Institutional and governance architecture as enablers of an evaluation culture: The experience of IFAD

This article discusses the building of an organizational evaluation culture and how independent evaluation can contribute to it, using the case of IFAD as an example. Creating an evaluation culture is distinct from introducing, enforcing and complying with norms, bylaws and processes. It goes beyond the formalities and procedures that may exist in an organization. The article introduces conceptual clarity on organizational culture, organizational evaluation culture, the key conditions and enablers, and how independent evaluation can support their development, before setting out the experience of IFAD.

Oscar A. Garcia, Director, and Fabrizio Felloni, Deputy Director, Independent Office of Evaluation, International Fund for Agricultural Development (IFAD)[1]

Key Messages

■ An organizational evaluation culture is one that demands and uses empirical information to assess performance and outcomes and orient managerial decisions.

■ In addition to the availability of data, there needs to be capacity to analyse and interpret them, deliberate interest in findings and a feedback loop leading to action and changes.

■ The presence of an independent evaluation function may not generate, on its own, an evaluation culture but represents an important building block.

■ Compliance with the evaluation procedures is not equivalent to evaluation culture. It is necessary to move beyond formalism and there is no substitute to the commitment of leadership (Management and Governing Bodies).

Clarifying the concepts

An **organizational culture** can be understood as a set of common values, beliefs, goals and business processes, irrespective of their level of formality (Ravasi and Schultz, 2006). Such culture permeates the strategies, style and modus operandi of an organization.

An **organizational *evaluation culture*** can be defined (Mayne, 2010) as a culture which appreciates and deliberately demands data and empirical information on its performance and outcomes and uses findings from such evidence in order to improve its performance and the achievement of outcomes. Mayne identifies some of the characteristics of an organization that possesses a strong evaluation culture: (i) engage in self-reflection and self-examination (i.e., self-evaluation); (ii) deliberately seek evidence on what it is achieving; (iii) use results information to challenge and support what it is doing; (iv) value candour, challenge and genuine dialogue; (v) engage in evidence-based learning and make time to learn; (vii) encourage experimentation and change.[2] Such evaluation culture clearly presents advantages in terms of capacity and timeliness to correct the course of action, adaptation to change and adherence to corporate mandate.

A consolidated practice of data gathering can help but, on its own, does not necessarily drive an evaluation culture. An important corollary of the above definitions is that the availability of empirical data and the capacity to analyse them is not a sufficient condition to establish an evaluation culture. Even when data are available, and if data quality or representativeness is not an issue (a heroic assumption in most cases), an evaluative culture would need:

■ capacity to analyse and interpret data, including at the higher decision-making levels;

■ deliberate interest in findings from data analysis to undertake an objective review of the operational performance and results, connected with some reward system to create and maintain incentives;

■ a feedback loop that links findings to action to modify current and future operations and business modalities. ▶▶

▶ In turn, this requires a strong drive from the organizational leadership, specific organizational arrangements and structural support, and a learning focus. As an example, the systematic conduct of impact evaluations, using rigorous methodologies for controlling sample bias, can set the premises to make decisions based on sound evidence. However, these evaluations would have to be intentionally utilised to support managerial decisions. As noted in a World Bank evaluation (IEG 2012), the results of impact evaluations were not used systematically as a basis for decision making, due to the absence of a clear feedback loop and system of incentives. A more recent evaluation by the Inter-American Development Bank (OVE, 2017) found that some impact evaluations were instrumental for the Bank to avoid ineffective projects, although the overall findings on the use of impact evaluations were uneven.

Evaluation in international organizations

In the experience of many international organizations, the presence of a dedicated and independent evaluation function was one of the building blocks to introducing or promoting an evaluation culture. As noted, leadership plays a key role in an evaluation culture. In larger organizations, leadership commitment to evaluation may be prone to the principal-agent problem (*Laffont and Martimort, 2002*)[3] with the related issues of moral hazard and conflict of interest. For this reason, one of the gold standards of evaluation practice [OECD/ DAC, Evaluation Cooperation Group and UN Evaluation Group][4] is that the responsibility for the evaluation of projects, programmes and policies should not be in the hands of persons or organizational units that were previously responsible for their design and implementation.

In many international financial institutions such as the multilateral development banks, the independent evaluation function is placed under the direct reporting line of the Executive Board. The Executive Board represents (a sub-set of) the membership of the organization (the principal). The fact that the evaluation unit does not functionally report to the management of the organization (the agent) addresses problems of moral hazard and conflict of interest.
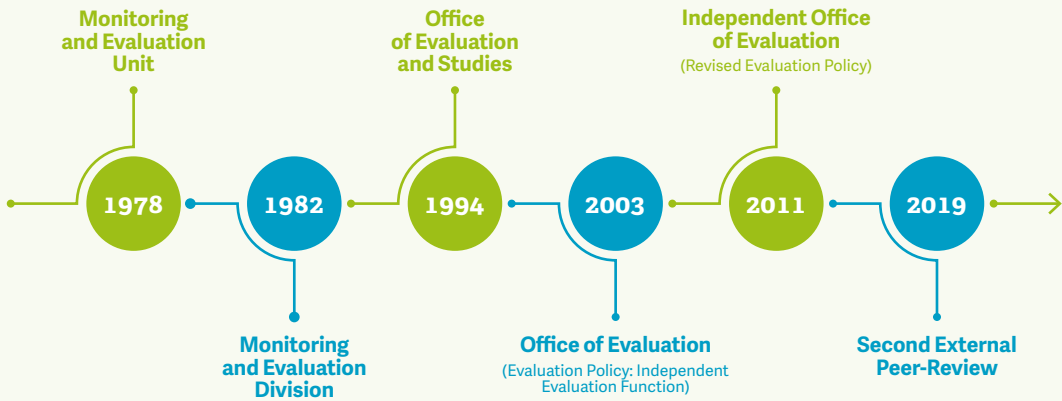
However, an important prerequisite for the system to function well is that the Executive Board (principal) continues to value and demand empirical evidence and independent analysis, and keeps management (agent) accountable for results and for implementing recommendations.

IFAD's experience

Historically, this has been the trajectory followed by the Independent Office of Evaluation (IOE) of the International Fund for Agricultural Development (Figure 1). The office started as a monitoring and evaluation unit (i.e. self-evaluation) in the Department managing development projects, and was then turned into a semi-independent unit, reporting only to the President (chief executive) of the organization. It became fully independent in 2003, under the direct reporting line to the Executive Board.

While independence can help deal with the principal-agent problem, it cannot be solely relied upon to build an evaluation culture in the whole organization. As noted, there needs to be a clear drive from the organizational leadership (both the principal and the agent) and an intentional use of data and information to assess, learn and re-orient programmes, strategies and internal processes. If the mandate to assess organizational activities is confined within an independent unit, there is no guarantee that it will create strong incentives across the organization to seek information on programme performance, set time to learn and innovate. Two practical constraints relate to the following: ▶▶

Figure 1: The trajectory of the evaluation unit at IFAD towards independence

Monitoring and Evaluation Unit — 1978

Monitoring and Evaluation Division — 1982

Office of Evaluation and Studies — 1994

Office of Evaluation (Evaluation Policy: Independent Evaluation Function) — 2003

Independent Office of Evaluation (Revised Evaluation Policy) — 2011

Second External Peer-Review — 2019

Source: Elaboration by the authors

▶ ▪ Flow and reach of information: in a large organization, the availability of an evaluation report does not automatically translate into shared knowledge on its findings and recommendations in all organizational units.

▪ Timing and timeliness of feedback: Depending on the stage of a programme or strategy within its lifecycle, the type of information and the level of analysis required vary (e.g., on outputs, intermediate outcomes, longer-term impacts). Most independent evaluations take place towards project end-line and tend to be summative (i.e. assessing the results), while implementers at an earlier stage need formative feedback (i.e., reviewing the implementation progress and process and its consistency with design).[5]

The importance and the role of the self-evaluation function. For this reason, many international development organizations distinguish between an independent and a self-evaluation evaluation function. While the independent evaluation function is placed under the supervision of the Executive Board, the self-evaluation function is a task of the operational units under the ultimate responsibility of the senior management.

Self-evaluation is a very important building block for developing an evaluation culture in an organization. A well-performing evaluation system is based on the premise that assessing performance and results is a duty of managers at all levels of the organization. Self-evaluation can also be an important building block for independent evaluation by providing useful secondary data for independent review and validation and triangulation with additional sources. The cost of generating, ex post, basic information on the implementation performance of a project or programme would be very high. In the words of a former Director General of the Independent Evaluation Group of the World Bank, "self-evaluation is to independent evaluation the equivalent of what accounting is for audit"[6].

Independent evaluation can play a key role in raising the standards of self-evaluation. Evaluating a project or programme is not an activity that is naturally familiar to everyone. It requires a certain level of skills, competencies, subject matter technical knowledge, intellectual rigour and mind-set for enquiries. Units that are primarily in charge of designing and supporting the implementation of programmes may not have the expertise, time and focus that are required to conduct high quality self-assessments. ▶▶

▶ In many international financial institutions, one of the functions of self-evaluation is to assess and report on completed projects. From its side, the independent evaluation function is typically responsible for: (i) helping management establish standards for self-evaluation and build capacity; (ii) assessing the quality of self-evaluation products; (iii) conducting independent validation work at the project level, while devoting most of its resources to higher-level and more strategic evaluation products; and (iv) periodically carrying out an independent assessment of self-evaluation systems in the organization, which can help assess the strength of the evaluation culture.
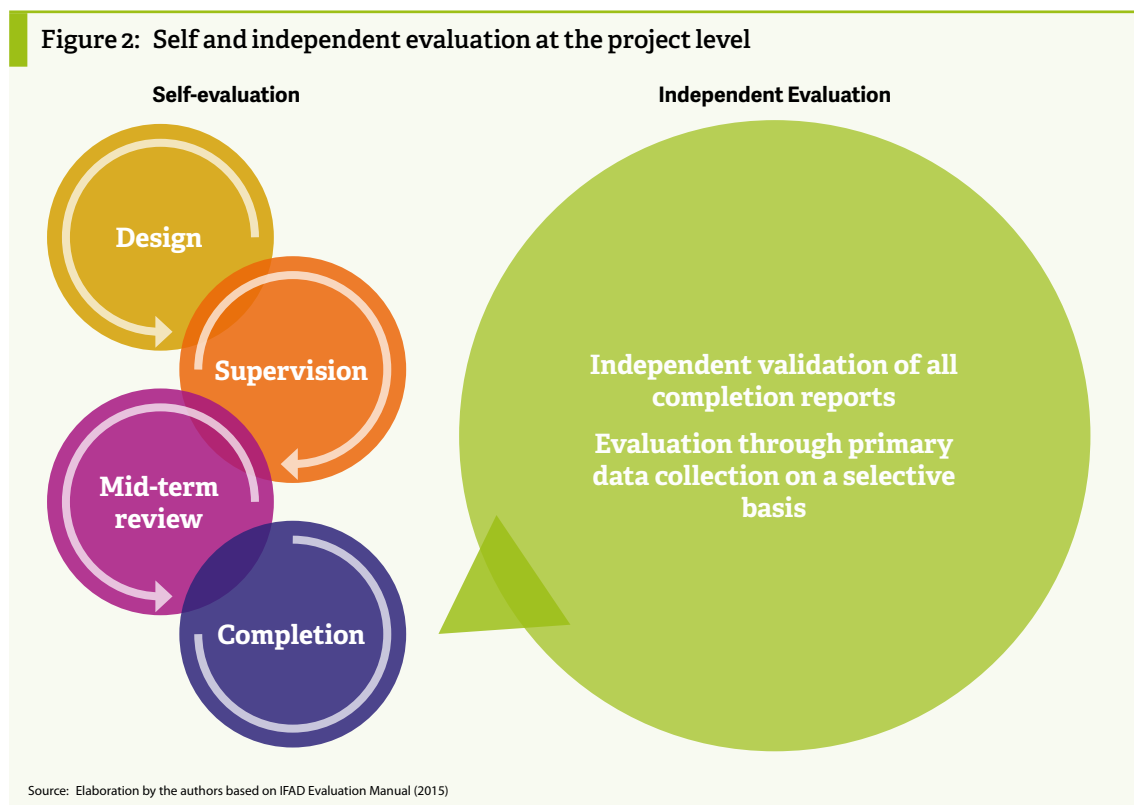
IOE at IFAD initially devoted most of its effort to evaluating projects. Self-evaluation had a limited role within the organization. The fundamental assumption was that project management teams, under the aegis of borrowing country governments, would establish their own monitoring and evaluation units and provide reporting during implementation and at completion point. At the corporate level, limited priority was given by management to assessing the quality of work of project M&E units and proactively reviewing progress made by projects.

From 2006, the first steps were made between IOE and IFAD management to adopt a common set of criteria for evaluating projects, based on similar definitions, as well as a common ratings system, based on a six-point symmetric scoring (harmonization agreement)[7]. This assessment system was based on the evaluation methodology of the independent evaluation function, which was progressively adopted and internalized by the management for its self-evaluation and became part of its business processes.

From 2010, after a peer review of IFAD's evaluation function, the Executive Board decided that IOE would conduct a validation of all project completion reports (prepared by the governments with IFAD management support), based on desk review of the documentation available (Figure 2). On a more selective basis, ▶▶

Figure 2: Self and independent evaluation at the project level

Self-evaluation

Independent Evaluation

Design

Supervision

Mid-term review

Completion

Independent validation of all completion reports

Evaluation through primary data collection on a selective basis

Source: Elaboration by the authors based on IFAD Evaluation Manual (2015)

▶ IOE would evaluate a smaller number of projects based on country visits. The three-fold purpose of this decision was to:

■ provide the evaluation function with an overview of the quality of self-evaluation activities;

■ through systematic scrutiny, generate incentives for Management to improve the quality and credibility of self-evaluation activities; and

■ obtain some form of independent assessment on all operations that are completed.

From individual self- and independent evaluation products to consolidated analysis. To provide a more consolidated review of findings, both the self- and the independent evaluation function prepare an annual synthesis report of their respective findings (Figure 3). At IFAD, these are known, respectively, as the Report on IFAD Development Effectiveness or RIDE (self-evaluation reporting) and the Annual Report on Results and Impact of IFAD Operations or ARRI (consolidating findings of independent evaluations). These are both presented to the Executive Board of IFAD and offer the two perspectives on the implementation performance and outcomes of the operations.

The preparation of two reports from the self- and the independent evaluation function is not a duplication but a reinforcement of the necessary interaction (feedback loops) and dialectics between the two functions. While both report at project level, where there is better comparability (thanks to a harmonization agreement on the methodology), the RIDE also reports on other levels of results, which are periodically agreed upon with IFAD member countries. On the other hand, the ARRI also reports on findings from higher-level evaluations (such as country level evaluations, thematic evaluations, corporate-level evaluations and evaluation syntheses).

It is inevitable that self- and independent evaluation will offer different perspectives and analytical angles. The comparison between the two, while potentially contentious, is also a source of knowledge. ▶▶

Figure 3: Consolidating self and independent evaluation findings
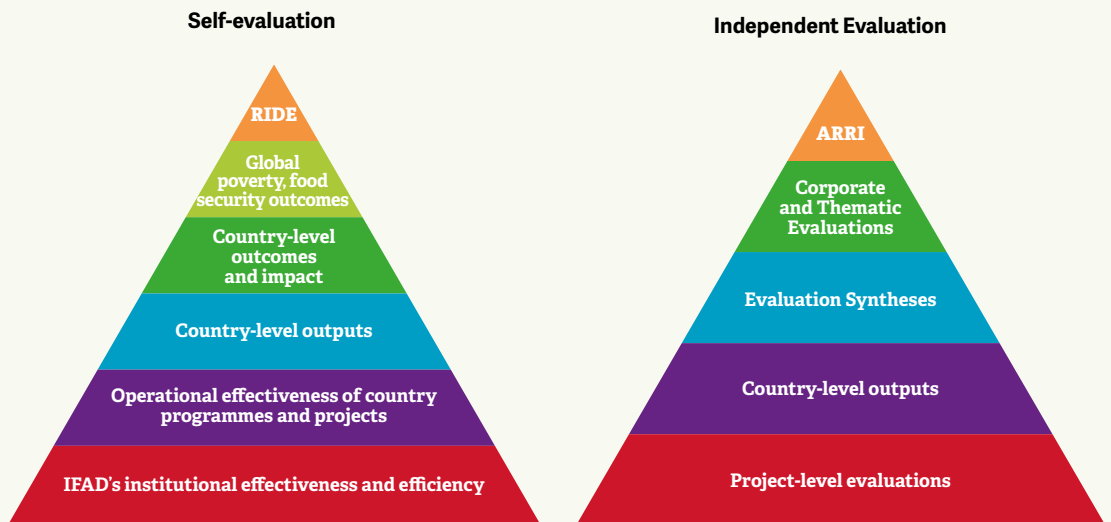
Source: Elaboration by the authors based on IFAD Evaluation Manual (2015)

Figure 4: Closing the evaluation learning loop at IFAD
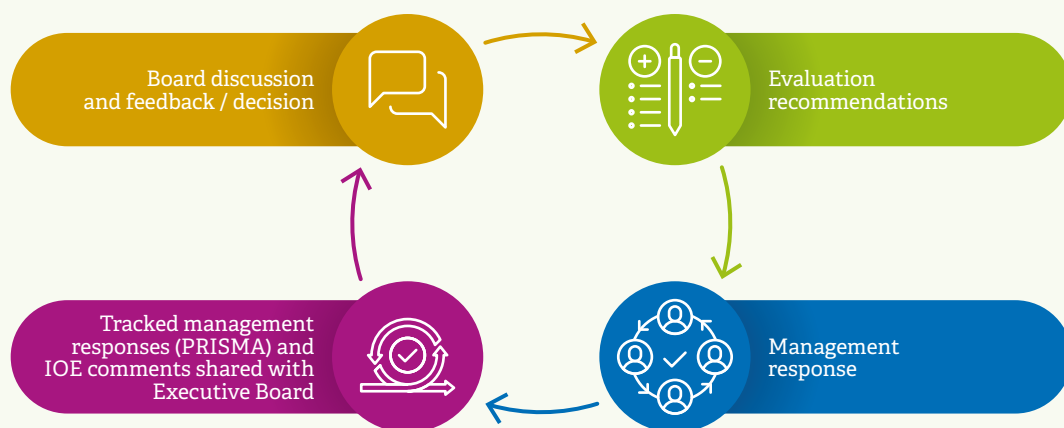
Source: Elaboration by the authors

▶ For instance, self-evaluation is more likely to take into account the constraints encountered during project design and the challenges related to the political environment in which projects operate. On the other hand, independent evaluations are typically more straightforward in assessing the level of achievement of objectives and results and in questioning the assumptions of the given projects.

Tracking responses to evaluations and their recommendation helps signal the urgency of implementing recommendations. An evaluation culture ultimately aims at producing evidence on performance and results in order to drive improvements and change. The conduct of an evaluation, even if of high quality, does not guarantee that recommendation will be heeded. A system that helps track recommendations and transparently records their follow-up can help enhance commitment to recommendation follow-up.

At IFAD, a streamlined process exists to follow up on the recommendations and help close the learning loop (Figure 4). All recommendations of evaluations completed in a given year are tracked by IFAD management (PRISMA report) and, along with the proposed action plan, are presented to the Executive Board in the following year. IOE provides comments on the quality of follow-up, highlighting gaps, if required. Table 1 illustrates the IFAD Management tracking of recommendation follow-up for the period 2010-2018. It highlights that 65 per cent of the recommendations had been fully followed up, 25 per cent had an on-going follow-up and 3 per cent partial follow-up. Outright disagreement had happened in 1 per cent of cases. ▶▶

Table 1: Follow-up status of evaluation recommendations at IFAD (2010-2018)

	Full Follow-up	Not applicable*	Not agreed	Not yet due	Ongoing	Partial	Pending	Total
Number	1027	39	11	27	402	44	38	1588
%	65%	2%	1%	2%	25%	3%	2%	100%

* Not applicable recommendations relate to project investments that IFAD was initially envisaging but have not been pursued

Source: IFAD (2019)

► Reflections

It is necessary to move beyond formalism. The above general features and the specific examples from IFAD relate to a rather mature evaluation environment, where processes and roles are well defined. However, on its own, complying with the processes does not ensure the presence of a strong evaluation culture. To take an example, while management responses may show wide acceptance of evaluation recommendations, this refers to a future course of action and it cannot be taken for granted that the follow-up will happen as presented. In principle, this could be assessed through an independent review exercise looking retrospectively at the quality of evaluation follow-up. So far, there has not been a comprehensive retrospective assessment of the quality of follow-up to evaluations at IFAD.

As the conceptualization of Mayne (2010) illustrates, there is no substitute to the leadership commitment. This refers to the commitment of both the governing bodies (such as the Executive Board) and the management of the institution (i.e., both the principal and the agent) to create an operational culture that is nurtured by evidence on performance and results, where there is some tolerance for risk and errors, and a strong drive to make improvements based on evidence. Usefulness of evaluation ultimately depends on the organizational commitment to using findings for improved decision-making. In an institution like IFAD, assessment, self-reflection and continuous learning are key to address complex and dynamic challenges such as rural development, rural poverty reduction and rural transformation.

The presence of an independent evaluation function may not generate, on its own, an evaluation culture but can contribute to such culture through validation of evidence and analysis, generation of primary evidence, setting standards and conducting periodic evaluations of self-evaluation systems. eVALU

Endnotes

1. The authors are, respectively, Director and Deputy Director of the Independent Office of Evaluation of the International Fund for Agricultural Development (IFAD). The opinions expressed in this article are only those of the authors and do not necessarily reflect those of the organization for which they work. Contact author: f.felloni@ifad.org

2. Adapted from Mayne, 2010

3. Such problem arises when one stakeholder (the agent) is tasked to make decisions and choices representing another stakeholder (the principal), leading to risks of moral hazard and conflicts of interest.

4. OECD-DAC (2010); United Nations Evaluation Group (2016); Evaluation Cooperation Group (2012).

5. An early distinction between formative and summative evaluation was proposed by Scriven, M. (1967).

6. Prof. Robert Picciotto, personal communication.

7. Three scores for positive performance / results (6-highly satisfactory; 5-satisfactory; 4-moderately satisfactory) and three scores for negative performance / results (3-moderately unsatisfactory; 2-unsatisfactory; 1-highly unsatisfactory).

References

Evaluation Cooperation Group (2012) Big Book on Evaluation Good Practice Standards.

IEG. (2012). World Bank Group Impact Evaluations: Relevance and Effectiveness. Washington, DC: World Bank.

IFAD (2019). President's Report on the Implementation Status of Evaluation Recommendations and Management Actions (PRISMA). Volume I. Rome, Italy.

Laffont, Jean-Jacques; Martimort, David (2002). The theory of incentives: The principal-agent model. Princeton University Press.

Mayne, J. (2010). Building an Evaluative Culture: The Key to Effective Evaluation and Results Management. The Canadian Journal of Program Evaluation Vol. 24 No. 2 Pages 1–30.

OECD-DAC (2010). Evaluating Development Co-operation: Summary of key norms and standards, Paris

OVE (2017). IDB's Impact Evaluations: Production, Use and Influence. Office of Evaluation and Oversight, Interamerican Development Bank, Washington DC.

Ravasi, D.; Schultz, M. (2006). "Responding to organizational identity threats: Exploring the role of organizational culture". Academy of Management Journal.

Scriven, M. (1967). The methodology of evaluation. R.W. Tyler, R M. Gagne, M. Scriven (eds.), Perspectives of curriculum evaluation, pp.39-83. Chicago, IL: Rand McNally

United Nations Evaluation Group (2016). Norms and Standards for Evaluation. New York: UNEG.

Authors' profile

Oscar Garcia is the Director of the Independent Office of Evaluation at the International Fund for Agricultural Development (IFAD). Garcia is the Vice-Chair of the United Nations Evaluation Group, and he was the Chair of the Evaluation Cooperation Group of multilateral development banks in 2017. Before joining IFAD, Garcia served as head of the advisory services at UNEP - Technology, Industry and Economics Division, Paris, providing guidance to the Partnership for Action on Green Economy. Garcia holds a Master's degree in organizational change management, New School University, New York, and an MBA from the Bolivian Catholic University. He holds a BSc in Economics, University of Santa Catarina, Brazil. He authored several evaluation reports and papers on development evaluation.

Fabrizio Felloni is the Deputy Director of the Independent Office of Evaluation within the International Fund for Agricultural Development (IFAD). He was Lead Evaluation Officer at IFAD from 2010 to 2016. He was an Evaluation Specialist at the United Nations Development Programme from 2008 to 2010 where he prepared their first methodology for country-level evaluations. He holds a Master's degree in Agricultural Economics from Washington State University (USA) and a Master-equivalent degree in Social and Economic Sciences from Bocconi University (Italy).