which limits their capacity as a reference framework for the selection of samples in the evaluations. Some programs do not produce their own information while others fail to keep historical records to make comparisons of the case while still others fail to reveal the quality of the data in terms of scope and accuracy.

This is the case, for example, of the Food Support Program run by Diconsa, which needs mechanisms to facilitate the updating of the list of beneficiaries through information on local actors. The Social Joint Investment Program has problems focusing on beneficiaries due to the lack of an adequate identification of the latter, meaning that a population group that is not living in a situation of poverty is benefitting.
TABLE 1: Principal Problems of Basic Information in Certain Programs

<table>
<thead>
<tr>
<th>Secretariat/Program</th>
<th>Weaknesses in data bases</th>
<th>Weakness in information on list of beneficiaries In target population</th>
<th>Weaknesses Record system/Instruments Recording/system/information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SEDESOL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local Development Program</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>State Incentive Program</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>National Handicraft Promotion Fund (FONART)</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>HABITAT</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>National Milk Purchasing Program run by Liconsa</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>3 x 1 Program for Migrants</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Social-Joint Investment Program</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Rural Supplies Program run by Diconsa</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Savings, Subsidy and Credit Program for “Tu Casa”</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Progressive Housing Program</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food Support Program run by Diconsa</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Farm Day Worker Program</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Senior Citizen Care Program</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Rural Housing Program</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Youth for Mexico Program</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><strong>Public Education Secretariat</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Scholarships and Financing Program</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Rural Education Program</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scholarship Program to Support Intensive Practice and</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Social Service for Students of Seventh and Eighth Years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Terms of Public Teacher Training Colleges</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Health Secretariat</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevention and Care of Disabled Persons</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Health Communities Program</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programs for Care of Vulnerable Families and Population</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Child and Teenage Care Program</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><strong>Secretariat of Labor and Social Welfare</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farm Day Worker Strategy of Employment Support Program</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Get a Grant Strategy of Employment Support Program</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><strong>Secretariat of the Environment and Natural Resources</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clean Water Program</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Safe Drinking Water, Sewage and Drainage Program in Urban Zones</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mexican Social Security Institute</strong></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>IMSS-Oportunidades Program</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Secretariat of Finance and Public Credit</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indigenous Regional Funds</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Ecotourism-Alternative Tourism in Indigenous Zones</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Little Use of Activity Carried out by SNEIG

Likewise, little use is made of the work carried out by SNEIG as continuous producers of data for the M&E process, particularly through sectoral administrative records. Social programs' M&E activity will demand large volumes of data, meaning that, in addition to considering the information created by the registers of the programs themselves, it will have to be complemented by information from other sources such as censuses, surveys and particularly, information from sectoral administrative records.

In general, according to the main results of the external evaluations carried out in 2006, the main weaknesses in the basic information affect the records for the identification of target populations, lists of beneficiaries, information registration systems and available data bases. Table 1 shows the main problems faced by the information on some of the social programs evaluated.

STRATEGIES FOR IMPROVING AND EXPANDING STATISTICAL PRODUCTION FOR M&E

In a context in which the basic information required for facilitating the M&E process is characterized by:

- not being available in sufficient amounts to deal with the M&E requirements;
- being of too low a standard to meet M&E demands;
- not being organized into interconnected or documented databases;
- being affected by limitations in the technical capacity of the human resources responsible for its management and
- being rarely used by the normative, infrastructure and statistical productin coordination resources of SNEIG.

There is an obvious need to take action that will reverse these characteristics to ensue that the demand for information by the M&E process is adequately dealt with and that the lack or poor quality of information does not become a bottleneck that will affect the management by results initiative.

This is in keeping with the mandate of the General Guidelines for the Evaluation of Federal Programs of Federal Public Administration of having information or means of verification to estimate the matriz indicators of the logical framework of the program and permit a correct interpretation of the latter and therefore an objective evaluation of the programs.

The following strategic activities are suggested as a means of improving the availability and quality of basic information.

Strengthen the technical capacity to generate information
In the short term, efforts should be made to reinforce the technical capacity of the professionals in the programs to generate and use basic information. It is essential to improve both the quantity and quality of the information produced by the social programs as well as their use in order to estimate the performance indicators defined in the MI. To this end, it would be extremely useful to develop a plan of action that includes technical reinforcement activities for the units responsible for recording the information.

The activities in the plan of action should include continuous training for the program operators responsible for compiling and organizing the information into conceptual aspects regarding M&E, data collection, processing and validation standards, the design of indicators and the estimation and validation of indicators, among others. This training could be provided through courses, workshops and/or technical assistance from consultants that are experts on the issue.

In this strategic activity, CONEVAL could participate in the definition of common standards and rules for data collection and processing, so that the information produced would be in line with the technical requirements demanded by the M&E of social programs.

**Improve the production of information on the programs' administrative records**

Administrative records are a source of direct information on the programs and are extremely useful for obtaining information that enables one to estimate indicators, mainly at the activities and components level. To this end, CONEVAL could promote specific technical assistance to reinforce this capacity in social programs. Technical assistance should cover the process from primary or basic data capture to the processing and generation of integrated data bases.

In the short term, in order to improve the production of information from the administrative records of federal programs, actions such as the following could be carried out: (i) review and improve the processes for producing information on the basis of administrative records; (ii) define and implement common norms, criteria and methodologies for optimizing the use of records; (iii) review and improve the basic forms of recording information; (iv) establish specific training activities; (v) incorporate available informatics systems to optimize the production of information from records; (iv) make the necessary institutional arrangements to facilitate the production of information from administrative records at the central and sub-national level.

**Standardize the systematization of data bases for M&E**

Another recommendation for short-term action is the standardization of the organization and systematization of databases available for M&E. One of the aspects most heavily criticized by social program evaluators is the heterogeneity between the criteria used and the types of organization of the databases available for evaluating programs. All this makes it impossible to undertake a better, more effective evaluation of the programs.

Preventing this requires defining common technical standards to organize the data bases, promote the homologation of variables and definitions and determine the corresponding
updating, access and use protocols. It is also necessary to increase the interconnection of
data bases between all those that can be linked, once a technical analysis has been carried
out of their condition and the available technological and institutional facilities.

Criteria and standards should be determined to document the databases in an appropriate
manner. In this respect, one alternative would be to use the IT system of the Microdata
Management Tools (HGM) promoted by the International Network of Household
Surveys within the sphere of the World Bank and other international organizations. This
tool makes it possible to document all the statistical operations carried out (censuses,
surveys and administrative records) within common criteria, as well as the databases
created by these primary sources, so that the documentation can be edited and
automatically incorporated into a document in PDF format, or a CD-ROM or the
Internet \(^48\).

**Establish a strategic institutional alliance with the SNEIG**

In the medium term, a strategic institutional alliance should be established with the
SNEIG to take advantage of both their infrastructure and statistical production.

SNEIG is the organization responsible for producing and disseminating national
statistical information on censuses, surveys and sectoral administrative records. SNEIG
comprises INEGI and the main public organizations. According to the Law of Statistical
and Geographical Information, INEGI is authorized to coordinate and promote SNEIG;
participate as the Technical Secretariat of Norms for the development of SNEIG and
produce statistical and geographical information of national interest. To this end, SNEIF
has an infrastructure consisting of information production and coordination units
(technical committees), rules and standards (conceptual frameworks, methodologies, etc.)
statistical infrastructure (statistical frameworks, inventories, etc.) and information
services (databases, storage systems, etc.).

The work of social programs based on the capacity of INEGI and SNEIG together will
facilitate coordination between the sectors and users for identifying needs and the
availability of statistical information for M&E and will promote sectoral statistical
production in keeping with the quality norms, practices and standards required as well as
facilitating the construction of indicators.

**Promote the production of information on the basis of sectoral administrative
records**

In the mid-term, efforts should be made to promote information production on the basis
of the continuous sectoral administrative records. This recommendation is closely linked
to the previous one, because in the context of SNEIG, it is necessary to increase the
information in sectoral administrative records for statistical purposes useful for federal
programs.

\(^{48}\) The characteristics of the Microdata Management Tools (HGM) have been explained to INEGI, which
has expressed interest in them.
The M&E process will continuously consume a large volume of information. Although the information generated by social programs is a major source, this is not sufficient for carrying out the estimates of certain indicators, especially the MI purpose and goal ones. Although household surveys and special studies are also viable sources of information, budgetary restrictions pose an obstacle to their continuous creation. Given this situation, the sectoral administrative records kept by public organizations as an instrument for their own management are one of the largest, most important complementary sources of data for obtaining statistical information.

This is a low-cost source of statistics (since the records have already been established). It offers continuous statistics or statistics for multiple time intervals and above all, facilitates information for total geographical coverage of the universe of study. The thematic coverage of the records is extremely broad. In Mexico, records are produced on life events (birth death), civil events (marriages and divorces), health, education, violence and crime, judicial events, together with work, economic, trade and environmental records.

Sectoral records are therefore a source of information that could be taken advantage of to facilitate the M&E of social programs. To this end, coordination agreements and commitments should be established in conjunction with sectoral offices. A first alternative would be to carry out this activity with the participation of the main sectors committed to the social area, such as education and health, within the framework of the institutional alliance with the SNEIG.

Utilizing administrative records for statistical purposes requires carrying out technical activities using well-defined criteria, in addition to establishing agreements and a solid, inter-institutional coordination. In this respect, political will is crucial.
REFERENCES


