



A Blueprint for a Regional Public Goods Monitoring and Evaluation System

Laura Bocalandro y Rafael Villa

**Inter-American
Development Bank**

Sector of Integration
and Trade

TECHNICAL NOTES

No. IDB-PB-164

September 2011

A Blueprint for a Regional Public Goods Monitoring and Evaluation System

Laura Bocalandro y Rafael Villa



Inter-American Development Bank

2011

© Inter-American Development Bank, 2011
www.iadb.org

The Inter-American Development Bank Technical Notes encompass a wide range of best practices, project evaluations, lessons learned, case studies, methodological notes, and other documents of a technical nature. The information and opinions presented in these publications are entirely those of the author(s), and no endorsement by the Inter-American Development Bank, its Board of Executive Directors, or the countries they represent is expressed or implied.

This paper may be freely reproduced provided credit is given to the Inter-American Development Bank.

A Blueprint for a Regional Public Goods Monitoring and Evaluation System

Laura Bocalandro
Rafael Villa



Inter-American Development Bank
2011

<http://www.iadb.org>

The Inter-American Development Bank Technical Notes encompass a wide range of best practices, project evaluations, lessons learned, case studies, methodological notes, and other documents of a technical nature. The information and opinions presented in these publications are entirely those of the author(s), and no endorsement by the Inter-American Development Bank, its Board of Executive Directors, or the countries they represent is expressed or implied.

This paper may be freely reproduced.

Table of Content

Abstract

Introduction

PART I: Diagnostic of the Monitoring and Evaluation of Regional Public Goods

Chapter 1: Governance, Structure and Performance of the M&E Function for RPGs

- Governance and mandate for an RPG self-evaluation M&E system
- The Bank self-evaluation system and RPGs
- Current performance of the M&E function for RPG projects

PART II: Strategic and Conceptual Considerations for a RPG Monitoring and Evaluation System

Chapter 2: Strategic Considerations for an RPG M&E System

- Monitoring and evaluation: functions and tasks
- Strategic approach to a RPG M&E system

Chapter 3: Conceptual Considerations for an RPG M&E System: the chain of results

- Conceptual basis for a chain of results framework for RPG
- A chain of results for RPG

PART III: Blueprint of a RPG Monitoring and Evaluation System

Chapter 4: A blueprint for the RPG M&E System

- Gauges and indicators
- RPG M&E Instruments

Conclusions

Bibliography

Abstract

This technical note presents a blueprint for a Monitoring and Evaluation System for the Regional Public Goods Program. The blueprint took into consideration the improvements and continued deficiencies of the monitoring and evaluation function currently applied to the RPG Program and projects, and the state of the art in monitoring and evaluation of integration programs, particularly those focused on the software aspects of integration.

The scope of the proposed system includes both monitoring and evaluation of impact and performance, and auditing results and processes (or technologies). Its design follows a principled pragmatic approach, paying particular attention to the cost effectiveness of the M&E function and a balanced governance and collaborative process in each proposed activity. The strategic approach for its implementation is that of risk management orientation, with a focus on sustainability and scalability, promoting feedback and engagement, and taking advantage of self-monitoring and collaborative evaluation.

The blueprint for the RPG M&E system consists of a platform with six instruments and a series of gauges reporting on pre-identified indicators along a results chain path. The six instruments are: (i) a repository for documents and data; (ii) a quality at entry checklist; (iii) a managerial situation dashboard; (iv) the performance accountability system (PMR-like system); (v) the sustainability assessment report (XPMR-like system); and (vi) case-based evaluation of interventions on the results chain. The gauges will be parked at locations along the path of the results chain, and will report on the performance of indicators with respect to inputs, activities, and outputs, including intermediate and final ones related to impact.

Introduction

Monitoring and evaluation are perceived by many as tasks to be performed, along with many other administrative chores that comprise the execution of an operation. This Technical Note recognizes this reality and yet was prepared from a different perspective: one that views the monitoring and evaluation function as a tool to focus work while keeping an eye on the larger picture; one that will allow leveraging the power of the monitoring and evaluation function to show results and enable managing by results; one that will serve as a learning tool for developing real solutions to development problems and for identifying the source of accomplishments.

The need to solve problems or take advantage of current opportunities is as strong as ever. Globalization, decentralization, technological advancement and external forces, such as environmental pressures (e.g. climate change) and cultural and economic pressures (e.g. migration issues and security issues) present new complexities and increased spillovers, moral hazards, and counterincentives that make planning for and measuring of results much more difficult. Scarcity of resources and increasing demands continue to intensify. Events such as the financial crisis and enormous demands for aid, such as the famines in Africa or the outbreaks of contagious diseases around the world, increase the need to prioritize investment in development interventions. Promoting, identifying, isolating and proving positive effects from a given intervention have become key tasks for fulfilling the promise of development.

The Inter-American Development Bank has embraced this challenge and pledged to increase its ability to document, check and communicate its interventions as fast and efficiently as possible to increase the likelihood of positive results. As an IDB instrument to promote development and as a key element of the Strategy for Global Competitiveness of the GCI-9, the Regional Public Goods (RPG) Program is must face this challenge, and it is up to it.

This Technical Note is not an evaluation of the M&E function under the RPG Program nor an application of a M&E methodology. The task addressed with this Technical Note is that of developing a blueprint for an M&E system for the Regional Public Goods Program.

The blueprint is grounded in a strategic approach for an RPG M&E system that consists of doing what works, working at different levels and at different points in the chain of results, with multiple tools, and under a theory of change that follows a logical and cumulative chain of results (). This strategic approach is paired with a set of principles based on the OECD standards for Global and Regional Projects and Programs (GRPP) and the broader standards and conceptual evolution for development evaluation. The expectation is that a good design will invite all RPG stakeholders (countries, executing agencies, project's strategic partners, team leaders, sector and division chiefs, the RPG team and others involved) to tackle this perceived obstacle as a means of identifying what works.

PART I: Monitoring and Evaluation of RPGs - Diagnostic

Chapter 1: Governance, Structure and Performance of the Monitoring and Evaluation Function for RPGs

“The evaluation system has a key role to play in contributing to IDB’s ability to deliver results where they matter—on the ground—and for the organization’s overall goal of effective development.”¹

To date, the M&E function for RPGs has relied solely on the Bank-wide institutional and operational structure. The definition of responsibilities, procedures, as well as systems and incentives, has been the same as those used by the Bank for its operations. Economies of scale were expected from a general design.

In the last few years, the Bank has renewed its centralized effort to enhance monitoring and evaluation, devoting a substantial amount of resources to improving its systems and procedures. This has imposed a set of conditions on separate M&E efforts, such as those of the RPG Program, which range from inheriting the governance structure of the self-evaluation system (i.e. definition of responsibilities, roles, accountability systems, checks and balances, etc.) to the usage of poorly designed and disconnected systems for the particular needs of the RPG Program and projects. The lack of flexibility has negatively affected the performance of the RPG M&E function by contributing to poor compliance with deadlines and commitments, and a lack of conclusiveness of the Program. Improvement of the M&E system has not met expectations yet, and there are no incentives to comply with RPG M&E activities. In addition, reporting overlooks the reality of RPG projects, preventing any meaningful lessons to be retrofitted into the RPG Program.

Governance and mandate for a RPG self-evaluation M&E system

The original document pursuant to which the Regional Public Goods Program (GN-2275-5) was created sets forth the objective and tasks for M&E of the RPG projects and Program. Paragraph 77 provides that “this monitoring will avoid diffuse actions and will allow the use of common guidelines and standards in the selection of proposals and for the supervision, monitoring and evaluation of the operations of the Initiative” and “[it] will permit the opportune identification of situations that merit the cancellation of multiyear operations with a high risk of not achieving their development objectives.”² According to the original mandate, M&E should be measured by efficiency and the likelihood of achieving results and development objectives. Other M&E priorities hinted were reliability and transferability with the use of common guidelines and standards for the whole project cycle.

Pursuant to the original document, the RPG Program coordination team (initially located in the INT/RTC division) was charged with the “responsibility for monitoring the execution of the projects” and “as part of the monitoring and reporting requirement for the Initiative as a whole, each responsible department will provide the required information to the coordinating Bank unit (INT/RTC).” Responsibility at the project level, however, was divided as the document provided that “operations will be implemented

¹ Draft report conclusion by an external review group commissioned by the Bank to assess the quality of evaluations and make recommendations for strengthening the role of evaluation in achieving development effectiveness. Inter-American Development Bank, *IDB Commission on Evaluation Report*, Inter-American Development Bank, 2011).

² Inter-American Development Bank, *Regional Public Goods: proposal for an expanded Bank role. Final approved version* (Document GN-2275-5), Inter-American Development Bank, 2005.

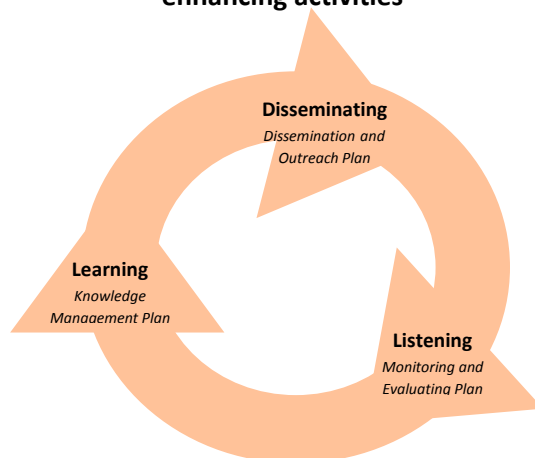
under the supervision of the Bank department responsible for the area of activity supported by the RPG.”³

Further, the original document tied RPG M&E to the governance and procedures of the Project Performance Monitoring Reporting System (PPMR) system (and subsequent incarnations thereof) curtailing further independence in the original mandate. Thus, it provided that “the current policies governing the use of funds, procurement and monitoring of the Bank’s projects will be applied to the operations funded through the Initiative”, and that “The Project Performance Monitoring Reporting System (PPMR) will be used for monitoring.”⁴

As a result, INT/RPG was responsible for monitoring and evaluating the RPG Program overall while each team leader, embedded in her own Sector, was responsible for monitoring and evaluating each project. Any strain arising from this two-headed approach was dealt with through the participation of the RPG team as a member in each project team and by building proactive collaboration to coordinate monitoring activities. Such collaboration was more active in connection to monitoring and less developed in the case of evaluation activities.

The Coordination team at INT has adjusted its strategy to deal with this double governance over M&E tasks and responsibilities. In its Report to the Board of Directors (BOD) on the 2008 call for proposals, INT/RPG presented a framework for enhancing Program activities that articulated the different priorities mandated by the original document. Document GN-2275-14 indicated that “in order to enhance effectiveness at the Program and project levels, and taking advantage of the Bank’s ability to play an active role in the cross-fertilization of knowledge and triangulation in South-South cooperation, Management has designed and is continuously fine-tuning a virtuous circle of listening, learning and disseminating.” The Monitoring and Evaluation Plan (MEP) focuses its activities on (i) **promoting a constant dialogue between the Program and the projects as well as among the projects**; (ii) identifying shared challenges both at the technical and strategic level and potential best practices; (iii) monitoring international conceptual developments and practices; and (iv) developing a solid framework for external and self-evaluation at the project and the Program level. The plan links with the knowledge management plan (KMP), where the Program provides incentives for the production of knowledge and for adapting lessons from the monitoring process.” (Paragraph 2.2, emphasis added). This MEP plan evidences the importance given to proactive collaboration to coordinate monitoring activities with each team project.⁵

Graph 1. RPG virtuous circle for enhancing activities



Despite the advancement of the MEP, the tools available for performing the M&E function were the Bank-wide tools. The same RPG Report to the BOD indicated that achievement of each project’s

³ Idem

⁴ Idem

⁵ Inter-American Development Bank. *Regional Public Goods - 2009 Report and recommendations for financing* (Document GN-2275-14), Inter-American Development Bank, 2009.

objective is measured through Bank tools, including a logical framework, PPMRs with day-to-day scorecard monitoring by the Loan Management System (LMS), the Operations Management System (OPMAS) and other Bank systems, as well as mid-term and final external evaluations, final Management Review Meetings (CRGs) and Project Completion Reports (PCRs) supporting the outcome-based approach. Some of these tools have been in transition or under construction for the last five years.⁶

The Bank-wide M&E framework also requires collaboration; this time between team members at Headquarters and specialists in Country Offices and between them and the executing counterparts. OA-200 establishes that in order to implement appropriate mechanisms for effective monitoring, support, and evaluation of the execution of operations “the Bank seeks to ensure appropriate and on-going participation by its staff, based on effectively coordinated actions and mutual cooperation between Headquarters and the Country Offices, and between them and project executing agencies in borrowing member countries.” It continues by stating that “to ensure that the Country Offices are in a position to fulfill their part of the commitment, they have been delegated the necessary authority to enable them to provide timely responses...”⁷ This delegation has effectively consisted in transferring operational governance of the M&E function to Country Offices.

The Project Monitoring Report (PMR) (previously called PPMR) and the Extended Project Monitoring Report (XPMR) (previously called PCR) are the two main instruments of the Bank’s M&E system. The Operation Manual unequivocally assigns responsibility for these instruments to the Team Leader. CO-303 provides that “The PMR must be updated by Project Teams and sent for validation by Project Team leaders twice a year.”⁸ It further provides that validation of progress achieved and reported through this system is the responsibility of both the Division Chief and the Representative. The operational policy for XPMRs (CO-309) is more comprehensive because of the integration of several internal stakeholders: “Regional Operations Departments are responsible for the timely preparation and processing, as well as the overall quality, of the PCRs of all Bank-financed operations. While specialists in the Country Office or at Headquarters are the principal authors of the PCR, it is expected that other Bank staff, with knowledge of the project context (members of the original Project Team) and/or its implementation, participate in the preparation of the PCR, independent of their current assignment.”⁹ The Project Team includes staff from the Regional Departments, specialists (or Team Leader) and other relevant specialists. In the end, however, the Team Leader is the main author of this report.

In summary, the BOD gave the overall mandate to INT/RPG to coordinate and ensure the monitoring and evaluation of the RPG portfolio while the particular responsibility to monitor and evaluate each operation remained solely with the Team Leader. Under this governance structure, the best approach for a successful M&E system was to work in coordination and collaboration.

⁶ Idem

⁷ Inter-American Development Bank, Policy OA-200, *Monitoring and Evaluation of Project Execution: Chapter Summary*, 1998, <http://manuals/go.cfm?do=Page.View&pid=1059> (Accessed Oct. 5, 2011).

⁸ Inter-American Development Bank, Policy CO-303 *Progress Monitoring Report*, 2010, <http://manuals/go.cfm?do=Page.View&pid=470> (Accessed Oct. 5, 2011).

⁹ Inter-American Development Bank, Policy CO-309 *Project Completion Report – PCR*, 2006, <http://manuals/go.cfm?do=Page.View&pid=484> (Accessed Oct. 5, 2011)

The Bank self-evaluation system and RPGs

Since 1995, when the first Bank Evaluation System (BES) was created, the Bank self-evaluation system has been evolving towards an approach that is more integrated with the Development Effectiveness Framework (DEF). The Commission on Evaluation Report stated that “Conceptually, the self-evaluation models have evolved over the years, from systems focused primarily on measuring inputs toward systems that more adequately measure outputs, outcomes and impacts.”¹⁰ The current framework (the DEF) comprises five main products or instruments, four of them at the operations level and one at the portfolio level. These instruments link to other managerial and accounting systems. The four operations level instruments are the Development Effectiveness Matrix (DEM), the Project Monitoring Report (PMR), the Project Completion Report or Extended Project Monitoring Report (XPMR), and the Loan Results Report (LRR). At the portfolio level the main instrument is the Development Effectiveness Overview (DEO) which reports annually on the Bank’s development effectiveness.

The RPG M&E framework includes three out of the five instruments: Development Effectiveness Matrix (DEM), Project Monitoring Report (PMR), and the Project Completion Report (now called Extended Monitoring Report - XPMR). The DEM was partially integrated and only during 2010 as a pilot activity. The Project Monitoring Report (PMR) and Extended Project Monitoring Report (XPMR) were incorporated from the beginning following Bank policies for technical cooperation. The other two instruments, LLR and DEO, were developed more recently Bank-wide and therefore have not been formally integrated yet into the RPG framework. However, from the beginning, the function performed by both instruments was covered by two other home-grown instruments applied to the RPG Program and projects. The function of the Loan Results Report was performed with a mandatory external mid-term evaluation for each RPG project; it fine-tuned the project design about half way through execution as the LRR does. The RPG Annual Report (series GN-2275) partially fulfilled the function of the Development Effectiveness Overview.

The main Bank-wide instruments used by the RPG self-evaluation M&E framework are the PMR and XPMR. The following describes each of them:

1. *Project Monitoring Report (PMR)*: its objective, according to CO-303, is “to monitor the progress of project implementation.” This report purposely seeks “to enhance management decision-making by serving as an early alert mechanism for project implementation.” To do that “it focuses on the delivery of outputs and the achievement of outcomes using both quantitative and qualitative data.” Additionally, “collecting and storing this data in a single system with the capability of aggregating data in various ways provides management with a powerful tracking and reporting tool.”¹¹ According to the Commission on Evaluation Report, this instrument is at the heart of monitoring under the self-evaluation framework. The same document reports that “these reports are widely perceived as a useful and relevant tool by operational staff.”¹²

The PMR system was put in place in 2009. It replaced the PPMR 3.0, which functioned from 2007 to 2009. The PPMR 3.0 system was only operational for regional technical cooperation operations, including RPGs, during the first semester of 2008. The PPMR 2.0 which came

¹⁰ Inter-American Development Bank, *IDB Commission on Evaluation Report*, 2011.

¹¹ IDB, *Progress Monitoring Report*, 2010.

¹² IDB, *IDB Commission on Evaluation Report*, 2011.

before the PPMR 3.0 system did not include any provision for technical cooperation projects or RPGs in particular. The current PMR system has only been implemented for loan operations, and the PMR portal indicates that there will be a system for Technical Cooperation and Knowledge and Capacity Products after Q2-2011.¹³ To date RPG projects have been unable to use this system and as a consequence it has no PMR reports for RPG projects in it.

2. *Project Completion Report or Extended Project Monitoring Report (XMPR)*. According to CO-309, the objectives of this report are “to assess the results achieved by the project, to foster its sustainability, and to extract lessons learned to improve the design and execution of future operations.” This report purposely goes deeper into the evaluation functions to delve into issues of sustainability and transferability. To that end, it “must be prepared in the context of a joint Bank - Borrower / Executing Agency participatory evaluation process.”¹⁴ The Commission on Evaluation Report was of the opinion that “while a large majority of respondents view these reports as relevant to their work (up to 80 percent in some units), the evidence in support of their usefulness and influence was weaker.”¹⁵

The Extended Project Monitoring Report (XMPR) was unveiled together with the DEF implementation in 2008. This report is slowly replacing the old Project Completion Report (*PCR*). The survey conducted by the Commission on Evaluation “revealed a strong sentiment that the new extended project monitoring report is an improvement.”¹⁶

The whole set of instruments are interconnected. Particularly there is a strong link between the DEM, the PMR and the XPMR. The following paragraph extracted from the Commission’s report describes this link: “A Development Effectiveness Matrix (DEM) is prepared for all operations. The DEM feeds into a Project Monitoring Report that assesses project progress and updates information as needed in the matrix. At the end of the project cycle, the progress report feeds into the Extended Project Monitoring Report, which is intended to provide relevant closing details such as lessons learned and results achieved.”¹⁷ The same degree of interconnection is not available yet for the RPG framework, partially because this entry section provided by the Bank-wide DEM is not fully developed for RPGs.¹⁸ The blueprint introduced in part 3 will include a proposal for an instrument such as the DEM that interconnects design with monitoring and evaluation. The following graph presents a comparison between the Bank-wide self-evaluation main instruments and the RPG framework.

¹³ *Progress Monitoring Report Web Portal*. <http://hqpapmr02.idb.iadb.org/pmrptl/index.do> (Accessed Oct. 5, 2011).

¹⁴ Inter-American Development Bank, Policy CO-309, *Project Completion Report – PCR*, 2006.

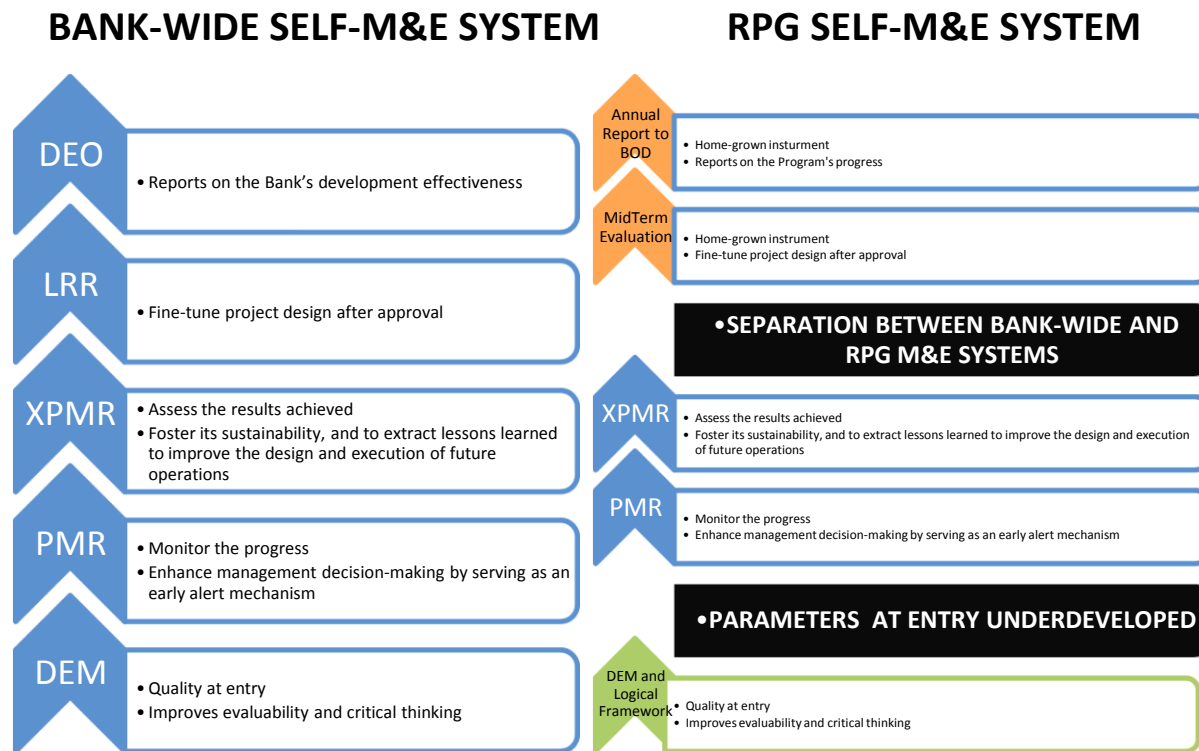
¹⁵ Inter-American Development Bank, *IDB Commission on Evaluation Report*, 2011.

¹⁶ *Ibid.*

¹⁷ *Ibid.*

¹⁸ A first test to comply with DEM matrices for each project was run in 2010, in which formats for KCP projects were included as annexes to all projects approved. However, the format does not provide any valuable information for quality at entry and does not permit to improve evaluability. An adequate requirement should include the specific whole process for RPG monitoring and evaluation.

Graph 2. Bank-wide self-M&E system vs. RPG self-M&E system



In addition to Bank-wide instruments, the RPG M&E framework includes a portfolio monitoring report that acts as an alert mechanism at the portfolio level and complements the PRM function. This instrument articulates disbursement data with some anecdotal milestone accomplishments for each project and administrative information about project execution such as execution period, compliance with contractual clauses and other variables from Bank-wide systems. The PMR monitoring function at the project level completes this framework with administrative tools such as the semiannual report from the executing agency, the Annual Operating Plan (POA), and the financial auditing report. The final external evaluation that assesses results, sustainability and lessons learned complements the XPMR. The RPG alerts system and the final external evaluation, together with the mid-term external evaluation and the RPG annual report to BOD, are home-grown instruments to monitor and evaluate the portfolio as a whole and the projects individually.

At the program level, the annual report summarizes the performance of projects as a whole, including information on trends across the portfolio. Document GN-2275-3 established that “in addition [to the use of PPMR], an analysis will be conducted of the effectiveness and efficiency with which the resources assigned to the Initiative have been used, as well as of the procedures involved. Bank Management will submit to the Board of Directors annually a report describing the results of the evaluations at the same time that information is submitted for new projects selected by means of the annual call for proposals.”¹⁹ This report to the BOD performs a function similar to that of the DEO in the Bank-wide self-evaluation system.

¹⁹ Inter-American Development Bank, *Regional Public Goods: proposal for an expanded Bank role. Final approved version* (Document GN-2275-5), Inter-American Development Bank, 2005.

Current performance of the M&E function for RPG projects

The RPG M&E system has fallen short of accomplishing accountability and learning purposes. This judgment could well refer to the Bank-wide M&E self-evaluation system or the Technical Cooperation M&E subsystem, both of which are overarching systems to the one used by RPG. Regarding the Bank-wide system, the Commission report found that “until the Development Effectiveness Framework (DEF) was put in place in 2008, IDB lacked a comprehensive self-evaluation system, despite a long history of failed attempts to establish one.” Moreover, the same report concluded that “self-evaluation processes at IDB have been predisposed to fail, which means the current system, still less than three years old, should be viewed as a somewhat fragile framework even though it’s on the right track.” Looking forward, the commission was of the opinion that “the new system offers IDB a real opportunity to build a sustainable self-evaluation process for the long term.”²⁰ Both pitfalls and promises of the Bank-wide system could be transferred to the RPG system as the RPG M&E framework inherently lacks comprehensiveness and is still fragile. Taking advantage of both world-wide technological advances and internal opportunities, such as those presented by the DEF and the new Global Competitiveness and Integration Strategy, can make a difference to the RPG M&E function and the Program as a whole.²¹

The Technical Cooperation M&E subsystem, where the RGP M&E system resides, was described by the Office of Evaluation and Oversight (OVE) as a non-operational, acutely non-evaluable program, not built for achieving results. OVE concluded that “the Bank still lacks an adequate mechanism to systematically track TC results”²² and it identified one of the main sources of the problem at the entry level, when they found that “evaluation is compromised early on at the design stage: in 2008 only 59% of TCs had M&E activities planned and only 34% had logical frameworks.” The TC M&E monitoring function at the project level relies solely on the PMR report, the evaluation function relies exclusively on the XPMR (or previously named PCR) report.

According to OVE’s report on Technical Cooperation (RE-364), the rate for PPMR completion for closing projects in 2007 and 2008 was 7%. Later on the same report OVE states that “In practice PCRs are required only for ‘operations for which a ... PPMR is prepared.’ Similar to PPMRs, compliance with the PCR requirement is low, and there is no tracking of their timely submittal.” Moreover OVE concluded that “...less than 3% of the TCs closed in 2007 & 2008 logged specific deadlines for the production of their PCRs in the Bank’s systems, and none reported whether deadlines were met.” The same challenges still remain: sole reliance on the PMR and XPMR, lack of compliance, lack of quality in the reporting activities, and particularly lack of quality at entry.

The RPG performance with PMR compliance is slightly better in almost all of these compliance measurements. Collaborative efforts between the INT/RPG Coordination team and the Team Leaders towards completion have been ongoing. The results of this collaboration can be seen when comparing the 7% rate of completion for comparable TC projects with the 28% PMR completion rate for RPG projects (see graph 3). The vacuum of system and lack of clear rules for PMR reports in the transition

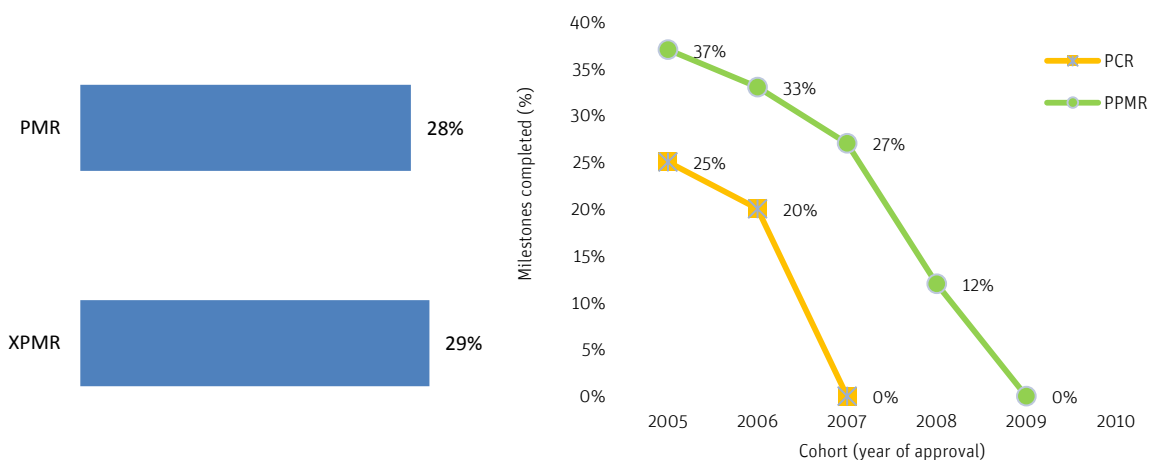
²⁰ Inter-American Development Bank, *IDB Commission on Evaluation Report*, 2011.

²¹ Inter-American Development Bank, *Sector Strategy to Support Competitive Global and Regional Integration* (GN - 2565-4), IDB, 2011.

²² Office of Evaluation and Oversight (OVE), Inter-American Development Bank, *Evaluation of the Bank’s processes for managing technical cooperation* (RE-364), IDB, 2010.

from the PPMR 3.0 (2007) to the new PMR system (2011) is noticeable. Up to the first semester of 2011, 57 PMR reports had been filled out. In order to bypass the lack of a reporting system, several PMRs (completed after 2007) were stored under the old system and others were filed in Word and published through the IDBdocs tool. This environment prevents these reports from reaching the division level management so that they can approve the PMR report.

Graph 3. Total compliance of PMR and XPMR and evolution of compliance by annual cohort



Regarding XPMRs or PCRs, as of mid-March 2011, only 7 projects were closed and declared completed. The Bank acted as executing agency in 2 of the 7 projects. As it is known, and highlighted by OVE in its Report RE-364, the dynamic of monitoring and evaluation for IDB-executed project is in practice very different from those executed externally. Even when Policy CO-309 does not exclude these projects from PCR completion, it only gives specific attention to “the context of a joint Bank -- Borrower – Executing Agency” relationship. Moreover, there is evidence that under these circumstances, at least one (RG-T1153) of the two IDB-executed RPG projects had a final review meeting similar to a final QRR for PCR. Among the five externally executed projects, two have a PCR. As graph 3 shown, this means 40% of all externally-executed projects or 29% of all completed and closed projects by mid-March 2011 had PCRs.²³

Recognizing that the problem of low PCR completion is structural (see recommendations from OVE to Initiative for the Integration of Regional Infrastructure in South America (IIRSA) and Plan Puebla-Panamá (renamed as Proyecto Mesoamerica) Programs),²⁴ the RPG Program introduced other measures aimed at increasing accountability: (i) The RPG program followed the implementation of the DEM for KCPs from the time the DEM was launched; and therefore, all the RPG projects approved during 2010 had a fully completed DEM. (ii) Since inception, the RPG Program implemented the practice of requiring and enforcing external evaluations (Intermediate evaluations in several cases and a Final one in all the cases)

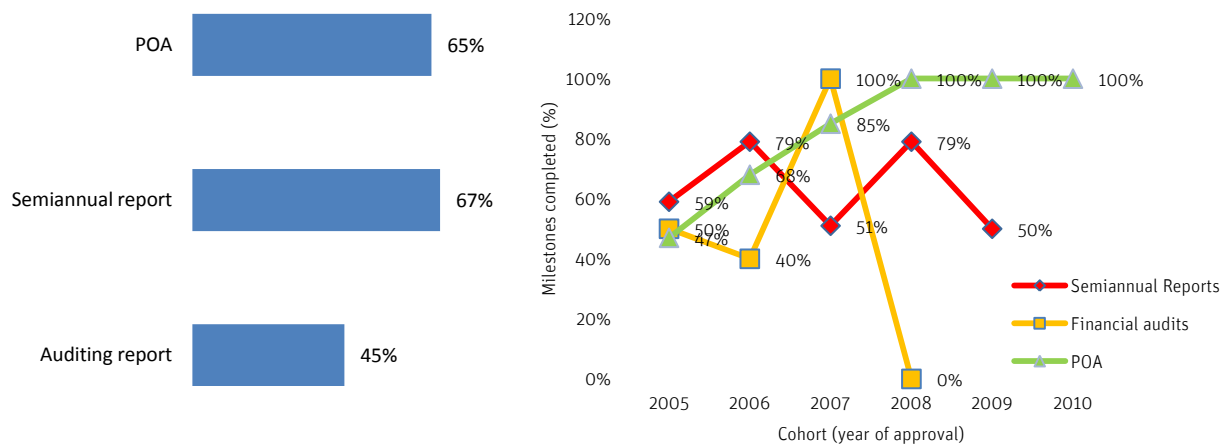
²³ Idem

²⁴ Office of Evaluation and Oversight, OVE, Inter-American Development Bank, *Evaluation of IDB action in the initiative for integration of regional infrastructure in South America (IIRSA)(RE-338)*, IDB, 2009; and Office of Evaluation and Oversight (OVE), Inter-American Development Bank, *Evaluation of IDB’s support to the Plan Puebla Panama Initiative (RE-350)*, IDB, 2009.

in an effort to increase M&E capacities with tools within each project’s control. This is how, as of mid-March 2011, all the RPG projects closed and finalized had an external evaluation. Although the RPG Program instruments do not replace the Bank-wide tools, such as the PCR, they complement and support them while providing a TC-focused approach to Bank-wide M&E tools. Moreover, compliance with external evaluation or any other M&E instrument is a necessary but not sufficient condition for having an effective evaluation.

RPG administrative efforts (POAs, bi-annual reports from the executing agency, etc.) are completed with greater frequency but are still not fully integrated with processes for monitoring the progress of execution. External semiannual reports prepared by the executing agency are carried out most often with a compliance rate of 67%, followed by planning tools, such as POAs, with a rate of 65%. Financial auditing reports are undertaken by 45% of projects. The evolution of POA completion is highly influenced by project cohorts, and is consistent with the timeframe for implementation of the POA instruments. New projects that follow new policies are required to plan ahead with POA instruments. Older projects are required to make a steady transition towards the prescriptions and guidance of the new fiduciary policy. The transition has been slower than expected, especially the incorporation of POAs for older projects.

Graph 4. Total compliance of POA, Semiannual report and Auditing report - Evolution of compliance by annual cohort

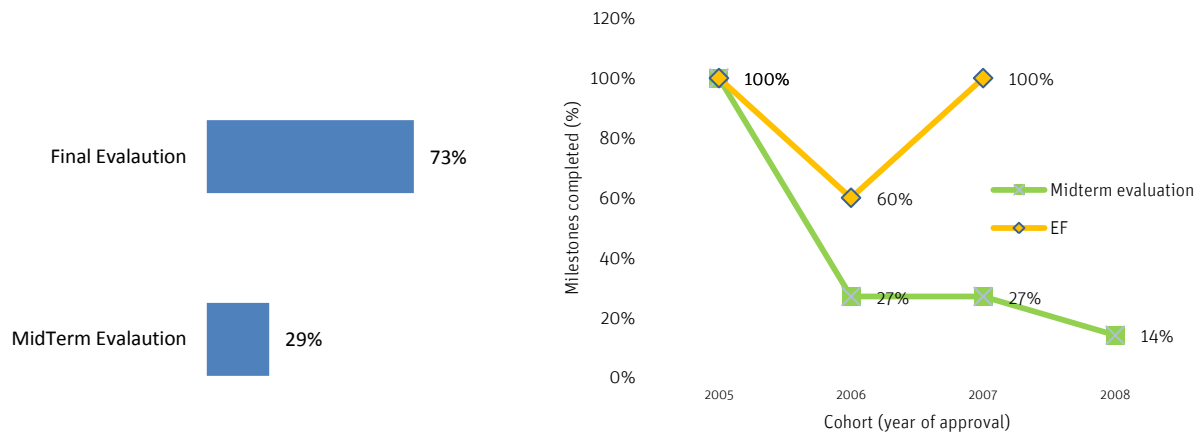


Home-grown M&E tools, such as external evaluations, have a mixed performance with respect to their timely production. Many mid-term evaluations were performed after midpoint of execution. Final evaluations are carried out in a more timely fashion possibly because they are included in the OPMAS system and are a prerequisite for the administrative closing of the project. Compliance with mid-term external evaluations is approximately 29%, with a high variance between cohorts and a significant lag time. To date, all projects approved in 2005 have done a mid-term evaluation while projects approved in 2006 and 2007, almost all of which already have passed the midpoint, have yet to fulfill this requirement. This situation handicaps the usefulness of the mid-term evaluation. Final external evaluations reach approx. 73% compliance with only 2 projects from the 2006 cohort missing one.

It is interesting to note that the perception of the different stakeholders in the Project Team contrasts with actual levels of compliance. According to a general survey conducted by the RPG Coordination

team in 2011, half the project teams believed that they met all the milestones (POA, AUD, Semiannual reports, PMR, XPMR, External Evaluations etc.), while the other half believed they missed some. This perception is shared by the Team Leader, the Project Coordinator (from the Executing Agency) and the Fiduciary Specialists. The Team Leaders' perception about compliance with milestones is slightly more positive.²⁵

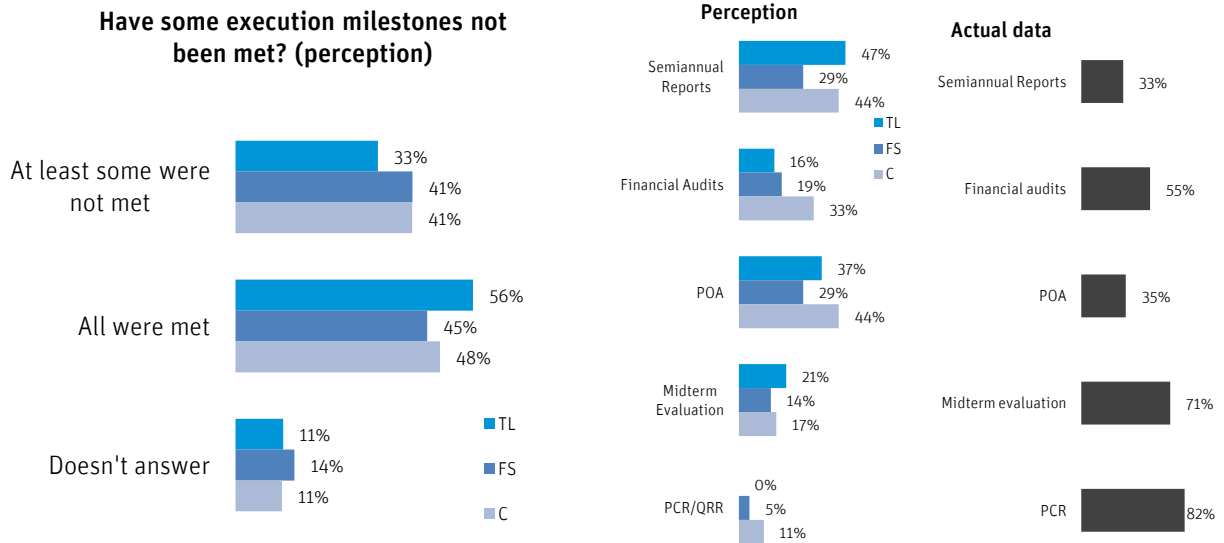
Graph 5. Total compliance with Mid-Term and Final Evaluation - Evolution of compliance by annual cohort



Disaggregating responses by the project's degree of maturity, one can observe lower that Team Leaders and Project Coordinators underestimate compliance with semiannual reports. The Team Leaders' perception of a 47% rate of non-compliance with the completion of semiannual reports contrasts with a more positive reality of a 33% non-compliance rate. On the contrary, Team Leaders were overconfident in respect to the degree of compliance with Mid-term Evaluations. While only 21% reported this instrument to be missing, in reality 71% were not compliant.

²⁵ Close to 10% of project teams did not answer this question.

Graph 6. Real lack of compliance vs. perception of lack of compliance of M&E milestones



This technical note takes into account preliminary recommendations from OVE and aims to respond to the challenges identified with the introduction of a more comprehensive framework. An external evaluation of the RPG Program and a draft OVE evaluation have coincided in indicating that monitoring and evaluation based on results is not sufficiently mature and does yield poor indications of accountability results and lessons to learn. Further, OVE recommended that “much more attention be placed on the monitoring and evaluation of RPG projects, from the design of these tasks to the collection of the relevant data, and to the production of all the required reports,”²⁶ and that indicators be included in the monitoring and evaluation provisions that correspond to outcomes and that measure the various concepts contained in the so-called RPG technology and their contribution to the provision of the regional public good.²⁷

²⁶ Guillermo Fernández de Soto, “Final Report of the Self-Evaluation Program Regional Public Goods,” 2010.

²⁷ Idem.

**PART II: Strategic and Conceptual
Considerations for a RPG Monitoring
and Evaluation System**

Chapter 2: Strategic Considerations for a RPG M&E System

“Designing and implementing a multi-level M&E framework for a range of activities operating in diverse contexts is complex”²⁸

The \$70 billion Ninth General Capital Increase and the 2010 Cancun Declaration prescribed “results-based” management as a practice for a better Bank, a practice that gives priority to interventions that promote development. While delivering results is the goal, strengthening M&E is the means to achieve it. The commitment to stronger M&E capabilities and systems has spread across the whole development system.

The OECD Development Assistance Committee (OECD-DAC) summarized the importance of M&E in the context of getting results and promoting development by stating that “evaluation can help us better understand how to achieve development goals – how to get more children educated or improve access to clean water in rural communities. Evaluations provide useful information about which programmes work best and help explain what factors contribute to their success or failure. But it is not enough for evaluators to investigate these topics: their findings must be shared and used.”²⁹

As early as 1991 the OECD countries reached consensus on the first set of principles for development assistance evaluation. Many non-OECD countries and institutions have pledged to follow the OECD principles, as they were periodically revised and updated. The Bank, as part of the DAC Network on Development Evaluation (a subsidiary body of the OECD-DAC), has adopted these principles and good practices for development evaluations.

Monitoring and evaluation: functions and tasks

Monitoring and evaluation (M&E) are separate yet complementary activities that together allow development stakeholders to promote, identify, isolate and prove positive effects from interventions. The DAC Evaluation Network included the activity of monitoring within the challenge of evaluation. Evaluation is “making a judgment about the amount, number, or value of something.”³⁰ Evaluations provide a rigorous assessment of “the significance and worth of development projects, policies, and programs and can be used to improve ongoing and future programs”³¹ while supporting transparency and accountability. It has the power to explain and to be conclusive based on evidence. While evaluation proves the positive effects of an intervention, monitoring scouts characteristics or behaviors, isolating each of them for evaluation and reintroducing lessons from the evaluation findings.

²⁸ Independent Evaluation Group (IEG), World Bank, *Sourcebook for evaluating global and regional partnership programs: indicative principles and standards*, World Bank, 2007, <http://www.worldbank.org/ieg/grp>.

²⁹ Organization for Economic Co-operation and Development, The DAC Network on Development Evaluation Factsheet, OECD, 2010, <http://www.oecd.org/dataoecd/38/42/42500661.pdf> (Accessed Oct. 5, 2011).

³⁰ The Oxford English Dictionary electronic version, Oxford University Press, 2011, <http://www.oed.com> (Accessed Oct. 5, 2011)

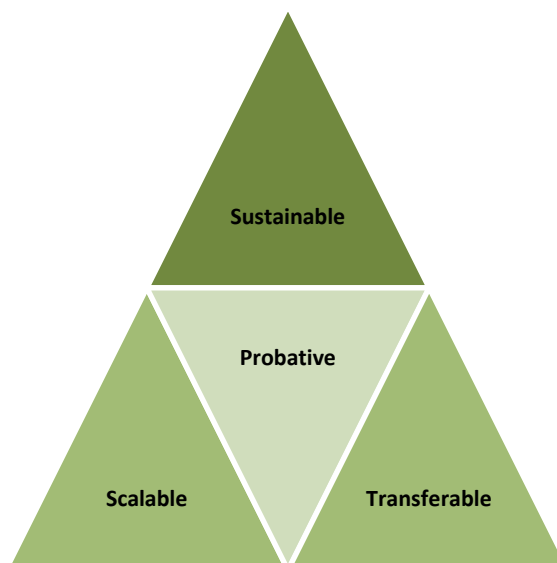
³¹ Independent Evaluation Group (IEG), World Bank, *Sourcebook for evaluating global and regional partnership programs: indicative principles and standards*, World Bank, 2007, <http://www.worldbank.org/ieg/grp>.

According to the Oxford Dictionary, monitoring is the ongoing process of “observing and checking the progress or quality of (something) over a period of time.”³² Further, monitoring is characterized as “a continuing function that uses systematic collection of data on specified indicators in order to provide management and the main stakeholders of an ongoing development intervention with indication of the extent of progress and achievement of objectives and progress in the use of allocated funds.”³³ Given its continuous, systematic and managerial-tool characteristics, monitoring provides substantial value to the promotion of results and feedback for strengthening management. Monitoring without evaluation is not conclusive; evaluation without monitoring may be judgmental and come too late.

The main purpose of an evaluation system, according to the Report prepared by an IDB Commission on Evaluation, is “to contribute to an organization’s effectiveness and its ability to deliver development results. An effective system addresses both the accountability and learning dimensions of evaluation. Accountability refers to the attribution of responsibility for developmental results. Learning means applying the knowledge gained and lessons identified by evaluations to achieve better results.”³⁴

Any M&E system will have manifold functionality, and entail multiple tasks. What makes evaluation an evaluation is its probative function; and therefore it is essential that an M&E system performs the probative function. Beyond this function, it should promote sustainability, transferability and scalability. These other goals supplement the usefulness of assessing an intervention’s positive or negative results.

Graph 7. Multiple functions of a Monitoring and Evaluation system



³² OECD, The DAC Network on Development Evaluation Factsheet, 2010.

³³ Organization for Economic Co-operation and Development, *Glossary of key terms in evaluation and results-based management*, OECD, 2010.

³⁴ Inter-American Development Bank, *IDB Commission on Evaluation Report*, 2011.

An intervention with proven positive results attributable to it should be sustained, while an intervention with no results or negative results should be discontinued summarily. Thus, an intervention assessment should evaluate its sustainability by looking into the possibility of continuing the intervention and at the long-term effects of such continuation. To that effect, the M&E system must be able to identify those elements of the intervention necessary for its sustainability.

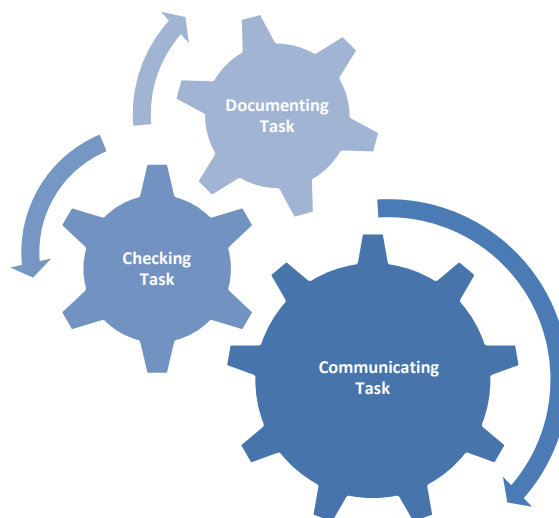
The case for scaling up an intervention is similar to the quest for sustainability in that the M&E system must be able to identify the key elements for scaling up or downsizing it based on its assessed results. The assessment must provide a solid argument in favor of scaling up interventions, focusing on those key elements ready to be scale up as well as a justification for an investment that could yield benefits beyond its original design.

A good intervention could be transferred to a different environment if the conditions of the new environment are also conducive to good results and the intervention is able to adapt to the new stakeholders' needs without losing identity. M&E assessment capabilities must satisfy both conditions for the transferability function.

These four functions - probative, sustainability, scalability and transferability- require the systematic performance of multiple tasks that address the accountability and the learning dimensions of evaluation. These different tasks may be compressed into three: documenting, checking and communication.

An M&E system should be able to document observations collected throughout the monitoring process and the evaluation assessment, in such a manner that allows for transferability and comparison. This comparison is key for the second task: systematic verification or checking. The M&E should be able to check documented observations against performance standards (monitoring) and against logical expectations (evaluation). This task is at the core of the accountability dimension. In order to fulfill the learning function, the findings resulting from the monitoring and the evaluation processes should provide feedback to the intervention, be presented to the stakeholders and to a wider community, and be transferable.

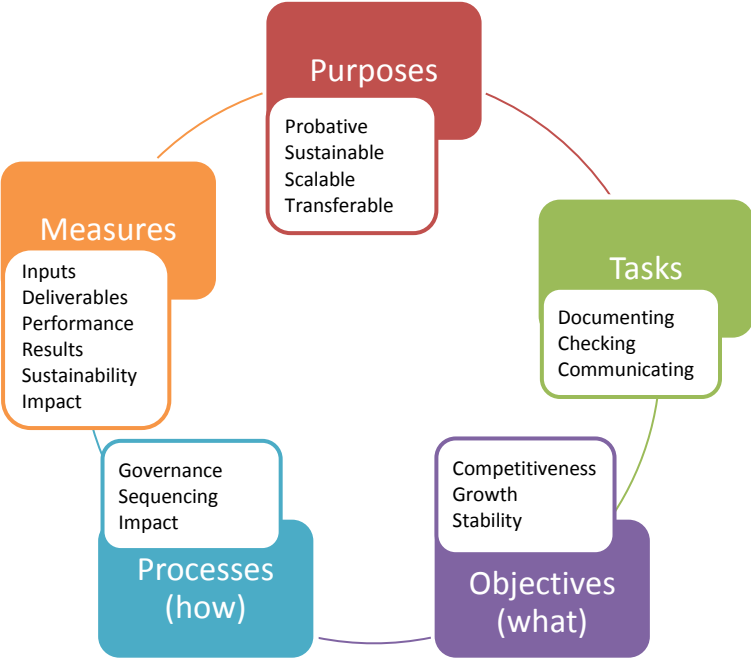
Graph 8. Multiple tasks of a Monitoring and Evaluation system



In summary, it is necessary to prioritize interventions based on their results. Projects and programs require result-based assessments focused on development objectives. Those assessments, however, must include more than a judgment on the condition, quantity and attribution of results. The assessment should include consideration of the different conditions and scenarios for sustaining, scaling up and transferring the interventions with positive results and discontinuing those with negative or neutral results as early as possible. It should be pointed out that in the case of an evaluation of Global and Regional Partnership Programs (GRPP), the OECD (2009) acknowledged that it is particularly difficult to detect attribution.³⁵ This is particularly true for RPG too as all RPG Program goals reflect the means rather than the ends of an intervention.

Therefore, an M&E system for RPGs will have to perform multiple functions; it should support the ability to prove the achievement of results and to identify sustainability and scalability elements related to the Program’s positive effects and the conditions for transferability. In order to perform these functions, the M&E should incorporate at least three tasks: documenting events, checking them against standards, and communicating results and lessons learned. These tasks should be able to measure the ‘what’ and the ‘how’ of an intervention, and plan what to measure and how to measure it. In addition to measuring, such a multi-level M&E framework should support the ability to learn, provide feedback and adjust. This will require flexibility and the power to fine-tune projects and the Program overall.

Graph 9. Multiple dimensions of an M&E system for RPGs



³⁵ Independent Evaluation Group (IEG), World Bank, *Sourcebook for evaluating global and regional partnership programs: indicative principles and standards*, World Bank, 2007, <http://www.worldbank.org/ieg/grp>.

Strategic approach to a RPG M&E system

“Designing and implementing a multi-level M&E framework for a range of activities operating in diverse contexts is complex.”³⁶ A strategic approach to an RPG M&E system will prioritize establishing clear objectives and clear measurement plans from the beginning of each project. Unclear or vague objectives make the task of evaluation impossible. This prioritization will be addressed by quality-at-entry checking. Once clear objectives are identified, it will be necessary to tackle the complexity of multi-level monitoring and evaluation. For a multi-level M&E framework, such as the RPG M&E system, the strategic approach will be based on three premises: (a) the M&E system has to reflect **principled pragmatism** (also called agnosticism by Levy) focused on what works under a set of basic principles rather than pursuing complex rule-based activities that could not yield effective results; (b) the M&E instruments should be able to incorporate a **mixed approach** to address the multi-level nature of the framework; and (c) the M&E framework should respond to a clear but **simple theory of change** that links inputs and outputs with results and impact.

A. Principled pragmatism approach

Any M&E for RPG should be grounded on pursuing what works, on being pragmatic to monitor and evaluate based on principles rather than ideologies and open to different kind of results and approaches to explain the results.

The OECD (2007) supported the proposition of having guiding principles, affirming that “while some variation in evaluation approach and design is to be expected, some standards for evaluation of GRPPs are necessary to ensure credibility, and a minimum frequency is necessary to meet accountability objectives.” Further, it clarified that “the evaluation design, scope, coverage, and methodology may also differ according to the governing body’s purpose in conducting an evaluation at a particular point in time, the maturity of the program, the portfolio size, and the type of activities supported.”³⁷

According to Levy (2010) “the task for development practitioners is to move away from *a priori* prescription, and to become more empirical, learning and adopting what works best as the ways to move forward, with the lessons certain to vary from country to country. What this approach lacks in grandiosity, it may more than make up for in feasibility, and as such prove more effective.”³⁸

Complementing Levy’s pragmatic/agnostic view, in the context of a discussion of information and evaluation systems for integration, De Lombaerde and Van Langenhove (2004) point out that “especially when systems are multi-dimensional (multi-disciplinary) it is useful to have some generally applicable

³⁶ Ibid.

³⁷ IDB, *IDB Commission on Evaluation Report*, 2011

³⁸ The term principled agnosticism applied to public policy was introduced by Brian Levy in his paper “The Case for Principled Agnosticism,” *Journal of Democracy* (2010). This term is defined as “the corollary of adopting an evolutionary approach to development,” with the aim, Levy argues, “to nudge the system so that, ten years on, it performs better in at least some (not all) domains—with the opportunity set for the subsequent ten years to be more attractive than it was at the outset.” Further, Levy compared this principled agnosticism with a rather futile tactical agnosticism where everything is valid everywhere. In a broader sense principled agnosticism is referred to as “the practice of a kind of humility.” The humility towards knowledge and the discovery of what works. What works is based on a set of principles that work as cornerstones to the task of discovering truth.

criteria.”³⁹ These generally applicable criteria are the principles that should govern the application of the M&E function. Such criteria however should not be a straight-jacket. There are more ‘pilot-leds’ that can guide the M&E activities towards the objectives of accountability and learning.

For example, pragmatism as a strategic approach could be used in the design of an evaluation. In this case, the decision on the instrument for capturing data will include considerations on whether to use a rich and detailed system or a simpler, less costly mechanism. The pragmatic approach will encourage the utilization of a “right to the point” instrument for capturing data, i.e. an instrument that is commensurate with the total amount of the project and the resources allocated for the evaluation. A five hundred thousand dollar project will not have a one hundred thousand dollar system to capture data, nor will it have a too sophisticated mechanism for its implementation.

*OECD Principles*⁴⁰

The RPG M&E system should follow international guidelines, norms and standards adopted by the OECD DAC Committee, of which the Bank is member. The sourcebook for regional programs states that “the purpose of the indicative principles and standards... is to improve the independence and quality of program-level evaluations of GRPPs in order to enhance the relevance and effectiveness of the programs.” These principles focus primarily on GRPPs that aim at international development and provide public goods. The RPG should adhere to the OECD DAC set of principles, prioritizing governance, cost efficiency and effectiveness, while also adopting the other OECD general principles for monitoring and evaluation, such as credibility, usefulness and feedback.

Governance: The OECD describes governance as “the structures, functions, processes, and organizational traditions that have been put in place within the context of a program’s authorizing environment to ensure that the [program] is run in such a way that it achieves its objectives in an effective and transparent manner.” It is essential to assess “the legitimacy and effectiveness of the governance and management arrangements” mainly “because the formal programmatic partnership represented by these governance structures is the *raison d’être* of a GRPP.”⁴¹

The OECD report points out that “GRPPs differ from other programs and projects because they have distinct governance mechanisms and processes that affect results. It is important to regard the way in which these mechanisms and processes work in practice, as well as any changes in them over time, as a part of the results chain.” At the same time, “GRPPs have more stakeholders and more diversity among stakeholders than country and local-level programs and projects. Hence, there are more perspectives on the achievement of results and objectives that need to be taken into account.”

As a principle, the RGP M&E will monitor the governance conditions and will derive indicators and measurements of deliverables, results and impact for governance processes and outcomes. At the same

³⁹ Philippe De Lombaerde and Luk Van Langenhove, “Indicators of regional integration: conceptual and methodological issues,” UNU-CRIS Occasional Papers, United Nations University, 2004.

⁴⁰ This section draws from OECD, *Sourcebook for evaluating global and regional partnership programs* and the Organization for Economic Co-operation and Development, *Principles for evaluation of development assistance*, OECD, 1991.

⁴¹ According to the OECD, for regional partnership programs it is also important to assess the distribution of the benefits and costs of the program among the beneficiary partners.

time, monitoring and evaluation activities will be framed under the governance structures internal and external to the Bank, and all program and project stakeholders should be incorporated in the M&E functions, as best as possible.

Cost efficiency: Efficiency is defined by the OECD as “the extent to which the program has converted or is expected to convert its resources/inputs (such as funds, expertise, time, etc.) economically into results in order to achieve the maximum possible outputs, outcomes, and impacts with the minimum possible inputs.” Cost effectiveness is a related concept that the OECD defines as “the extent to which the program has achieved or is expected to achieve its results at a lower cost compared with alternatives.” Given the fact that funding from the RPG program is low in relation to the demand from RPG proposals, the comparison with other alternatives is a principle to follow.

The OECD provides that “GRPP evaluations need to assess the efficiency of the interventions to the extent feasible and to make recommendations for improving the efficient use of resources.” This is more important given that GRPP evaluations require a longer timeframe and a larger budget to achieve a sufficient level of data collection and stakeholder participation and consultation, because of the program’s wide geographic scope, large number of beneficiaries, and multiple operational levels.

Moreover, financial as well as human resources for M&E are a scarce resource too. This condition demands cost efficiency not only as a factor in the monitoring process but also as a principle in the implementation of M&E activities. The key to the cost-effectiveness principle is in its comparative nature, so that M&E activities can be prioritized on those that are more cost-effective. The efficient use of available resources and the selection of evaluation methods and tools will be prioritized as well.

Effectiveness: According to the DAC Glossary and Independent Evaluation Group (IEG) evaluation criteria, “effectiveness (or efficacy) is the extent to which the program has achieved, or is expected to achieve, its objectives, taking into account their relative importance.” As a principle, “All GRPP evaluations need to include an assessment of the effectiveness of the program in order to demonstrate to stakeholders (a) the degree to which the original objectives are being met, (b) whether the program should adjust or restate its objectives or strategies to reflect changing circumstances, or (c) whether the program needs to put in place additional safeguards or compensatory measures to mitigate any negative unintended results.” Moreover, the OECD is emphatic when it requires that “assessments of the achievement of objectives, and of other unintended results, should be evidence-based.”

In addition to assessing the accomplishment of the Program’s objective, effectiveness as a principle applies to the effectiveness of the M&E function as well. The monitoring and evaluation function should acknowledge the need for an effective monitoring system, both to provide the information required for scheduled reporting to the governing body on the use of resources, the progress of activities, outputs, and outcomes, and to provide the information necessary for future evaluations.

As a principle, the RPG M&E will focus on effectiveness assessment. At the same time, monitoring and evaluation activities will be conducted to search for effective positive results as well as to mitigate negative results resulting from the M&E function. Thus, those M&E activities that yield more positive and timely results will be prioritized while adjustments and compensatory mechanisms should be part of the toolbox of M&E recommendations, so that the M&E function supports the accountability and learning M&E objectives.

Other OECD evaluation principles: Among other principles, credibility,⁴² usefulness,⁴³ and reporting, dissemination and feedback⁴⁴ should also be taken into consideration.

The following table presents a glossary of the different concepts articulated around the OECD principles for result-based monitoring and evaluation

B. A Mixed Approach

A mixed approach would allow taking advantage of working at different levels, at different points in the chain of results, and with multiple tools. In relation to regional integration M&E systems, De Lombaerde and Van Langenhove predict that “Independently from the choice of overall objectives, dimensions and contents of the system, a whole range of problems are likely to occur at the moment of implementation.” The problems they recognize “range from the difficulty of measuring a specific variable and the quality of a particular data source to the problems related to managing and funding the system. Other problems relate to the availability, generation and frequency of the data needed to feed the system, and the management of a workable and sustainable system of indicators. Sustainability is likely to involve some degree of co-responsibility from the side of the data-generating institutions. The quality of the data is needed to guarantee comparability.”⁴⁵

In order to manage these risks without shortchanging the scope of the monitoring and evaluation goals and to integrate different dimensions of analysis, it would be necessary to use a mixed approach that would blend the application of criteria and principles with what actually works from the pragmatic/agnostic approach. These dimensions of analysis encompass elements such as the level in the results chains (outputs, outcomes and impacts etc.), the level of aggregation of activities (project or program), the level of implementation of activities (global, regional, national, and local), the categories of stakeholders (owners, supporters, strategic partners and allies), among others.

The OECD recommends measuring “the program’s inputs, progress of activities, outputs, outcomes, and impacts at all levels — global, regional, national, and local...” and finding “a way to present in summary form the results from the local and national levels and the way in which they affect results at the regional and global levels.” It also recommends “identifying the various categories of stakeholders early in the planning for a GRPP evaluation and taking account of their diverse interests.”⁴⁶

⁴² The credibility of evaluation depends on the expertise and independence of the evaluators and the degree of transparency of the evaluation process. It requires that evaluations report successes as well as failures. Recipient countries should, as a rule, fully participate in evaluation in order to promote credibility and commitment.

⁴³ To have an impact on decision-making, evaluation findings must be perceived as relevant and useful. They should fully reflect the different interests and needs of the many parties involved in development co-operation. Easy accessibility, clarity and being concise are also crucial qualities to usefulness. The evaluation process itself promotes a further clarification of objectives, improves communication, increases learning, and lays the groundwork for follow-up action. Evaluations must be timely in the sense that they should be available at a time which is appropriate for the decision-making process. This suggests that evaluation has an important role to play at various stages during the execution of a project or program and should not be conducted only as an ex post exercise.

⁴⁴ Evaluation reporting should be clear, as free as possible of technical language and include the following elements: an executive summary; a profile of the activity evaluated; a description of the evaluation methods used; the main findings; lessons learned; conclusions and recommendations

⁴⁵ IDB, *IDB Commission on Evaluation Report*, 2011

⁴⁶ OECD, *Glossary of key terms in evaluation and results-based management*, OECD, 2010.

The mixed approach will allow the M&E System to assess multiple dimensions in order to have a comprehensive view of results. For example, in the case of a project for coordination of SME policies, it is not only valuable the dredge of coordination among public agencies but the public value created for those SMEs served by the public agencies. One dimension from which the mixed approach will capture gains will include the public sector as well as the private sector or civil society sector. Another dimension that complements the sector dimension is the program level and the project level dimension, where the evaluation will capture the degree of coordination as an outcome of the program dimension and the policy or policy instrument as an outcome of the project dimension

C. Systematic Logic Chain Approach

As the back-bone to a pragmatic approach for M&E, the OECD recommends applying a theory of change that logically links inputs to outputs to results and impact. The OECD sourcebook for GRPP states that “a theory-based approach helps to track the influences at different points in the results chain and to enhance understanding of when or why the program works well or not.”⁴⁷ Carol Weiss (1995) defines a theory of change quite simply as a theory of how and why an initiative works. Following Weiss’ definition, the RPG M&E must declare why and how RPG interventions produce results common to all projects, and then focus the monitoring and evaluation activities on proving if this was done (or not). Principled pragmatism requires a theory flexible enough to allow learning and well-grounded enough to ensure credibility.⁴⁸

Authors such as Connell and Kubisch concurred that “an evaluation based on a theory of change... identifies what to measure, ultimate and interim outcomes and the implementation of activities intended to achieve these outcomes, and helps to guide choices about when and how to measure those elements.”⁴⁹ De Lombaerde and Van Langenhove highlight the scope of an M&E system for regional integration issues and the need to complement it with “choices regarding the theoretical framework used.” “Decisions on whether to include or not specific variables into a [System of Information of Regional Integration] SIRI will also be dependent on the theoretical affinities of the SIRI” they point out.⁵⁰ More importantly, Connell and Kubisch found that “the [theory of change] approach helps avoid the risk that evaluations will be driven by the tools themselves.”⁵¹

Connell and Kubisch identified three attributes as prerequisites to proceed with a theory of change, and argued that these should be endorsed by stakeholders and revisited during the implementation and evaluation of the initiative. As prerequisites to its application, a theory of change should be **plausible** (Do evidence and common sense suggest that the activities, if implemented, will lead to desired outcomes?); **doable** (Will the economic, technical, political, institutional, and human resources be available to carry out the initiative?); and **testable** (Is the theory of change specific and sufficiently complete for an evaluator to track its progress in credible and useful ways?).⁵²

⁴⁷ Idem

⁴⁸ Carol Hirschon Weiss, “Nothing as Practical as Good Theory: Exploring Theory-based Evaluation for Comprehensive Community Initiatives for Children and Families,” in *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, ed. James Connell et al. (Washington, DC: Aspen Institute. 1995).

⁴⁹ James P. Connell and Anne C. Kubisch. *Initiatives: Progress, Prospects, and Problems* (Washington, D.C.: Aspen Institute, 1998).

⁵⁰ IDB, *IDB Commission on Evaluation Report*, 2011.

⁵¹ Connell and Kubisch. 1998.

⁵² Idem

The OECD identified the following main evaluation issues connected with a theory of change.

Box 1. Basic groups of evaluation issues for a theory of change approach to M&E

Rationale: Does the undertaking make sense? Are the objectives relevant and realizable? Should alternative objectives be considered?

Objectives Achievement: To what extent were the original objectives achieved, or likely to be achieved? What were the major factors influencing the achievement or non-achievement of objectives? If objectives were not stated clearly enough to allow for an evaluation of goal achievement, an assessment of impact and effects of the activities undertaken should still be attempted.

Impacts and Effects: What has happened as a result of the project/program? This involves not only direct outputs but, very importantly, the basic impacts and effects on the social, economic, environmental and other development indicators resulting from the activity.

The Overall Results: How successful was the undertaking? Why? Do impacts and effects justify costs? Were the objectives achieved within time and within the budget? Were there any major shortcomings? Were there major achievements?

Sustainability: The question of whether achievements are sustainable in the longer run is of critical importance.

Alternatives: Are there better ways of achieving the results?

Lessons Learned: What are the general lessons which can be drawn and which should be borne in mind when embarking on future programs?

The construction and implementation of the RPG M&E theory of change will be based on a blend of the Connell and Kubisch attributes and the OECD evaluation issues. An operational translation of the theory of change is called the chain of results. The next chapter analyses the chain of results in more depth.

In other words, a strategic approach for an RPG M&E system should consist of doing what works (principled pragmatism/agnosticism), taking advantage of working at different levels, at different points in the chain of results, and with multiple tools (a mixed approach), and working within a framework that specifies a logical and cumulative chain of results (theory of change). This strategic approach should be paired with a set of principles based on the OECD standards for GRPP and the broader standards for development evaluation.

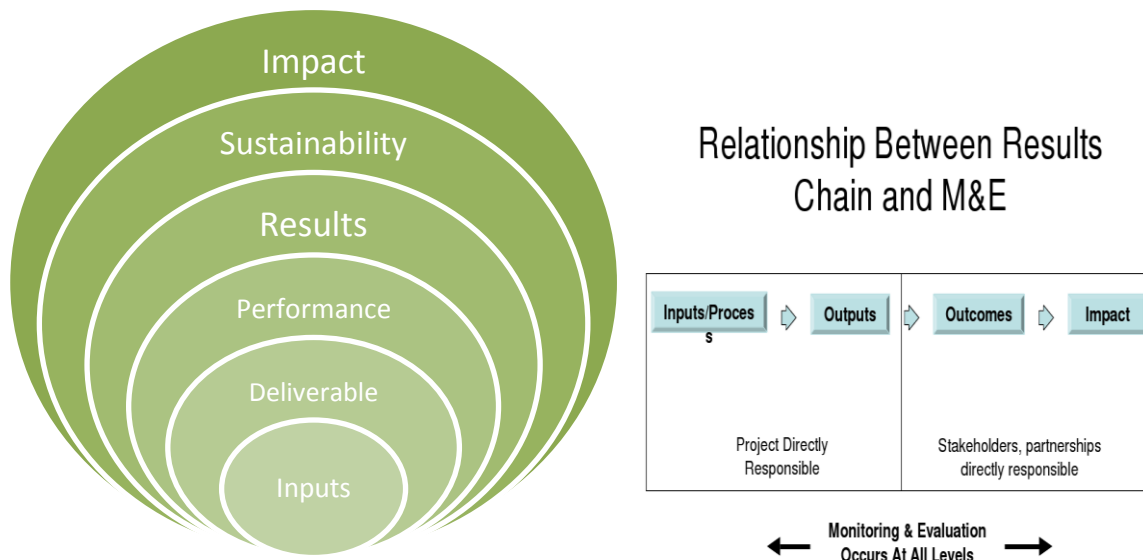
Chapter 3: Conceptual Considerations for a RPG M&E System: the chain of results

“Conceptually, the self-evaluation models have evolved over the years, from systems focused primarily on measuring inputs toward systems that more adequately measure outputs, outcomes and impacts”⁵³

The literature identifies at least three different kinds of evaluation that focus on measuring different aspects: (i) formative evaluation, which focuses on improved performance before and during implementation (project, program or policy); (ii) summative evaluation, which focuses on outcomes (consequences); and (iii) prospective evaluation, which focuses on particular questions such as whether this program/project/policy is worth evaluating, or whether the gains will be worth the effort/resources to be expended. Moreover, the debate on measurement has evolved from measuring inputs to measuring impact, passing through the measurement of deliverables, performance, results and sustainability.

For all these three different kinds of evaluations, measurements interconnect by a chain of results, based on a theory of change, and become a practical tool for M&E. The Independent Evaluation Group of the World Bank Group (IEG) defined the results chain as “the causal sequence for a development intervention that stipulates the necessary sequence to achieve desired objectives - beginning with inputs, moving through activities and outputs, and culminating in outcomes, impacts, and feedback.”⁵⁴

Graph 10. Chain of Results



⁵³ IDB, *IDB Commission on Evaluation Report*, 2011

⁵⁴ World Bank, “The results chain: types of results framework, M&E training for MoHSW & Business Coalition on HIV/AIDS,” http://gametlibrary.worldbank.org/FILES/447_The%20Results%20Chain%20Introduction.pdf (Accessed Oct. 5, 2011).

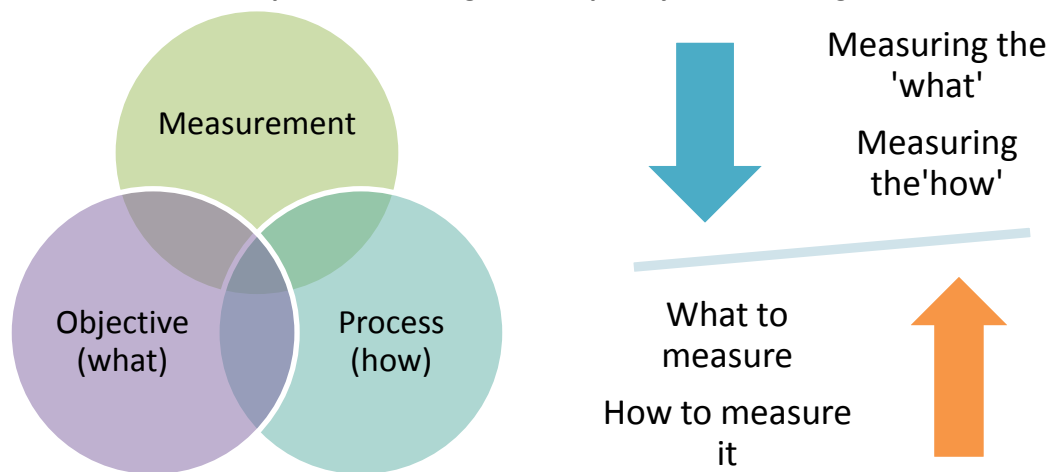
Conceptual basis for a chain of results framework for RPG

As introduced in Chapter 2, all RPG Program goals reflect means rather than ends. Monitoring and evaluating means poses a twofold challenge: in contrast with ends, the means usually have to be measured through proxies rather than by direct measurement; and secondly, although means associated with process do contribute to impact, they are not the impact itself. Thus progress may be measured but not results.

The use of a proxy to capture the progress in the means towards the end adds another layer of complexity. The RPG Program requires measurement of two different dimensions: the 'what is done' and the 'how it is done'. Measuring (i) processes (how) and (ii) results from processes (what) by any M&E system entails multi-dimensional architecture that multiplies the complexity of the M&E function.

The debate on the 'what' includes approaches focusing on competitiveness, growth, stability and volume, among others, as proxies for the impact of integration. Regarding the 'how', the current debate emphasizes taking into account the process of decision-making and governance, the legitimacy of the process, its sequencing, and its ability to make long-lasting changes.

Graph 11. The triangle of complexity for measuring means



Institutionalization and integration could be used, respectively, as proxy measures for capturing the 'how' and the 'what' dimensions of the public goods promoted under the RPG Program. Advancing integration, strengthening institutionalization, or even if narrowed down to promoting public goods, are means towards the benefits of consuming those goods, collectively and from a strong institutional platform. These two proxies, integration and the institutionalization, encompass changes to the 'what' and the 'how' as observed in the following debates on these dimensions.

Performing the function of assessing RPGs results on the 'what' dimension will require a prioritization of the formation and consolidation of relationships that permit transactions to occur. This systematization

of transactions through institutions established for that purpose resembles the strengthening of market institutions, not only for the real economy but for the political economy as well. This approach rests on the correction of market failures, particularly of coordination failures, through the creation of institutional capabilities. It serves two objectives: (i) it closes the divide between the ‘what’ and ‘how’, linking integration with institutionalization; and (ii) it provides well-defined ends (transactions and platform for transitions) that appear early on the chain of results the M&E system will monitor.⁵⁵

The ‘how’ debate is as wide-open and complex as the ‘what’ debate. The process for achieving results from the RPGs is commonly linked to the process of institutionalization. Yet, when regarding institutionalization as a means, what to measure is not a straightforward proposition. This aspect “is by far the most ambiguous and difficult for which to develop satisfactory indices” affirmed Nye (1968) and can only be measured indirectly.⁵⁶

Scholars have explored using several proxies to capture progress in respect to institutionalization, including legitimacy, degree of formalization and enforcement capabilities. Brian Levy presented an articulated framework of the regulatory function applied to the global environment. His framework focused on the “effectiveness [of institutional arrangements] in achieving their intended purposes, whatever these may be.”⁵⁷ Using as a point of departure the work of the 2009 Nobel prize-winning economist, Oliver Williamson, Levy articulated three sets of functions for governance structures: rule-making, monitoring, and enforcement. Levy is of the opinion that for measuring the process of institutionalization, these three functions will use three key channels. The **content** of the rules, the **comprehensiveness** of coverage, and the **credibility** of compliance with the regulatory regime will influence the effectiveness of globalized regulation.

Levy’s set of functions could be completed by a fourth function proposed by Roth (2002), **matchmaking**, to build a complete measurable structure. Roth introduced the matchmaking function (or designing marketplaces) as the function to fix market failures. In this regard, he says that “as we have dealt with more market failures it has become clear that the histories of [particular] markets, and the market failures that led to their reorganization into clearinghouses, are far from unique. Other markets have

⁵⁵ A systematic transaction platform (or operating markets at the micro or macro levels as well as at the real or political economy) is seen as the basis for growth through the different channels highlighted by Sala-i-Martin and is the conclusion of institutionalization following Levy’s framework. In this sense, an “encapsulated local system of tasks, transfers and agents can become an autonomously governed organization in the larger economy” (Baldwin and Clark (2006)) linking this approach with the means of institutionalization as well. Regarding Sala-i-Martin see Carliss Y. Baldwin and Kim B. Clark “Where Do Transactions Come From? A Network Design Perspective on the Theory of the Firm.” [Harvard NOM Working Paper No. 06-12](#), 2006.

⁵⁶ Nye distinguished four different types of political integration that have been prominent in the literature: institutional, policy interactive, attitudinal, and the security community concept. See Joseph S. Nye, Comparative Regional Integration: Concept and Measurement. *International Organization*, Vol. 22, No. 4 (1968): 855-880.

⁵⁷ For the purposes of M&E, this approach is as conducive to accountability as the normative one, but more forward-looking and richer in terms of the potential for learning. Levy himself contrasted his approach with a more normative one, as an alternative focused on the normative challenge – the putative gap between market and socially-preferred outcomes (the classic welfare economics rationale for economic regulation). See Brian Levy. “Innovations in globalized regulation: opportunities and challenges,” Policy Research Working Paper no. 5841. World Bank. October 2011.

failed for similar reasons, and some have been fixed in similar ways.” It is the efforts that an institution makes to fix the market that give rise to the matchmaking function.⁵⁸

Of the three channels identified by Levy, the content relates to the ‘what’ of institutionalization, linking the dimension of the ‘how’ with the dimension of the ‘what’. The same link was proposed by Sala-i-Martin with the description of his seventh channel (institutions) for affecting the overall growth rate with integration. It refers to the “positive effect that openness has on institutions, policies and the political process itself.”⁵⁹ Roth adds to this by noting that “transactions and institutions matter at a level of detail that economists have not often had to deal with.”⁶⁰

The two other channels described by Levy, comprehensiveness and credibility, relate to governance itself. Comprehensiveness is quite straight-forward and questions the coverage or scope of the governance structures. “Rule comprehensiveness influences effectiveness via the extent to which the globalized rules cover those countries and companies that significantly affect the outcome that is being regulated.” Credibility may be more complex. Levy identified several means by which credibility influences effectiveness: (a) Rule-process legitimacy: the extent to which the process strengthens the perceived legitimacy of the globalized rule-making effort among the full range of relevant stakeholders; (b) Monitoring quality: Credible and cost effective (both financial and transactions costs) observation of whether those who have agreed to abide by the rules are indeed following through on their commitments; (c) Enforcement quality: enforcement influences effectiveness directly insofar as credible costs of non-compliance create a credible incentive for compliance; (d) Monitoring and enforcement legitimacy: credible monitoring and enforcement enhance the perceived legitimacy of regulation – strengthening the incentive for self-enforcement, thereby both reducing transactions costs of compliance and extending voluntary participation, resulting in more widespread compliance.⁶¹

Viewing governance as an integrated process, the OECD concludes that regional programs “employ a diverse array of governance models associated with the history and culture of each program. Therefore, it is not practical to base the assessment of governance and management on a particular governance model. Rather, the report suggests that the assessment should be based on compliance with seven generally accepted principles of good governance: legitimacy, accountability, responsibility, fairness, transparency, efficiency, and probity.” Assessing governance based on principles avoids complexity and bias of one model over another.⁶²

The method described by Kaufmann, Kraay and Zoido-Lobaton (1999) could be used to bring the principles to operational measurement. Arranging different observations around thematic clusters,⁶³ or dimensions, of the governance process allowed the creation of a composite measurement that accounts holistically for most of the process. Some of the topics highlighted by Kaufmann, Kraay and Zoido-Lobaton were (i) the process by which those in authority are selected and replaced (linked to the ease of

⁵⁸ Alvin E. Roth, “What have we learned from market design? Hahn Lecture,” *Economic Journal*, 118 (2008): 285-310

⁵⁹ Carliss Y. Baldwin and Kim B. Clark, 2006.

⁶⁰ Roth, 2008.

⁶¹ Brian Levy, *Innovations in globalized regulation: opportunities and challenges*, Policy Research Working Paper No. 5841. World Bank. October 2011.

⁶² Organization for Economic Co-operation and Development, *Glossary of key terms in evaluation and results-based management*, OECD, 2010.

⁶³ The clusters are: a) voice and accountability, b) political stability, c) effectiveness, d) regulatory burden, e) rule of law, and f) graft.

the group-formation dynamics and comprehensiveness), (ii) the ability to formulate and implement sound policies, and (iii) the respect of the citizens and the state for the institutions which govern their interactions (linked to credibility).⁶⁴

The operational recommendations from Kaufmann, Kraay and Zoido-Lobaton linked with Levy's functions and channels lead to a focus on strengthening governance dynamics for efficient matchmaking, rule-making and self-sustaining collective action. These processes prioritize the 'how' dimension and could be measured by the smoothness of the group-formation process (selection, replacement, representativeness, etc.) plus comprehensiveness and credibility. Group-formation includes the formation and the consolidation of transactions by a group that acts as a platform for relationships. This articulation of the analysis of transactions and the analysis of group interactions finally connects the 'what' (transactions) and the 'how' (group-formations).

In summary, proxies, such as regional integration and institutionalization, address the 'what' and the 'how' dimensions of RPG interventions while the chain of results concept is the backbone to the measurement dimension. Using the concept of transactions and group-formation as an analytical framework to assess the 'what' and the 'how', the chain of results can focus on one main element: the solution to coordination failures. The following section presents the chain of results constructed on the concept of correction of coordination failures through groups as platforms for relationships.

A chain of results for RPG

RPGs operations are interventions to solve coordination failures through their collective action and active governance. At their early stage (as those promoted by the Bank under the RPG Program), RPGs take the form of policy interventions (rather than investment interventions which respond to later stage RPGs or other types of instruments). This approach aligns with the latest recommendations from leading scholars of transnational public goods (see, for instance, Kaul's work 2003-2010) who focus on governance and collective action. The IDB BOD prioritized this focus during the 2003-2004 discussions leading to the creation of the RPG Program. Under the Bank's program, Regional Public Goods are policy interventions to solve coordination problems by means of the collective action and joint governance of three or more countries.⁶⁵

The chain of results starts with its two main resources: money and commitment to explore coordination options. The logic that connects money and commitment with the benefits of coordination is based primarily on the work of two scholars. The first of these is Robert Axelrod, who in his seminal work "The Evolution of Cooperation" (1984), introduces the theory that cooperation is possible, desirable and

⁶⁴ Daniel Kaufmann, Aart Kraay, and Pablo Zoido-Lobaton, Aggregating governance indicators, Policy Research Working Paper No. 2195, World Bank, October 1999.

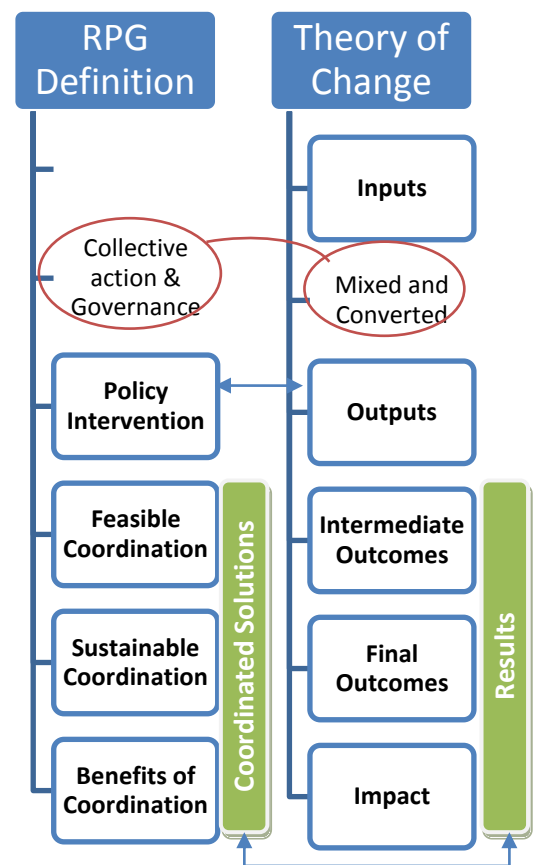
⁶⁵ The chain of results used by the M&E RPG System will not differentiate between transnational and common problems addressed by the intervention. Traditionally RPGs are associated with problems with cross-border effects. Although a differentiation between these problems and the common problems would lead to a more precise account of the efforts made to accomplish coordination, recent experiences and multidisciplinary applied research have shown that common problems can benefit as much as cross-border challenges from collective solutions such as RPGs. In order to keep the M&E system simple and effective, it will compensate by using different methods for valuation at the level of the gauges and not at the level of the logical theory of results.

requires simple conditions to occur and last.⁶⁶ Among those conditions, it just requires a long shadow of the future, the disposition to interact and the ability to recognize the other party. Alvin Roth, one of the leaders in Market Design studies and the design economics discipline, presents a conceptual framework paired with several tools for microeconomic engineering that target the process of designing marketplaces to fix market failures.

The two resources, money and commitment to explore coordination, are translated into outputs in order to obtain intermediate outcomes, final outcomes and later on, impact. These three products are hereafter called 'results' and connect with coordination, RPG's main goal. As a first stage, it is necessary to verify an increased probability that coordination will happen. Then, there is an expectation of feasibility of the coordination or at least a clearer expectation of attainable feasibility. Once it has been proved to be feasible, the coordination must also be sustainable. This sustainability could come from an unstructured but solid volume or steady flow of coordinated interactions between actors or more commonly from a structured platform designed for that purpose (generally a correction on an existing market platform or an adjustment on a new market one). After the coordination is attained and has been proven to be sustainable, the benefits expected since the beginning (regional and national public and semi-public benefits) are reckoned to measure the impact. This measurement must include the positive expected effects plus the positive spillovers and unexpected effects, discounting any negative effect expected or unexpected. Graph 12 represents the linkage between the definition of RPGs and the elements of the theory of change.

These two frameworks are blended to forge a solid and simple chain of results or logic chain that identifies several of the main requirements for Axelrod's cooperation through a systematized and structured process that targets coordination failure, a particular type of market failure.

Various scholars or practitioners have made recommendations for enhancing the chain of results at different points. Inge Kaul recommends (in relation to transnational public goods) adopting an expanded empirical definition of public good, framing the challenge as a governance challenge, integrating the multi-disciplinary approach for policy options and taking the full political process into account, among others. Brain Levy suggests differentiating small governance (small-g) interventions from big governance (Big-G) interventions (linking small g with innovation processes), and paying attention to bridging the island of effective collective action (or micro-groups among bigger groups). Another framework focuses on the the use of RPG technology, which compiles lessons learned through the years of the Program, based on the practitioners' experiences and findings. The technology focuses on two pillars (a) Collective action and (b) governance, plus four strategies (c) demand-driven prioritization, (d) south-south innovation, (e) bottom-up approach and (f) integrating strategic partners.



⁶⁶ Robert Axelrod, Evolution of Cooperation (New York, NY: Basic Books, 1985).

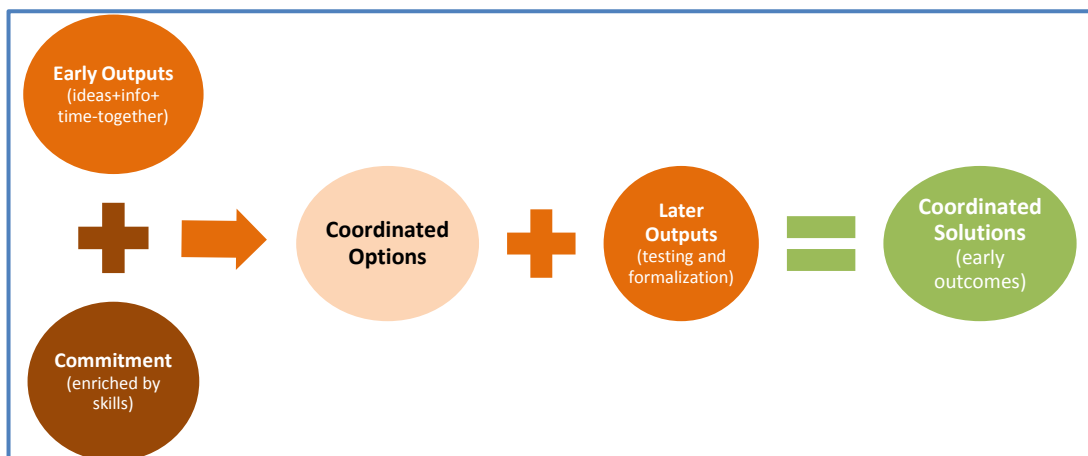
The following list presents a summary of the main conceptual frameworks and their contribution to the RPG logic of change:

- Cooperation Theories (Axelrod) – Requirements and strategies to foster cooperation and coordination
- Market Design (Roth) – engineered approach for correcting coordination failures and considerations for advancing markets
- Transnational public goods (Kaul) – explanation of particularities of RPGs and integration with the governance approach
- Principled Agnosticism Governance (Levy) – differentiation of types of policy reforms (interventions) and their strategic considerations
- RPG Technology (RPG Program and practitioners) – lessons learned for successful implementation.

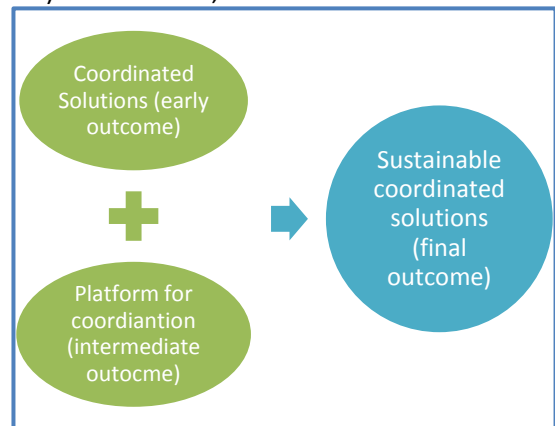
Money is the most versatile resource that goes into the chain. It is used to buy: ideas (that in time will support the development of coordination options); information that serves as a starting point for assessing ideas; coaching services to acquire the skills necessary to both process ideas and advance those ideas selected as options to try out; conditions necessary for discussing and coordinating options (buying time together); formalization arrangements and testing methods (and tools) to try out options or scaling up coordinated solutions.



Early outputs can be organized into three categories: (i) early outputs that feed the process; (ii) outputs that enrich other inputs; and (iii) later outputs that build-up the process. In the first category, there are ideas, information and time together; these are direct outputs that will mix with the other input: commitment. The second category is represented mainly by coaching services that complement the commitment input, in order to enrich it with skills required to address particular development problems. Finally, once ideas, information, time-together, commitment and skills have been combined to produce coordinated options, in order to move coordinated options (intermediate outputs) down the chain of results towards coordinated solutions (early outcomes), it will be necessary to incorporate testing methods and formalization methods in the mix.



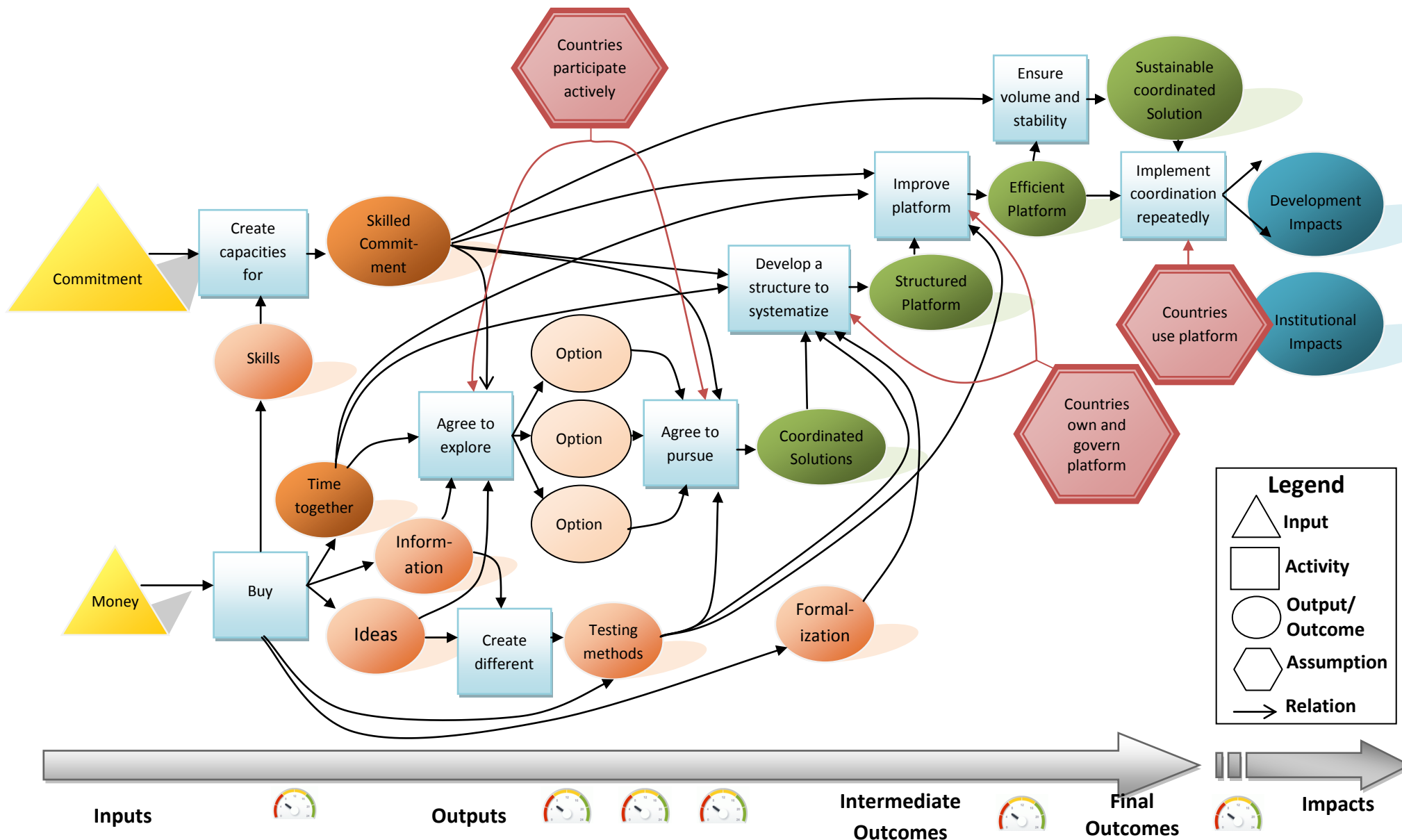
Agreeing on a coordinated solution implies a positive change in behavior from each of the participants. This process is costly and should be systematized to produce the coordination products in a cheaper, faster and more sustainable manner. Otherwise, as Roth (2002) reminded us, “if the market outcome is unstable, there is an agent or pair of agents who have the incentive to circumvent the match.” The systematization of the process will require activation of a platform (inexistent or previously inefficient) for coordination. The coordinated decision by the participants on the parts, functions and roles of the platform would be an intermediate outcome that specifically focuses on governance by the whole set of stakeholders. Once the platform is complete, making it sustainable would be the final outcome. Benefits from using the platform, i.e. the adoption of relatively low-cost and stable coordination among the participants, should produce an impact. Impact can be observed in financial savings, more access, better quality, etc. following the chain of results modeling technique introduced by John Thorp in 1998.⁶⁷



The following diagram presents a complete chain of results, from inputs to impact on results.

⁶⁷ John Thorp, *The Information Paradox: Realizing the Business Benefits of Information Technology*, Center for Strategic Leadership, 1998

Graph 16. Results Chain for Regional Public Goods



PART III: Blueprint of a RPG Monitoring and Evaluation System

Chapter 4: A blueprint for the RPG M&E System

The blueprint for the RPG M&E system consists of a platform with six instruments and a series of gauges reporting on pre-identified indicators along a chain of results path. The gauges will be placed at locations along the path of the chain of results, and will report on the performance of indicators with respect to inputs, activities, outputs, intermediate and final outputs related to impact.

As discussed in the previous section, the chain of results describes the conceptual foundation and connects the different pieces that lead to results and finally to impact through a logical discourse. Along the chain of results, the theory of change places a series of gauges that focus in on a set of key measurements in time: the indicators. In addition to the conceptual and logical framework, a set of instruments supports the systematic capture and processing of information, calling the alerts or disseminating the conclusions it processed.

The OECD recommends that impact evaluations be conducted for subsets of activities. This approach will result in better measurability than for the program as a whole, or and will support an assessment of impact to influence design adjustments or decisions on replicability and scaling up. Further, the OECD identifies a few different approaches to impact evaluation, including “quantitative impact evaluation, participatory impact evaluation, and theory-based (program logic) approaches” and recommends the use of all three approaches in order to produce good impact evaluations.

The RPG M&E blueprint incorporates these two recommendations in the design of its toolbox: (i) the instruments and gauges target a subsystem, representing a part of the whole chain of results, where it is possible to obtain conclusions while pursuing the objectives of the M&E overall function (monitoring inputs, tracking the progress of conversion of inputs into outputs, assessing the accomplishment of goals, ensuring sustainability and learning valuable lessons etc.). This allows the system to offer partial conclusions with valuable insight; and (ii) the toolbox contains several instruments that enable the use of different approaches, instead of having a central overarching instrument with a single approach.

Gauges and indicators

The set of gauges positioned along the chain of results function as control instruments for the process and the performance of each intervention. Following the theory of change method for evaluation, these sets of gauges are located at the inputs section, the output section and the outcomes sections. Each gauge will have two back-end inputs: (a) the variable to control and (b) the method to measure it; and one front-end input: the reporting method. Gauges used in the RPG M&E system will be multi-variable. The table below introduces the indicators proposed for each set of gauges. Each gauge will be wired to a particular measurement instrument, and each will have a specific reporting method. The standard methods used by the Bank for the M&E function will be the starting point in the design of a measurement method, enriching it when possible. Some gauges will report throughout a control panel, similar to a dashboard, while others will take a written report format. The following section gives more detail on each instrument.

Graph. 17 Types of



Table 1. Indicators for each set of gauges

Gauge Family	Indicator Family	Indicator
Early Inputs/Outputs	Fiduciary Alerts	Time to start disbursing
		Speed of disbursements
		Time since last disbursement
		Time before expiration date
	Documented Interventions	Improvement over time for last disbursement
Later Inputs/Outputs	Earn Value Method	Cost Performance Index
		Schedule Performance Index
		Performance Index
Early Outputs/Outcomes	Feasibility	Progress on the feasibility of coordination
	Probability	Increase in the probability of coordination
Later Outputs/Outcomes	Composition of the structure to produce outcomes in a sustainable manner	Completeness of the structure of the platform for coordination
	Efficiency of the structure to produce outcomes in a sustainable manner	Improvement in the efficiency of the platform for coordination (preparedness for long term net gains)
	Sustainability (market design conditions for stable solutions)	Attract a large enough proportion of the potential participants
		Overcome the congestion that thickness (volume) can bring
Impact	Institutional Impact	Institutional comprehensiveness
		Institutional credibility
	Development Impact	Accordingly to each intervention

RPG M&E Instruments

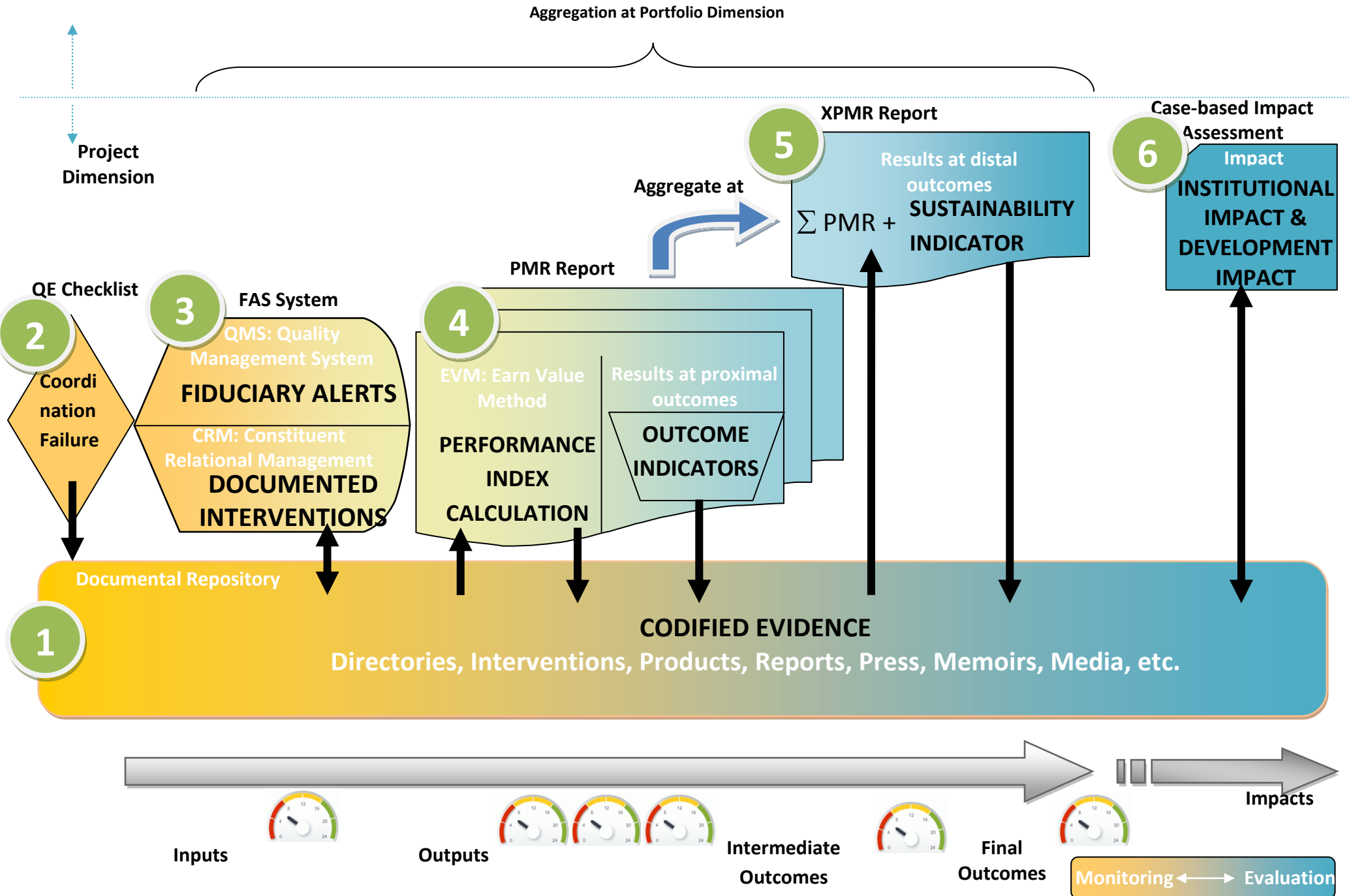
In addition to the series of gauges reporting on pre-identified indicators along the chain of results path, the blueprint for the RPG M&E system contemplates the use of six instruments. These instruments are: (i) a repository for documents and data; (ii) a quality at entry checklist; (iii) a managerial situation dashboard; (iv) the performance accountability system (PMR-like system); (v) the sustainability

assessment report (XPMR-like system); and (vi) case-based evaluation of interventions on the results chain. Each of these instruments represents a subsystem, described as follows:

- (i) Evidence Repository System – the ER System is an electronic repository of documents and data related to the projects included in the RPG Portfolio. It acts like a virtual library of different categories of information products.
- (ii) Quality at Entry Checklist – the QE Checklist is a simple but potent tool to ensure the identification of the coordination failure, the purpose for coordination and the degree of systematization that the coordination requires to become sustainable. It requests a clear and precise objective for the RPG project together with an effective set of evaluation indicators.
- (iii) Fiduciary Alert (fiduciary risk) System – the FAS System combines a Quality Management System (QMS) with a Constituent Relationship Management (CRM) system. The QMS supports the tasks of quality planning, quality control and quality assurance. This component tracks the fiduciary progress of each operation and uses the prediction models to assess the level of urgency to intervene. The CRM documents the quality assurance task and supports the quality improvement task. It tracks the interaction with the different stakeholders once interventions have been rolled out.
- (iv) Performance (operational risk) Reports – the PMR Report is the building block of the evaluation system. It tracks progress on the achievement of intermediate objectives and goals both at the sector and the institutional levels. In addition to monitoring progress at the output/outcome level, the PMR instrument will also identify bottlenecks and lessons learned.
- (v) Sustainability (continuity risk) Assessments – the XPMR Report, similar to the Bank-wide XPMR instrument, will present a complete picture of project performance, including tracking the history of all PMRs. Moreover, the RPG XPMR will focus primarily on the sustainability of achieved coordination.
- (vi) Case-based Impact Assessments – CIA Evaluations is in fact a set of different case-based impact evaluations, designed and implemented for each case.

The following graph (Graph 15) presents an overall view of the instruments toolbox for M&E. It resembles the chain of results, given its coordination with the main logical flow and each of the points of measurement (gauges). After the graph, profiles for each individual instrument will highlight aspects such as the expected usage for each instrument, key issues of technological implementation and the conceptual grounds which support them.

Graph 18. Map of M&E Instruments



RPG M&E Instrument Profile

Name: Evidence Repository System (ER)

Description:

The Evidence Repository System (ER System) is an electronic repository of documents and data related to the projects included in the RPG Portfolio. It acts like a virtual library of different categories of information products. Among the type of products included are: consultancy reports, meetings' aid memoirs, annual operational plans, bi-annual reports, external evaluations and auditing reports, etc. Categories will included operational documents, deliverables, institutional development evidence, results evidence, etc.

The repository will works in a relational manner, linking not only objects (documents, photos or any type of data) to projects, but also categories, media type, authors, comments or ratings of the products and usage in the subsequent M&E instruments. In this way, the ER system will works as an Experts Database, a CRM source of information and the link to Bank-wide systems such as IDBDocs.

Interconnection with other instruments:

Graph

Location:

This instrument serves a function at each gauge, mainly the recording and managing of documentary information related to it. It is a back-office instrument of the M&E System.

Graph in the results chain

Expected Usage:

The ER System will be used to file and organize any type of information related to each project. This system is the building block of any evidence-based M&E system. It will open the opportunity to track performance beyond the fiduciary aspects -recorded in the financial system of the Bank or the Executing Agency-, allowing the identification and analysis of documentary evidence for comparison of outputs and outcomes against project targets or market standards.

It is useful for operational (fiduciary, sector and programmatic) decision making as well as for evaluation assessment and learning.

Responsible:

Coordinator + Team Leader + RPG Portfolio

Graph organization map

Periodicity:

It will be used in a constant manner during the lifetime of the project and after closing for impact assessments.

Graph annual cycle

Technological implementation:

Relational Database that links the physical location of files both on the internal network, IDBDocs or the internet (servers, youtube, facebook etc.)

File locations are complemented with description of the object, the process of production and governance and the environment where it belongs (project, categories, authors, etc.)

Conceptual foundation:

- Knowledge Management
- Evidence-based Evaluation
- Object-Oriented Data Management

RPG M&E Instrument Profile

Name: Quality at Entry Checklist (QE)

Description:

The QE checklist ensures the identification of the coordination failure, the purpose for coordination and the degree of systematization that the coordination requires to become sustainable. It requires the articulation of a clear and precise objective for the RPG project together with an effective set of indicators for evaluation.

Interconnection with other instruments:

Graph

Location: At the entry of the chain of results

Graph in the results chain

Expected Usage:

This simple checklist will be used to identify the purpose for coordination and the degree of systematization that the coordination requires to become sustainable.

Responsible:

RPG Portfolio Administrator

Graph organization map

Periodicity: Before the approval (after QRR) or at the beginning of execution

Graph annual cycle

Technological implementation:

Through a web-based checklist, similar to the safeguards checklist system.

Conceptual foundation:

- Evidence-based Evaluation
- Market Design adaptation for non-financial markets (Roth)

RPG M&E Instrument Profile

Name: Fiduciary Alerts System (FAS)

Description:

The Fiduciary Alert System combines a Quality Management System (QMS) with Constituency/Client Relationship Management (CRM).

The QMS supports the tasks of quality planning, quality control and quality assurance. This component track the fiduciary progress of each operation, and through the use of prediction models assesses the level of urgency to intervene, signaling alerts as green (Ok), yellow (pay attention) and red (take action). Data visualization is provided at the portfolio, sector/division, sub region, team leader and project level.

The CRM documents the quality assurance task and supports the quality improvement task. This component tracks the interaction with the different stakeholders once interventions have been rolled out. The CRM component serves as a journaling tool of the relationship between the RPG coordination team and the project team (both inside and outside the Bank), and becomes a dynamic entry point for the ER system.

Interconnection with other instruments:

Graph

Location:

This instrument is located mainly at the first gauge, the input/output gauge, with interconnections with the successive ones. Both components (QMS and CRM) are back-office components with a front-end layer of data visualization on top of the QMS.

Graph in the results chain

Expected Usage:

The FAS system will be used to track fiduciary progress and to prioritize actions and interventions to improve performance. This system is the basic node of the monitoring task and serves as the basis for portfolio decision making. This instrument also serves as an early detection system of irregularities in deliverable and outcomes progress.

Responsible:

RPG Portfolio Administrator + Team Leader + Division Chiefs

Graph organization map

Periodicity:

It will have monthly

Graph annual cycle

It is useful mostly for operational (fiduciary, sector

<p>and programmatic) decision making.</p>	<p>updates and quarterly aggregate reporting for management</p>	
<p>Technological implementation:</p> <p>As mentioned above, it combines a tracking and measuring on-target compliance tool with a documentary database. For the first application it will use a multi-dimensional cube nurtured from the Bank's Data Warehouse analyzed by a statistical model built-in the data visualization tool. Documenting will use a relational database (linked as an entry point to the ER system) that will allow the recording of the interactions across interventions.</p>	<p>Conceptual ground:</p> <ul style="list-style-type: none"> • Quality management theories particularly the conceptual approach of “uniformity around a target value” (Taguchi) and the operational approach of Statistical process control (SPC) (Shewhart) • The input/output underlining of the Theory of Change concept 	

RPG M&E Instrument Profile

Name: Performance Report Instrument (RPG-PMR)

Description:

The RGP PMR instrument is the building block of the evaluation system. It tracks progress with respect to the achievement of intermediate objectives and goals both at the sector and the institutional levels. Besides monitoring the progress at the output/outcome level, the PMR instrument will also serve to identify bottlenecks and lessons learned.

This instrument is operationalized by a bi-annual report that integrates the ability to convert inputs into outputs with the assessment of declared achievements and gaps from the executing agency and the judgments of the team leader on both aspects.

Interconnection with other instruments:

Graph

Location:

This instrument interconnects two gauges, the input/output gauge and the output/outcome set of gauges. With respect to the input/output gauge, it will complement the FAS through a bi-annual assessment of the project's performance in converting input into outputs.

At the output/outcome level, it is responsible for measurement of: (a) progress on the feasibility of coordination, (b) increases in the probability of coordination, (c) completeness of the structure of the platform for coordination, and (d) improvement in the efficiency of the platform (degree of preparation for long term net gains)

Graph in the results chain

Expected Usage:

The PMR instrument will be used to assess the progress from outputs towards objectives and goals. To measure the performance in the conversion of input to outputs, the RPG PMR will adopt the Bank-wide instrument, the Earned Value Method (EVM), to produce a composite

Responsible:

Coordinator and Team Leader

Graph organization map

<p>Performance Index Calculation (see Progress Monitoring Report Concept Document, pages 8 to 10).</p> <p>As the four gauges at the output/outcome level indicate, the PRM will track progress on the feasibility of the coordination to take place, the overall probability to take place and to keep occurring, and the progress in building a durable and efficient platform for that coordination to occur. It will focus primarily on the (nature and outlook of) risk of underperforming at the four gauges.</p> <p>It is useful mostly for mid-term operational modifications and strategic decision making</p>	<p>Periodicity:</p> <p>Following Bank-wide practices, outcomes will be monitored according to the terms established in the project (annually, at mid-term or at completion), while Outputs will be monitored biannually (in September and March) during project execution.</p>	<p>Graph annual cycle</p>
<p>Technological implementation:</p> <p>A web-based form integrated with the ER system for bringing evidence that supports each gauge measured at the output/outcome level.</p> <p>The application will integrate a model for EVM calculation and part of the CRM component for follow-up activities (as another entry point to the ER system).</p>	<p>Conceptual foundation:</p> <p>At the input/output level:</p> <ul style="list-style-type: none"> • Operation Research methods such as Earned Value Method (EVM) – PERT/CPM approach <p>At the output /outcome level:</p> <ul style="list-style-type: none"> • Evidence-based Evaluation • Market Design adaptation for non-financial markets (Roth) 	

RPG M&E Instrument Profile

Name: **Sustainability Assessment Instrument (RPG-XPMR)**

Description:

Similar to the Bank-wide XPMR instrument, the RPG XPRM will present a complete picture of project performance; including tracked history of all PMRs. Moreover, the RPG XPMR will focus primarily on the sustainability of the coordination achieved.

Congruent with the PMR instrument, the XPMR instrument is operationalized through a final report at the closing stage of the project that integrates the declared achievements and shortfalls from the executing agency and the judgments of the team leader on both aspects.

Interconnection with other instruments:

Graph

Location:

This instrument is located at the fifth gauge of the output/outcome level: shadow of the future of the coordinated relationship. In this control position it will measure the sustainability of coordination.

Graph in the results chain

Expected Usage:

The XPMR will provide a holistic view of project performance, from onset through completion. It will compare actual results achieved against end-of-project targets as outlined in the Results.

This report will also consolidate the lessons learned by the different stakeholders of the project, using reverse engineering methodologies to incorporate feedback in new operations design.

It is useful mostly for accountability purposes and prospective decision making regarding the future of the intervention (the later stage regional public goods); It is

Responsible:

Team Leaders and Executing Agency

Graph organization map

Periodicity:

This report will be fulfilled by different stakeholders only once, at the closing stage of the operations.

Graph annual cycle

also useful as a basis for learning.		
<p>Technological implementation:</p> <p>This instrument will share the technological platform of the PRM instrument. Thus, it will also be a web-based form integrated with the ER system for collecting evidence that supports each gauge measure at the output/outcome level.</p> <p>The application will include a model for EVM calculation.</p>	<p>Conceptual ground:</p> <ul style="list-style-type: none"> • Operation Research methods such as Earned Value Method (EVM) – PERT/CPM approach • Evidence-based Evaluation • Market Design adaptation for non-financial markets (Roth) • Reverse engineering methods for lessons learned 	

RPG M&E Instrument Profile

Name: Case-based Impact Assessment (CIA)

Description:

This instrument is in fact a set of different case-based impact evaluations designed and implemented according to each case. For each evaluation, and as a cross-cutting methodology, it will use a before and after methodology for sector results, unless the project itself includes a control group or randomized methodology; and a control group methodology for institutional results according to the M-pie methodology (Multi-party institutional evaluation)

Interconnection with other instruments:

Graph

Location:

The CIA is located at the top level of gauges related to impact. The two measurement indicators tools will track the institutional value added and the developmental value added.

Graph in the results chain

Expected Usage:

In correlation with the two gauges the CIA will assess impact of the institutional and the developmental problem identified by the intervention. The first one will provide an account of the capacity development achieved and the second one will provide an account of the positive impact that capacity is producing. Institutional value added will be measured with similar methodologies for all cases, while developmental value added will be selected in accordance with the developmental goals of the project (such as infants better nourished, improvement in recovery from disasters, etc).

There will be a strategic selection of operations for CIA. The selection will be based on such criteria as the availability of resources, the feasibility of assessment, and the degree of integration with future operations, among others.

Responsible:

RPG Portfolio Administrator

Graph organization map

Periodicity:

This assessment can be prepared a considerable span of time after the closing stage of the operations.

Graph annual cycle

Technological implementation:

In accordance with the design of each CIA

Conceptual foundation:

- Evidence-based Evaluation
- M-pie methodology (Multi-party institutional evaluation)

Conclusions

This Technical Note has produced a diagnostic of the M&E function for Regional Public Goods based on the current mandates, governance, structure and performance. It further develops strategic and conceptual considerations for an updated M&E system for RPGs, including a strategic approach to such a system and a conceptual basis for a chain of results framework. Based on the diagnostic and the strategic and conceptual considerations, this Technical Note presents a blueprint for a RPG M&E system that would support the monitoring and evaluation of impact and performance and audit results and processes. To that end, it introduces a platform with instruments, measurement gauges and indicators.

In producing this blueprint, the Technical Note reaches punctual conclusions along the path. These include the following:

1. The Bank's general M&E system is weak but is improving. It does not incorporate monitoring characteristics for projects such as the RPGs and, therefore, it is too general to fully serve as the RPG M&E system. At this time, the weaknesses of the Bank's general M&E system are also evident in its application to the RPGs. It will be necessary to design a system for RPG M&E along lines that are consistent with the Bank's general model but that specifically focus on the characteristics that define RPGs.
2. The new model for an RPG M&E system should address the new monitoring and evaluation functions of sustainability, scalability and transferability, beyond the probative function. It should also explicitly incorporate principles of pragmatism, adopt a mixed approach and follow a logical chain.
3. This logical chain should be linked to the key element that defines the target of RPGs: the coordination failure.
4. Monitoring the process towards a solution to the coordination failure will need an integrated vision that entails a broader approach requiring looking at documentation, evaluation of administrative performance, generation of products, and achievement of results. These will be captured by a series of instruments that will build on existing ones.
5. A new RPG M&E system should be designed taking into consideration (i) the choice of technologies, (ii) efficiency gains, and (iii) the Bank's institutional culture. The design should also include a clear definition of roles and a series of incentives for its smooth functioning.
6. There is more to be done in deepening our knowledge on the Bank's institutional culture. There is a working group at the Bank that is exploring the issue of incentives. Their recommendations could feed the proposed blueprint for an RPG M&E system.

Bibliography

Axelrod, Robert, *Evolution of Cooperation*. New York: Basic Books, 1985.

Baldwin, Carliss Y. and Kim B. Clark. *Where do transactions come from? A network design perspective on the theory of the firm*. Harvard NOM Working Paper No. 06-12. Harvard University. 2006.

Connell, James P. and Anne C. Kubisch. *Initiatives: Progress, Prospects, and Problems*. Washington, D.C.: Aspen Institute, 1998.

De Lombaerde, Philippe and Luk Van Langenhove. *Indicators of regional integration: conceptual and methodological issues*. UNU-CRIS Occasional Papers. United Nations University, 2004.

De Soto, Guillermo Fernández. "Final report of the self-evaluation program Regional Public Goods." 2010

Inter-American Development Bank. *IDB Commission on Evaluation Report*. Washington, D.C.: IDB, 2011.

Inter-American Development Bank. Office of Evaluation and Oversight (OVE). *Evaluation of the Bank's processes for managing technical cooperation (RE-364)*, IDB, 2010.

Inter-American Development Bank. Office of Evaluation and Oversight (OVE). *Evaluation of IDB action in the initiative for integration of regional infrastructure in South America (IIRSA)(RE-338)*, IDB, 2009.

Inter-American Development Bank. Office of Evaluation and Oversight (OVE). *Evaluation of IDB's support to the Plan Puebla Panama Initiative (RE-350)*, IDB, 2009.

Inter-American Development Bank. Policy CO-303 *Progress Monitoring Report*, 2010. <http://manuals/go.cfm?do=Page.View&pid=470> (Accessed Oct. 5, 2011).

Inter-American Development Bank. Policy CO-309 *Project Completion Report – PCR*, 2006. <http://manuals/go.cfm?do=Page.View&pid=484> (Accessed Oct. 5, 2011).

Inter-American Development Bank. Policy OA-200 *Monitoring and evaluation of project execution: chapter Summary*. 1998. <http://manuals/go.cfm?do=Page.View&pid=1059> (Accessed Oct. 5, 2011).

Inter-American Development Bank. *Progress Monitoring Report Web Portal*. <http://hqpapmr02.idb.iadb.org/pmrptl/index.do> (Accessed Oct. 5, 2011).

Inter-American Development Bank. *Regional Public Goods - 2009 Report and recommendations for financing* (Document GN-2275-14). Inter-American Development Bank. 2009.

Inter-American Development Bank. *Regional Public Goods: proposal for an expanded Bank role. Final approved version* (Document GN-2275-5). Inter-American Development Bank. 2005.

Inter-American Development Bank. *Sector strategy to support competitive global and regional integration* (GN -2565-4), IDB, 2011.

Kaufmann, Daniel, Aart Kraay, and Pablo Zoido-Lobaton. *Aggregating governance indicators*. Policy Research Working Paper No. 2195. World Bank. October 1999.

Levy, Brian. *Innovations in globalized regulation: opportunities and challenges*. Policy Research Working Paper no. 5841. World Bank. October 2011.

Levy, Brian. "The Case for Principled Agnosticism." *Journal of Democracy* (2010).

Organization for Economic Co-operation and Development. The DAC Network on Development Evaluation Factsheet, OECD, 2010. <http://www.oecd.org/dataoecd/38/42/42500661.pdf> (Accessed Oct. 5, 2011).

Organization for Economic Co-operation and Development. *Glossary of key terms in evaluation and results-based management*. OECD, 2010.

Organization for Economic Co-operation and Development. *Principles for evaluation of development assistance*. OECD, 1991.

Oxford English Dictionary electronic version, Oxford University Press, 2011, <http://www.oed.com> (Accessed Oct. 5, 2011).

Roth, Alvin E. "What have we learned from market design? Hahn Lecture." *Economic Journal*, 118 (2008): 285-310.

Thorp, John, *The Information Paradox: Realizing the Business Benefits of Information Technology*. Center for Strategic Leadership. 1998.

Weiss, Carol Hirschon. "Nothing as Practical as Good Theory: Exploring Theory-based Evaluation for Comprehensive Community Initiatives for Children and Families." In *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, edited by James Connell et al. Washington, DC: Aspen Institute. 1995.

World Bank. Independent Evaluation Group (IEG). *Sourcebook for evaluating global and regional partnership programs: indicative principles and standards*. World Bank, 2007, <http://www.worldbank.org/ieg/grp>.