

FOCEVAL – Promoting Evaluation Capacities in Costa Rica: Smart(er) Implementation with Capacity WORKS?

Sabrina Storm

Executive Summary

The National Monitoring and Evaluation System of Costa Rica and its corresponding laws were established during the 1990s. Since then, the country has endeavored to implement monitoring and evaluation (M&E) activities as part of its public policy framework. Nevertheless, hardly any systematic evaluations had been conducted, and monitoring activities had been reduced mainly to the institutional self-reporting of implementation compliance. Persisting regional disparities and growing levels of inequality among the population raised the level of pressure on the government to present reliable information on the effectiveness of public interventions. Hence, results-oriented evaluations were promoted by some Costa Rican departments as tools that would support evidence-based policy-making while also increasing public-sector accountability.

This paper focuses on the strategies and steps pursued by the Costa Rican government – and supported by the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) program Strengthening Evaluation Capacities in Central America (FOCEVAL)¹ – to improve evaluation capacities within the public sector between 2011 and 2014. Mistrust and a lack of dialogue and communication between different units and departments were some of the challenges GIZ faced. Also, academic and political discourse on evaluation was detached from public-sector conditions. Soon, the program team had to learn that the initial operational plan – which included proposals for large-scale impact evaluations – was over-ambitious and that a much more incremental approach was needed. Commitment, alliances, and common understanding needed to be developed before pilot evaluations could be successfully executed. It is shown how creating space for communication, co-creation, and learning fostered a

1 GIZ implements the program on behalf of BMZ.

joint and feasible approach for implementation within the cooperation system. GIZ's management model Capacity WORKS – with its integral concept of capacity development and systemic orientation toward the success factors of strategy, cooperation, steering structure, processes, and learning and innovation – provided a framework for creating this space. The case study is structured around Capacity WORKS, and thus exemplifies how the management model of GIZ can provide support.

Introduction

When Costa Rica's former Minister of Planning Roberto Gallardo left office in 2014, he was asked in an interview: Which were the five most important issues he would like to hand over to his successor? Evaluation of public policies and programs was one of the topics he chose to put in this list of priorities. "For policy-makers, it is of vital importance to count with information that allows them to understand and assess an ever more complex reality" (Gallardo, 2013).

In the two years preceding this statement, the Costa Rican Ministry of National Planning and Economic Policy (MIDEPLAN²) had just executed the very first strategic evaluations within the framework of the National Monitoring and Evaluation System. It had developed and harmonized guidelines for results-oriented planning, monitoring, and evaluation; had overhauled the compilation process and methodological setup of the National Development Plan; and the ministry's evaluation unit had trained dozens of planning officials from sectorial ministries and agencies. Results-oriented evaluations were promoted as tools to address some of the persisting development barriers of the country more effectively. However, MIDEPLAN needed to align many stars in order to move closer to its vision of an institutional M&E practice, which would not only be technically sound but also useful for learning, improvement, and strategic decision-making in public administration and government.

Costa Rica's national M&E system is composed of an institutional network of at least 100 agencies in 14 government sectors. Although the system had already been created in the mid-1990s, hardly any systematic evaluations had been conducted, and monitoring activities had been

2 See <http://www.mideplan.go.cr/>

reduced mainly to institutional self-reporting of implementation compliance. Standards, procedures, and orientations for planning, monitoring, and evaluation – set out by the Ministry of Finance, the Supreme Audit Institution, and the Ministry of National Planning – were incoherent and caused frustration and confusion within the sectorial and institutional planning units. Persistent misconceptions and a highly diversified understanding of evaluation and its objectives could be observed, with “evaluation” often being used as a synonym for auditing, scrutiny, or control.

This paper focuses on the strategies and steps pursued by the ministry – and supported by the GIZ program FOCEVAL – to strengthen the evaluation component within the National Monitoring and Evaluation System (Sistema Nacional de Seguimiento y Evaluación – SINE) between 2011 and 2014. In 2009 the Costa Rican government, along with the University of Costa Rica, presented a proposal to the German Federal Ministry for Economic Cooperation and Development (BMZ) for a cooperation project to improve evaluation capacities in the country.

FOCEVAL’s first phase (2011–2014) was implemented by GIZ on behalf of BMZ and funded via a supra-regional cooperation fund for Evaluation Capacity Development, which had been set up as a response to the international agreements on aid effectiveness (i.e., mutual accountability and results-based management) in Paris and Accra. The first phase, which will be analyzed in this paper, focused on Costa Rica and regional knowledge-sharing; its objective was to improve institutional conditions for evaluations of public policies and programs. It was the first Evaluation Capacity Development (ECD) program of its kind within German cooperation and had a total budget of €3.3 million over the three years.

GIZ’s management model Capacity WORKS provided a structured approach for the project’s implementation process. Capacity WORKS was used

- as a framework for joint reflection and discussion within the cooperation system;
- as an organizing set of principles that helped to define the shape of the program;
- as a management toolbox; and
- as an orientation for internal and external communication.

This paper aims at illustrating how a deliberate approach toward Capacity WORKS’ five “success factors” (strategy, cooperation, steering structure,

processes, learning and innovation) supported adaptive management and results-oriented implementation.

Contextual conditions

The Republic of Costa Rica is one of the most stable democracies of the Americas. The Constitution of the Republic of 1949 and its political implementation by the respective governments fostered the provision of broad-based access to education and healthcare, as well as a robust system of checks and balances. Political stability and the successful transformation from an agrarian-based economy into a service industry provided the basis for the necessary investments in basic social services, while a good endowment of human capital and natural amenities has further contributed to socio-economic progress.

In recent years, a number of socio-economic challenges, together with the fragmentation of the political system and the public administration, have been putting pressure on the country's governance mechanisms and have resulted in decreasing levels of trust in public institutions (OECD [Organisation for Economic Co-operation and Development], 2015). Costa Rica's public administration is characterized by an important number of subsidiary bodies of central government ministries and a large, institutionally decentralized sector (e.g., semi-autonomous and autonomous bodies, state- and non-state-owned enterprises). As a recent OECD governance review (OECD, 2015) points out, the country's public administration has limited steering capacity by the center of government and limited accountability mechanisms. The center of government is composed of the Ministry of the Presidency, the Ministry of National Planning and Economic Policy, and the Ministry of Finance. Persisting regional disparities and increasing levels of inequality among the population require reliable information on the effectiveness of public interventions for evidence-based policy-making.

MIDEPLAN's primary duties include the preparation of the National Development Plan; the verification that public investment projects across government entities are aligned with the priorities set forth in the National Development Plan; and the approval of investment projects of public agencies when such projects are externally financed or government approval is required. The ministry is also responsible for the setup and coordination of SINE. Through these functions, the ministry gives techni-

cal and political advice to the Presidency of the Republic and other public institutions while it formulates, coordinates, monitors, and evaluates the strategies and priorities of the government.

Development challenge

The National Monitoring and Evaluation System of Costa Rica and its corresponding laws were established during the 1990s. Since then, the country has endeavored to implement M&E activities as part of its public policy framework. Critics contend that no efforts were undertaken to strengthen SINE as an institutional support network for research and strategic analysis, and that no systematic evaluations of public programs were carried out. Thereby, the national M&E practice transcended into a bureaucratic follow-up and institutional self-reporting system for the government's principal political agendas. Thus, although it served basic accountability needs, it was hardly used for decision-making, learning, and strategic management or the design and development of new policies (Sanchez & Storm, 2016).

Implementation challenges and hypotheses

FOCEVAL's objective was to create improved institutional conditions for the evaluations of public interventions. This paper focuses on how the program developed, executed, and adapted a joint approach with its national counterparts to deal with the following implementation challenges.

1. The Ministry of Planning was considered to be weak and as having limited steering capacities as well as being rather isolated within Costa Rica's highly fragmented public administration. Inside the ministry there were conflicting views on the institutional mandate toward monitoring and evaluation, and there was generally little dialogue and coordination between the different units and departments. In the wider institutional context, evaluation was perceived as an instrument for scrutiny and control that manifested in resistance and fears of personal scapegoating. The demand from civil society and parliament for evaluation results as a basis for evidence-based policy-making was still low, and evaluation was perceived as an additional administrative cost rather than as an instrument enhancing transparency or public adminis-

tration performance. Planning and reporting mechanisms were focused on activities and lacked results-orientation. Institutional information politics were restrictive, but evaluations that are only accessed by policy-makers provide a very limited kind of transparency – inwards and upwards. In this context, the challenge to be tackled was: How to create an enabling environment for a reform process that would require changes in organizational behavior and mindsets from a multitude of actors?

2. Many consultants and public servants had already been trained in evaluation methods but had no practical evaluation experience, and the actual public-sector programs did not match the classroom conditions. Academic discourse on evaluation was detached from public-sector conditions. Program theories and detailed objectives often needed to be reconstructed *ex post*; in cases where the data from monitoring systems or other statistics was available, it was often not reliable, incomplete, or impossible to disaggregate, posing difficulties for effect and impact-oriented evaluations. Inter-institutional processes and procedures for strategic evaluations as well as their setup and utilization were not yet in place. Officials in MIDEPLAN had no experience with managing evaluations and guidelines, standards, as well as the processes and steering mechanisms that needed to be established. Hence, the second implementation challenge for the program was: How to design, select, and set up pilot evaluations that, for them, become catalysts for greater interest and acceptance of evaluations in the public administration and among civil society and parliament while matching international standards?

The program's hypotheses of how to best approach these challenges were as follows:

- a) Establish an integral approach to capacity development that addresses the individual, organizational, and political levels simultaneously by combining training and sensitization on an individual level with organizational development and changes on a policy level.
- b) In order to create an authoritative environment for reforms, the program needs to create alliances with – and gain broad support from – political decision-makers, in particular. Also, strategic alliances between academia and public-sector institutions are considered important to enhance the relevance and usefulness of national training offers

on the one hand, and the legitimacy of evaluation processes on the other hand.

- c) As a pilot program, with little available information on what had worked where and why elsewhere, the program’s success is dependent on tight feedback loops involving information and perceptions from a variety of stakeholders and allowing room for reflection, deviance, and adaptation. Hence, an incremental and results-based approach needed to be part of the design in order to achieving the program’s objective.

These hypotheses remained relevant throughout the implementation process, but the program had to adapt its specific approaches and strategic responses during the course of implementation. GIZ staff and program stakeholders discussed the logic and strategy of implementation (“how” are we going to work and implement the program and “why”) on a regular basis. GIZ’s management model Capacity WORKS – with its integral concept of capacity development and systemic orientation toward the success factors of strategy, cooperation, steering structure, processes, and learning and innovation – provided a framework for these reflections and enabled discussion and decision-making within the cooperation system. The main themes of the five success factors are (GIZ [Deutsche Gesellschaft für Internationale Zusammenarbeit], 2015)³:

Strategy: The cooperation system will succeed if and when the cooperation partners agree on a joint strategy to achieve the negotiated objectives. This requires an investment of time and resources in a communicative and social process of strategy formulation with partners.

Cooperation: Trust, the negotiation of appropriate forms of cooperation, and clearly defined roles form the basis for a good cooperation. Only when the participants in a cooperation system acknowledge that they are all dependent on each other and play their parts can the system be successful.

Steering structure: A development project has to take decisions daily about management and needs an adequate structure to do this. The cooperation systems is guided by agreements on how the actors involved will go about jointly preparing and taking the decisions that affect them.

3 Please see the paper from Neill Hatton for further details.

Processes: Successful cooperation systems include a clear understanding of effective ways of delivering outputs, for which new processes are established or existing processes modified. The management (change processes) of the project as well as the permanent processes in the sectors and countries where the sustainable change is to take place need to be analyzed, designed, managed, and improved.

Learning and innovation: The cooperation partners create an enabling environment for innovation by boosting the learning capacities of the actors involved. The systemic concept not only focuses on learning needs on an individual level but also interdependencies and learning needs on an organizational and policy level as well.

Tracing the implementation process

This section aims at illustrating how decision-making and adaptive management took place and which lessons were learned with regard to the program's hypotheses and implementation challenges. It does this with reference to the three hypotheses identified above as being particularly relevant: (a) integral capacity-development strategy, (b) promoting coalitions and alliances, and (c) feedback loops and adaptive management. Although the mentioned success factors – as they are associated and interconnected – are all relevant throughout the analysis, the first part has a particular focus on strategy and processes, the second on cooperation and steering structure, and the third on learning and innovation as well as cooperation.

Strategy development

Implementation of FOCEVAL began in July 2011. One of the first activities GIZ and MIDEPLAN agreed upon was to set up a strategy process that included a series of short workshops with stakeholders from different institutions and sectors. The purpose of this process was to:

1. increase MIDEPLAN's visibility as a lead agency for the national evaluation system;
2. enhance understanding and information about barriers and existing good practices with regard to evaluation in the public sector;

3. shape the capacity-development strategy and the operational setup of the program;
4. promote the project and identify cooperation allies.

A core team was identified that was comprised of staff from GIZ and MIDEPLAN, which organized the strategy process. The process was conceived as a joint learning process, during which each of the team members assumed responsibility for specific tasks and topics. Impressions and hypotheses derived from the events were discussed within the team, and each member carried a learning diary. During this process, it became obvious that the project needed to address not only the supply side (technical capacities) – as it was established in the project offer (training, pilot evaluations, M&E systems) – but also the demand side (support and demand for evaluation from policy-makers and civil society representatives) in order to achieve its set objective, which was to create improved institutional conditions for evaluations of public interventions. The dimension of responsiveness toward evaluation and evaluation results by stakeholders in order to use the instrument of evaluation to facilitate change was underestimated by the appraisal team. However, the leeway given a GIZ program to adapt to changing circumstances during implementation allowed for the inclusion of this dimension in the operationalization of the program. The program was promoting a systemic approach toward capacity development that considered interventions at the individual, organizational, and societal levels. The table below illustrates how such a conceptual understanding of a systemic approach to capacity development is operationalized and leads to the identification of core activities for the program.

Table 1: FOCEVAL: Generic capacity-development strategy

| ECD | Supply | Demand |
|---------------------|---|--|
| Individual | Training of evaluators Training the trainer – formats Trainings on evaluation management for commissioning organizations Learning by doing formats | Sensitization of political decision-makers Advocacy directed at members of parliament and representatives of civil society organizations |
| Organization | Conduct pilot evaluations Harmonization of concepts and terminology (manuals, seminars) Adapt training offers to local demand and context and anchor them at local training providers | Establish incentives to develop and use M&E systems (e.g., quality award) Promote dissemination and accessibility of evaluation results and dialogue between civil society and the public sector on M&E results (e.g., monitoring reports: National Development Plan) |

| | | |
|---|---|--|
| <p>Cooperation and networks</p> | <p>Promote strategic alliances between public administration and academia Promote professional exchanges among networks of evaluators and national and regional professional associations (Community of Practices) Promote and support of inter-institutional evaluation projects</p> | <p>Foster civil society networks / NGOs working on transparency and accountability Promote knowledge-sharing with other national M&E authorities (Colombia, Mexico, El Salvador, Ecuador)</p> |
| <p>Political and societal conditions</p> | <p>Foster the evaluation component within the national M&E framework Foster the role of the executive M&E authority (MIDEPLAN): coordination mechanisms/steering groups, standards</p> | <p>Foster principles and mechanisms of accountability that are based on evidence and evaluation results Ensure institutional budgets for evaluations</p> |

Source: FOCEVAL/GIZ (Strengthening Evaluation Capacities in Central America/ Deutsche Gesellschaft für Internationale Zusammenarbeit, 2012a)

The joint strategy process fostered team spirit, trust, and communication within the core cooperation team. As a result of it, as many observers from other organizations commented, MIDEPLAN was viewed as taking a very different role than usual: asking questions, facilitating discussions, and listening to concerns. Overall, the strategy process was crucial for developing a joint understanding of context and specific challenges, promoting a systemic understanding of the implementation challenges, and establishing the foundations for the program’s cooperation strategy. It also strengthened the role of the leading agency, MIDEPLAN, vis-à-vis other agencies in the public administration, and thus raised the likelihood that MIDEPLAN could deliver the task it was mandated with. The process had fulfilled its purpose as an initial strategic loop within the implementation process.

However, though most of the identified intervention areas and change projects remained relevant throughout the project, assumptions and specific approaches had to be tested and developed while moving along with the implementation process. Although the strategy process helped to shape an understanding of the key processes that the program had to address within the institutional setup as well as gain clarity on the process map, the key change, strategy, and support processes within the program sequencing of – and entry points for – change initiatives could not be fully understood through analysis and discussion alone. Through learning by doing and joint reflection within the cooperation system, the program gained an increased understanding of the underlying barriers and stabilizing factors for existing processes and could allocate its resources more

efficiently. Shifting forces, interests, and coalitions in the wider institutional context needed to be assessed and monitored on a regular basis. Retrospectively, the initial change model and operation plan of the project were far too complex and overloaded with measures and indicators. A more iterative, flexible “learn as you go” approach while testing the temperature and staying open to strategic changes and alternatives would have been more appropriate. Also, it took a lot more time “to work the territory” than initially planned before concrete, feasible evaluation projects could be identified and initiated. Relationships and trust needed to be established, institutional contexts assessed, and political and technical support ensured. One of the initially foreseen evaluation projects had to be cancelled in the end because these factors had not been taken into account properly.

Building alliances

The operational plan and objectives for the first year contained numerous initiatives and change projects. The elaboration and publication of a manual for strategic evaluations by MIDEPLAN had been an important milestone. Now it had to be put into practice and prove its usefulness in real evaluation processes, but this process moved slowly. Health had been selected as a sector for an evaluation pilot. MIDEPLAN had received a formal request from the Costa Rican Department for Social Security to support the enhancement of the institutional M&E system. The background for this request was a report from the Supreme Audit Institution. However, discussions about an evaluation of primary healthcare were conflictive and, after a series of meetings and several months, it was still not possible to agree on the scope and purpose of evaluation. Other activities/change projects were still more on a level of general proposals and required a considerable amount of clarification, discussion, and negotiation on how to approach them. After a year of back and forth without breakthrough, frustrations and worries among the program team increased concerning the achievement of visible progress, particularly with regard to concrete evaluation projects.

These worries were reiterated in a workshop with the wider group of stakeholders, who had been invited to serve as a sounding board to the project: The lack of national evaluation cases that could be used to promote evaluations and serve as good examples was seen as a big problem

by all stakeholders. Promotional activities remained rather theoretical without national “show cases” that could demonstrate the benefits and challenges of evaluation.

The project team discussed the implementation strategy by making use of the five Capacity WORKS success factors and asked for feedback. The main recommendation that emerged from the workshop was: Create more spaces for inter-institutional discussion and coordination; allow more actors to take an active role in the program, not only as beneficiaries of program activities but also as active change agents; and improve communication on program activities and their intended results.

The GIZ team discussed these recommendations and their impressions and hypotheses internally and with MIDEPLAN. Supported by the feedback from the stakeholders’ meeting and seeing the need for more inclusive inter-institutional coordination, MIDEPLAN agreed to the revision and adaptation of the program’s steering structure.

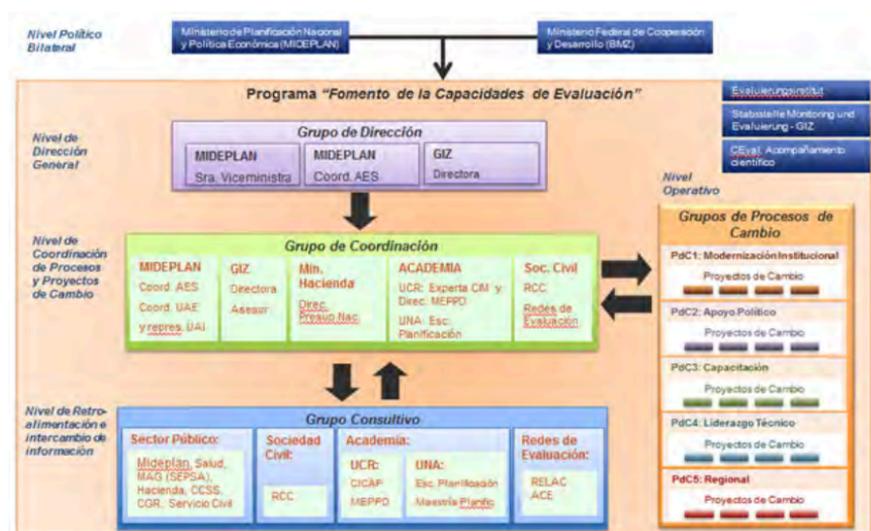
There were concerns that flexibility and promptness of decision-making could be lost if more voices were to be heard and included. Final decision-making should remain within MIDEPLAN and GIZ, whereas the steering structure should serve as a platform for strategic consultation and mutual information, and an advisory board for improved coordination and implementation. The coordination group included representatives from public administration, academia, and civil society (see below for details). Most of the participants were familiar with the program and had been involved in specific program activities. Monthly meetings were set up and hosted by GIZ and MIDEPLAN in the FOCEVAL facilities. The coordination group became an important motor for innovation and joint activities within and outside the scope of the program. Seminars and presentations were promoted and organized together. Representatives from MIDEPLAN and the Ministry of Finance engaged actively in the preparation and development of new training offers at the Training Center for Public Administration. The discussions facilitated a common understanding and language with regard to evaluation. Naturally, every organization also brought its particular interests to the table and looked for opportunities to benefit from the program’s resources. However, through joint discussion, these became more transparent, resources could be pooled, and synergies (e.g., joint trainings, disposition of rooms, equipment, facilitators) could be identified more easily. In other contexts, a lean steering structure might have been the most efficient one: In this context, the coordination group was not only a means to an end for improved implementation but also an intervention

that addressed deficits of inter-institutional communication and lack of coordination. It facilitated a common understanding of the challenges and objectives and strategic alliances with regard to a new and controversial issue.

One of the first joint activities became the promotion of an “evaluation challenge” and the subsequent selection of pilot evaluations.

The following chart of the FOCEVAL steering structure illustrates the multiple purposes the steering structure had, next to decision-making: The steering structure became an important instrument for building understanding of the issues among a wider group of stakeholders and forged alliances among stakeholder groups via its focus on consultations.

Figure 2: FOCEVAL – steering structure



Source: FOCEVAL/GIZ (2012b)

Learning by doing, learning from failure, and learning from data

A tracer study conducted by the program reviewing the professional development of graduates from postgraduate training courses in evaluation and the University of Costa Rica’s master’s program on the evaluation of projects and programs confirmed the existence of a significant gap between “theory and practice.” Hardly any graduate was able to apply his

or her acquired evaluation knowledge in the job they obtained after their degree. There was no demand nor opportunity to apply these skills. The organizational conditions were not yet in place. Though most graduates of specific evaluation courses confirmed the usefulness of the content learned, overall, the longtime capacity-building activities had not manifested in institutional M&E practices. No systematic evaluations that had been executed or commissioned by public institutions could be identified or were publicly available. In addition, many graduates also expressed the need for further training and did not feel prepared to design and conduct evaluations of public policies, plans, or programs.

One of the program's conclusions in light of this situation was that the program needed to promote institutionally embedded evaluation pilots that would allow for learning by doing under "real life" conditions, and thereby enhance individual and organizational evaluation capacities. The health pilot had not been a good choice in that sense. The topic of primary healthcare was extremely complex, with lots of controversies, conflicting views of a multitude of actors, and high levels of political and public attention. The Costa Rican Department of Social Security was already under high amounts of public scrutiny at the time the evaluation was discussed, and it had been shaken by several institutional affairs. It can be assumed that the evaluation was perceived as an additional threat. Although considerable resources (predominantly in terms of time and energy) had been invested in getting the health-sector evaluation to the starting point as a pilot, the activity was finally cancelled by the Minister of Planning. This decision was based on recommendations from the working group on healthcare evaluation and FOCEVAL's steering committee. It was followed by an exchange of tense institutional notes. Frustrations among all involved actors had accumulated during the tenacious negotiation process. On the other hand, it allowed for important learnings. The program team discussed the lessons learned and how they should be considered in the following evaluation pilots: 1) The evaluation should be voluntary and not imposed externally. Participating institutions should have the opportunity to present themselves as "early movers" and models, with an interest in learning and transparency. 2) A clear commitment from technical staff and leadership was needed. Institutional representatives should be involved in the whole evaluation process.

FOCEVAL decided to organize an evaluation challenge. Organizations were invited to compete to "win" an evaluation by presenting proposals for evaluations of public programs. Criteria were established beforehand

and included: relevance, evaluability, as well as political and technical support. The selection process gave MIDEPLAN and GIZ the time and opportunities to assess the viability and context of the evaluation proposals – which had been a significant deficit in preparation of the healthcare evaluation.

The challenge was promoted through the steering groups and the networks the project had developed during the first year of implementation. It led to about eight proposals from different ministries and agencies. The proposals were assessed by a team from MIDEPLAN and GIZ. Two proposals from two different sectors seemed eligible and were selected. Both were handed in from former participants of FOCEVAL trainings: one from the Institute of Social Support, one from the Ministry of Education. The selected proposals were presented in a press conference with participation of the institutions leadership and the Minister of Planning. Formal inter-institutional agreements were signed, making detailed notes of roles, contributions, and responsibilities of each organization and the commitment to publish and discuss evaluation results.

Now that pilot evaluations had been identified, the according processes needed to be designed in a way that would generate individual and institutional learning experiences and allow for practicing roles and processes. FOCEVAL proposed to follow a highly participatory learning-by-doing approach: Each evaluation team consisted of functionaries from the Ministry of National Planning, GIZ advisors, and the institution responsible for the program that was evaluated. Technical quality was ensured via external Costa Rican evaluation consultants and backstopping from the Centre of Evaluation from Saarland University in Saarbrücken, Germany. Representatives from the involved organizations discussed and developed the evaluation design jointly with the evaluation team and were informed regularly on evaluation progress which ensured transparency and continuous feedback during the whole process. Conflicts and disagreements were treated in a steering committee that had been set up for each evaluation. Though MIDEPLAN was not supposed to execute evaluations itself, it was considered important that the staff, who were supposed to coordinate, commission, and ensure the quality of evaluations in the future, be involved to gain detailed insights into the whole evaluation process. Thus, officials from MIDEPLAN formed part of the evaluation teams throughout all stages of the evaluation process, including data collection and analysis.

By the end of 2012, several pilot evaluations were on their way: MIDEPLAN had decided to lead by example and was preparing a strategic evaluation of the methodological setup and compilation process of the National Development Plan. A team from MIDEPLAN, GIZ, and the Ministry of Agriculture was working on a feasibility study for an impact evaluation of a large irrigation project. People from different areas of MIDEPLAN had begun to discuss and align their concepts and guidelines for planning, monitoring, and evaluation, and they used the experiences from the pilot cases to adapt its manual for strategic evaluations.

There was one last area of concern the program had to address at this point. Though the GIZ team had established a close cooperation with its main counterpart, MIDEPLAN, and engaged in regular conversations and joint reflections, tensions in the cooperation became apparent: Feedback from interviews that had been conducted after the first year of implementation revealed that misunderstandings and frustrations regarding roles and cooperation had arisen within the ministry. In a two-day workshop organized and facilitated by the GIZ team, GIZ and MIDEPLAN discussed these findings and their mutual impressions. In the first part of the workshop, the group reflected on what they had achieved so far and what had worked and why. In a second part, the teams from MIDEPLAN and GIZ made their difficulties with cooperation visible in a drawing: MIDEPLAN chose to draw a football stadium with an ongoing game as a symbol for the cooperation. On the field were players from MIDEPLAN and GIZ – however, it was not always clear if they were playing on the same team or competing against each other, nor what the exact rules of the game were. Meanwhile, there were a lot of spectators at the margins – watching, cheering, and relaxing, and occasionally throwing new balls onto the playing field – representing other institutions/stakeholders of the program, but also staff from MIDEPLAN's other areas.

The teams realized that they were in danger of reproducing the patterns and becoming part of the problems they were trying to solve: competition, miscommunication, and lack of coordination. They decided on a series of measures. These included explicit agreements on roles and rules of the game. For example, in all project activities or change projects, one person was appointed as the official coordinator or focal point, who would serve as the entry point for requests from other institutions and shared all information within the team. The team also organized Q&A sessions within the ministry and improved communication with – and the involvement of – other areas from MIDEPLAN.

As a program that was working on accountability, transparency, and organizational learning, the program itself tried, and needed, to be a model in how it gathered, processed, and made information and data available. The project paid much attention to the setup of its own monitoring system, which was supposed to serve as a good practice in itself. The program's M&E system consisted of various elements:

- process- and results-indicators assessed by different methods such as focal groups, surveys, document analysis, etc.;
- a comprehensive study that consisted of a series of in-depth interviews with relevant stakeholders, external observers, and experts. Interviews were conducted by an external independent consultant and designed as a panel. The interviews focused on implementation progress/results but also on the project setup and internal functioning. They included an analysis and visualization of actors and relationships within the wider institutional system;
- regular assessment and evaluation of Capacity WORKS “success factors”: Where do we stand with regard to the five success factors? Which success factors require our attention? Which modifications are necessary in order to achieve our objectives/results more effectively?

Feedback that was obtained by different means and from different sources was shared and discussed within the steering group. Lessons learned in all components were discussed and documented on a regular basis. An important source of feedback was the panel study, which included not only interviews and perceptions from stakeholders and beneficiaries but also external observers and independent experts as well. Critical observers, who were not directly involved in project implementation and had no immediate stake in particular activities, were an important source of information to assess blind spots and biases of the project team. Regular joint reflection within the cooperation system – combining evidence from external data with internal perceptions, lessons learned, and hypotheses – were important loops for shaping the implementation process.

By the beginning of 2014, all four pilot evaluations had been finalized and results were discussed and published. Institutional management responses and plans were developed. MIDEPLAN set up a national evaluation agenda for the upcoming years.

The coordination group organized an international knowledge-sharing event in March 2014. All actors of the coordination group assumed a very active role during the preparation and execution of the conference. During

this event, the Minister of Planning officially submitted the reports of the country's first strategic evaluations to the corresponding authorities, and MIDEPLAN presented its achievements and learnings together with a group of committed allies.

An evaluation of FOCEVAL, conducted in September 2013, came to the following assessments (GIZ, 2013): In comparison with other comparable technical cooperation programs with similar funds, the program achieved to position the topic in the sector in a relatively short time; the good cooperation with the Ministry of National Planning and the commitment of other stakeholders affected the cost-effectiveness relationship positively. The connection of the program with its political counterpart was exemplary. The Ministry of National Planning has actively pursued the institutionalization of strategic evaluations.

Conclusions

Was FOCEVAL's implementation process smart? In retrospect, many of the program's decisions were not smart. The decision to engage in the health sector evaluation was not smart. However, there appeared to be good reasons for it at the time. In other countries (e.g., Mexico), large-scale, strategic evaluations that received broad public attention had served as icebreakers for developing national evaluation capacities. Nevertheless, the evaluation case, its implication, and its context needed to be assessed and understood, and the decision was made – and announced – too early (due to political pressures from different sides). Strategic evaluation was introduced as a new process to the public sector; roles and procedures – though existing on paper – needed to be practiced. Practice requires room for detours, delays, and possible failure. The pressure that built up with regard to the first evaluation was counterproductive and tied up too much of the program's attention for too long.

The strategy process can be considered as being somehow smart – it served different purposes, increased ownership among counterparts, and helped the program to establish the foundations and contacts for program implementation – but the outcome (operational plan and results model) was overambitious. The program team and their counterparts would have been wise to listen to more critical voices initially. However, as it can happen when a new project is set up (and especially when most of the people involved have a strong planning and/or evaluation background), the team

fell into a planning illusion and believed it had to identify only the right paths and patterns. From today's perspective, the advice of the program team would be to start small, think big, and grow as you go. The initial illusion was that the program could tackle all relevant processes at the same time and began with the most audacious challenges (healthcare evaluation).

Probably the smartest thing about FOCEVAL was that the project was able to attract and include an increasingly large number of diverse voices over time and to build a platform for collective learning and impact. This did not happen by chance – it was supported by structured reflections and also by making use of tools from Capacity WORKS. The project identified and approached cooperation partners deliberately and opened options for their participation and involvement (e.g., through the steering structure). Spaces and processes for reflection, learning, and cooperation were created. However, a high level of intensity of cooperation and a multitude of actors often imply high levels of conflict as well. A cooperation system has an underlying set of different organizational logics, interests, and cultures at its base, with different needs and requirements for more or less formalization, communication, or information. This has been a recurring issue in the program. The lesson learned is possibly that there is no optimal outcome that will satisfy all but rather only an approximate good compromise. In the course of the program, it was important to reflect on when and where it was adequate and eligible to change the visible and invisible rules of the game, and where it was necessary to enforce and not inadvertently undermine them, for example by establishing temporary, parallel structures and processes. For example, it was considered crucial that MIDEPLAN was – and was perceived as – the owner of the pilot evaluations and set the rules for them.

There is no one-size-fits-all approach for a cooperation system and steering structure. Different means and options – and their advantages and disadvantages – should be considered. In the case of FOCEVAL, the coordination group was not only a means to an end for improved implementation but also an intervention that addressed deficits of inter-institutional communication and the lack of coordination. Also, there were no other relevant donor activities in the sector; therefore, no other coordination and transaction costs were imposed on program staff and counterparts. The discussions and joint activities within the coordination group improved the common understanding of challenges and objectives as well as strategic alliances with regard to a new and controversial issue. In this context, it

was a crucial catalyst for the other success factors of learning and innovation, strategy development, and cooperation.

Finally, did Capacity WORKS make FOCEVAL's implementation process smart(er)? Capacity WORKS was an integral and important part of the program's management approach. It provided orientation and helped to facilitate a joint learning process within the cooperation system, which enhanced trust and enabled critical discussions over time. It was furthermore helpful to establish a common language and approach in order to discuss implementation challenges within the cooperation system. This required resources (time and space) and commitment, as well as a high level of interaction. It was an enabling factor that there were people among the program's main counterparts who felt immediately attracted to the systemic ideas underlying the model and who assumed visible leadership and supported its joint application throughout the implementation process. In the end, any management model can only be as smart as the people working with it.

References

- Gallardo, R. (2013, July 20). Columna de opinión. *El financiero*.
- FOCEVAL/GIZ (Strengthening Evaluation Capacities in Central America/Deutsche Gesellschaft für Internationale Zusammenarbeit). (2012a). *ECD strategy*. Internal documents. Costa Rica: Author.
- FOCEVAL/GIZ. (2012b). *Steering structure*. Internal documents/workshop documentation. Costa Rica: Author.
- GIZ (Deutsche Gesellschaft für Internationale Zusammenarbeit). (2013). *Informe de evaluación. Programa FOCEVAL*. Internal documents.
- GIZ. (2015). *Cooperation management for practitioners. Managing social change with Capacity WORKS*. Wiesbaden: Springer Gabler.
- OECD (Organisation for Economic Co-operation and Development). (2015). *Costa Rica. Good governance, from process to results*. Retrieved July 29, 2016, from <http://www.oecd.org/countries/costarica/costa-rica-good-governance-from-process-to-results-9789264246997-en.htm>
- Sanchez, S., & Storm, S. (2016). Political use of evaluation results in Central America. In R. Stockmann & W. Meyer (Eds.), *The future of evaluation. Global trends, new challenges, shared perspectives* (pp. 204–213). London: Palgrave Macmillan.